Enjeux juridiques de l'utilisation de l'intelligence artificielle dans le secteur de l'électricité

Thomas Le Goff

Université Paris Cité

ED 262 - Sciences juridiques, politiques, économiques et de gestion

CEDAG EA 1516, Centre de droit des affaires et de gestion de l'Université Paris Cité

Enjeux juridiques de l'utilisation de l'intelligence artificielle dans le secteur de l'électricité

Par Thomas LE GOFF

Thèse de doctorat de Droit privé

Dirigée par Mme Nathalie MARTIAL-BRAZ

Présentée et soutenue publiquement le 24 mars 2023

Devant un jury composé de :

Mme Alexandra BENSAMOUN, Professeure à l'Université Paris-Saclay (examinatrice)
Mme Brunessen BERTRAND, Professeure à l'Université de Rennes 1 (rapporteure)
Mme Lucie CLUZEL-METAYER, Professeure à l'Université Paris Nanterre (présidente du jury)
Mme Nathalie MARTIAL-BRAZ, Professeure à l'Université Paris Cité (directrice de thèse)
M. Yannick MENECEUR, Maître de conférences associé à l'Université de Strasbourg (examinateur)
M. Samir MERABET, Professeur à l'Université des Antilles (rapporteur)

Titre : Enjeux juridiques de l'utilisation de l'IA dans le secteur de l'électricité

Résumé : Dans le secteur de l'énergie électrique, le recours à des systèmes d'intelligence artificielle (IA) permettrait de répondre à de nombreuses problématiques structurelles dans un contexte de transition écologique (prédiction de la production des énergies renouvelables et intermittentes, optimisation de la charge des véhicules électriques, modernisation et automatisation des réseaux de transport et de distribution...). Toutefois, leur développement apparaît contraint par le corpus juridique existant. Si certaines contraintes juridiques sont justifiées au regard des risques que peut générer le recours à l'IA, d'autres peuvent être levées à condition de bâtir un cadre de régulation adapté. La présente thèse analyse les obstacles résultant de l'application par défaut du corpus existant, qu'il convient d'adapter aux spécificités de l'IA, et propose des pistes concrètes pour construire un cadre juridique nouveau et nécessaire, conciliant promotion de l'innovation et prévention des risques. La réflexion menée sur la régulation de l'IA dans cette thèse, bien qu'abordée sous le prisme du secteur de l'énergie électrique, est transposable à d'autres secteurs hautement régulés. En effet, la démarche adoptée, visant à questionner les règlementations sectorielles à la lumière des apports potentiels de l'IA, peut être répliquée dans d'autres domaines tels que la finance, l'automobile ou l'aéronautique. De plus, certaines problématiques, notamment celles relatives à l'éthique ou à l'environnement, ne sont pas spécifiques au secteur étudié et peuvent donc être traitées de manière transversale.

Mots clefs : intelligence artificielle – énergie – électricité – régulation

***

Title : Legal issues raised by the use of AI in the electricity power sector

Abstract : In the electric energy sector, the use of artificial intelligence (AI) systems could address many structural issues in a context of ecological transition (prediction of renewable and intermittent energy production, optimization of electric vehicles charging, modernization and automation of transportation and distribution grids...). However, their development appears to be constrained by the existing legal framework. While some legal constraints are justified regarding the risks that the use of AI can generate, others could be lifted if an appropriate regulatory framework is built. This thesis analyzes the obstacles resulting from the default application of the existing body of laws, which should be adapted to the specificities of AI, and proposes concrete recommendations to build a new and necessary legal framework, reconciling the promotion of innovation and the prevention of risks. The reflection on the regulation of AI in this thesis, although approached under the prism of the electric power sector, is transposable to other highly regulated sectors. Indeed, the approach adopted, aiming at questioning the sectoral regulations in the light of the potential benefits of AI, can be replicated in other fields such as finance, automotive or aeronautics. Moreover, some issues, such as those related to ethics or the environment, are not specific to the sector studied and can therefore be addressed in a transversal manner.

Keywords : artificial intelligence – energy – electricity - regulation

REMERCIEMENTS

Bien qu'on dise d'elle qu'elle est un travail solitaire, la thèse est avant tout une formidable aventure humaine. On apprend, on rencontre, on s'y découvre, on endure, bref on grandit. C'est pourquoi je souhaite adresser mes pensées les plus chaleureuses à toutes les personnes qui ont rendu ce projet possible, qui m'ont aidé à lui donner vie ou qui m'ont soutenu dans sa réalisation.

Tout d'abord, je tiens à exprimer ma reconnaissance à Madame la Professeure Nathalie Martial-Braz pour la confiance accordée et l'opportunité donnée d'accomplir cette thèse.

Je remercie ensuite les membres du jury de thèse, Mesdames les Professeures Alexandra Bensamoun, Brunessen Bertrand, Lucie Cluzel-Métayer, Monsieur le Professeur Samir Merabet et Monsieur Yannick Meneceur pour leur temps et leur précieuse expertise.

Je souhaite également remercier tous mes collègues de la Direction Juridique d'EDF SA et en particulier Julie Amiot et Sabine Le Gac pour m'avoir donné la chance de réaliser cette thèse dans le cadre du dispositif CIFRE et pour m'avoir accompagné dans ce beau défi depuis sa genèse.

Enfin, j'adresse mes remerciements à mes proches, mes amis et ma famille pour leur soutien indéfectible tout au long de ce parcours. Mes pensées les plus tendres vont à Nina, que j'ai cessé de remercier tant aucun mot ne permet d'exprimer toute la reconnaissance que j'ai pour elle.

AVERTISSEMENT

La faculté et l'employeur de l'auteur n'entendent donner aucune approbation ni improbation aux opinions émises dans cette thèse ; ces opinions doivent être considérées comme propres à leur auteur.

LISTE DES PRINCIPALES ABRÉVIATIONS

ADEME	Agence de la transition écologique
aff.	Affaire
AI	*Artificial Intelligence*
AIE	Agence Internationale de l'Energie
AIEA	Agence Internationale de l'Energie Atomique
AIV	Activité d'importance vitale
AJDA	L'Actualité juridique : Droit administratif
al.	Alinéa
AMF	Autorité des Marchés Financiers
ANITI	*Artificial and Natural Intelligence Toulouse Institute*
art.	Article
ASN	Autorité de Sûreté Nucléaire
AWS	Amazon Web Services
CA	Cour d'appel
Cass. 1re, 2e, 3e civ., com., soc.	Cour de cassation, première, deuxième, troisième chambre civile, chambre commerciale, chambre sociale
CE	Conseil d'État
CEPD	Comité européen de la protection des données
CJCE	Cour de Justice des Communautés européennes
CJUE	Cour de Justice de l'Union européenne
CNIL	Commission nationale Informatique et Libertés
Coll.	Collection
CRE	Commission de régulation de l'énergie
D.	Revue du Recueil Dalloz
Dir.	Sous la direction de
Ed.	Edition
et al.	*et alius*
Fasc.	Fascicule
GAFAM	Google, Apple, Facebook, Amazon et Microsoft
GES	Gaz à effet de serre
IA	Intelligence artificielle
Ibid.	*Ibidem*
INB	Installation Nucléaire de Base
JCP	JurisClasseur périodique (Semaine juridique)
JORF	Journal officiel de la République française
JOCE	Journal officiel des communautés européennes
JOUE	Journal officiel de l'Union européenne
LGDJ	Librairie générale de droit et de jurisprudence
LIL	Loi Informatique et Libertés
LPA	Les petites affiches

N°	Numéro
Notamm.	Notamment
Obs.	Observation
OIV	Opérateur d'importance vitale
op. cit.	*Opus citatum*
p. ; pp.	Page ; pages
Prec.	Précité
Préf.	Préfacé par
PUF	Presses Universitaires de France
RGPD	Règlement général sur la protection des données
RLDI	Revue Lamy droit de l'immatériel
RTD	Revue trimestrielle de droit
s.	Suivants
SIIV	Système d'information d'importance vitale
ss.	Sous
TIPCE	Tribunal de première instance des communautés européennes
UE	Union européenne
V.	Voir
Vol.	Volume
VVQI	Vérification, validation et quantification des incertitudes

SOMMAIRE

INTRODUCTION GÉNÉRALE

« Un certain jour de marché, Xantus, qui avait dessein de régaler quelques-uns de ses amis, commanda [à Esope, son esclave] d'acheter ce qu'il y aurait de meilleur, et rien autre chose. Je t'apprendrai, dit en soi-même le Phrygien, à spécifier ce que tu souhaites, sans t'en remettre à la discrétion d'un esclave. Il n'acheta donc que des langues, lesquelles il fit accommoder à toutes les sauces ; l'entrée, le second, l'entremets, tout ne fut que langues. Les conviés louèrent d'abord le choix de ce mets ; à la fin ils s'en dégoûtèrent. Ne t'ai-je pas commandé, dit Xantus, d'acheter ce qu'il y aurait de meilleur ? Eh ! qu'y a-t-il de meilleur que la langue ? reprit Ésope. C'est le lien de la vie civile, la clef des sciences, l'organe de la vérité et de la raison : par elle on bâtit les villes et on les police ; on instruit, on persuade, on règne dans les assemblées, on s'acquitte du premier de tous les devoirs, qui est de louer les dieux. Eh bien ! dit Xantus (qui prétendait l'attraper), achète-moi demain ce qui est de pire : ces mêmes personnes viendront chez moi ; et je veux diversifier.

Le lendemain Ésope ne fit encore servir que le même mets, disant que la langue est la pire chose qui soit au monde : c'est la mère de tous débats, la nourrice des procès, la source des divisions et des guerres. Si on dit qu'elle est l'organe de la vérité, c'est aussi celui de l'erreur, et, qui pis est, de la calomnie. Par elle on détruit les villes, on persuade de méchantes choses. Si d'un côté elle loue les dieux, de l'autre elle profère des blasphèmes contre leur puissance. »

J. de La Fontaine, « La vie d'Esope le phrygien », in *Fables*, Bernardin-Béchet, Libraire-Éditeur, 1874, pp. 13-26.

1. **L'intelligence artificielle e(s)t la langue d'Esope.** L'intelligence artificielle (IA) est à la société moderne ce que la langue était à Esope le phrygien dans le mythe éponyme. Elle est

parfois présentée comme la « meilleure » des choses, la quintessence technologique[1], la « clef des sciences » et « l'organe de la raison ». Ses applications permettent, comme la langue, de « bâtir des villes » intelligentes[2], ou encore de les « policer »[3] de façon prédictive[4]. Toutefois, elle est, dans le même temps, la « pire », « mère de tous débats » philosophiques ou moraux[5] , « nourrice des procès » ou autres affaires médiatiques[6] et « source des divisions » entre ceux qui voient en elle la solution à tous les problèmes[7] et ceux, raisonnables, qui en perçoivent les risques[8]. Comme le soulève Frédéric Rouvière, « *si le mythe est le nom de tout ce qui n'existe et ne subsiste qu'ayant la parole pour cause, alors l'intelligence artificielle apparaît mythique à certains égards* »[9]. Toutefois, le caractère mythique de l'IA et les nombreuses controverses qui l'entourent ne doivent pas rebuter le juriste mais justifient au contraire l'intérêt qu'il faut lui porter. En effet, la force du mythe est de toujours comporter une part de vérité[10] et, une fois dépassé le stade de la science-fiction, doit venir le temps de la réflexion[11].

---

[1] CNRS, *Comment l'intelligence artificielle va changer nos vies*, dossier thématique, disponible en ligne : <https://lejournal.cnrs.fr/dossiers/comment-lintelligence-artificielle-va-changer-nos-vies>, consulté le 3 mars 2022 ; DGE, *Intelligence artificielle : État de l'art et perspectives pour la France*, rapport commandité par le Pôle interministériel de prospective et d'anticipation des mutations économiques, février 2019, 333 p.

[2] R. E. Hall, B. Bowerman, J. Braverman, *et al.*, « The vision of a smart city », *2nd International Life Extension Technology Workshop*, Paris, 28 septembre 2000 ; A. Camero, E. Alba, « Smart City and information technology: A review », *Cities*, 2019, vol. 93, pp. 84-94.

[3] L. Kalfon, « Des machines à gouverner ? », Le Monde, 11 octobre 1980, archives.

[4] B. Benbouzid, « Quand prédire, c'est gérer. La police prédictive aux États-Unis », *Réseaux*, 2018, vol.5, n° 211, pp. 221-256.

[5] E. Awad, S. Dsouza, R. Kim, *et al.*, « The Moral Machine experiment », *Nature*, 2018, 563, 59-64 ; E. M. Clare, « Moral Crumple Zones: Cautionary Tales in Human-Robot Interaction », *Engaging Science, Technology, and Society*, 23 mars 2019, n°5.

[6] D. Wakabayashi, « Self-Driving Uber Car Kills Pedestrian in Arizona, Where Robots Roam », *New York Times*, 19 mars 2018, disponible en ligne : <https://www.nytimes.com/2018/03/19/technology/uber-driverless-fatality.html>, consulté le 10 septembre 2019 ; The Guardian, *The Cambridge analytica files : a year long investigating into Facebook, data, and influencing elections in the digital age*, dossier thématique, 2018, disponible en ligne : <https://www.theguardian.com/news/series/cambridge-analytica-files>, consulté le 10 septembre 2019.

[7] D. Gronier, *L'intelligence artificielle en action : Santé, environnement, énergie... ce que l'IA change concrètement*, Eyrolles, 2020, pref. C. Villani, 220 p.

[8] G. Koenig, *La fin de l'individu : voyage d'un philosophe au pays de l'intelligence artificielle*, L'Observatoire, coll. De Facto, 189 p. ; E. Ntoutsi, P. Fafalios, U. Gadiraju, *et al.*, « Bias in data-driven artificial intelligence systems : an introductory survey », *WIREs Data Mining Knowlede Discovery*, décembre 2020, 10:1356.

[9] F. Rouvière, « L'intelligence artificielle au risque du mythe », *Revue Trimestrielle de Droit Civil*, Dalloz, 2020, p. 990. L'auteur fait ici référence à une formule de Paul Valéry dans P. Valéry, *Variétés II*, La *Petite Lettre sur les Mythes*, Paris, Gallimard, 1930, pp. 243-258.

[10] C. Atias, *Philosophie du droit*, PUF, 4ème éd., 2016, p. 331.

[11] A. Bensamoun, G. Loiseau, « L'intelligence artificielle : faut-il légiférer », *Recueil Dalloz*, 2017, p. 581.

2. **Plan de l'introduction.** Afin de donner les clés de compréhension au lecteur, il apparaît utile de commencer par identifier la « part de vérité » en définissant l'IA (**I**). Il faudra ensuite situer la présente étude dans son contexte doctrinal (**II**). Après cet effort de contextualisation, viendra le « temps de la réflexion » nécessitant de présenter le sujet de la thèse (**III**), à savoir les enjeux juridiques soulevés par l'utilisation de systèmes d'IA dans le secteur de l'électricité. Cet état des lieux permettra de problématiser le sujet (**IV**) avant de présenter la démonstration proposée (**V**).

I – La définition de l'IA

3. **Plan.** La compréhension des principaux éléments de définition technique de l'IA (**A**) est un préalable à sa qualification juridique (**B**).

A/ La définition technique de l'IA

4. **Origine et définition générale de l'IA.** Si la formule « intelligence artificielle » a été utilisée pour la première fois par John McCarthy en 1956 à la Conférence de Dartmouth[12], il faut remonter aux années 1940 pour trouver les origines de cette discipline dans les débuts de la « cybernétique »[13]. À cette date, les réflexions d'Alan Turing, célèbre pour avoir déchiffré le code Enigma utilisé par l'Allemagne nazie pendant la deuxième Guerre mondiale, visaient à faire reproduire des actions ou comportements humains par des machines[14]. Ces travaux ont donné naissance aux fondements de l'informatique telle qu'on la connait aujourd'hui. La formule « intelligence artificielle » est donc venue se greffer à un champ de recherche pluridisciplinaire déjà bien établi, mobilisant notamment mathématiques, robotique, mécanique et même philosophie. C'est pourquoi le Conseil de l'Europe la définit aujourd'hui comme

[12] Archives de la conférence disponibles en ligne : <http://raysolomonoff.com/dartmouth/>, consulté le 13 septembre 2020.
[13] M. Triclot, *Le moment cybernétique. La constitution de la notion d'information*, Champ Vallon, Paris, 2008, 384 p.
[14] A. Turing, « Proposal for Development in the Mathematics of an Automatic Computing Engine (ACE) », in *The Essential Turing: Seminal Writings in Computing, Logic, Philosophy, Artificial Intelligence, and And Artificial Life plus The Secret of Enigma*, dir. J. Copeland, Oxford University Press, New York, 2004.

« *l'ensemble de sciences, théories et techniques dont le but est de reproduire par une machine des capacités cognitives d'un être humain. Les développements actuels visent, par exemple, à pouvoir confier à une machine des tâches complexes auparavant réalisées par un humain* »[15].

5. **Les techniques « d'IA ».** Les techniques employées pour parvenir à l'objectif général de reproduction de l'intelligence humaine ont scindé la communauté scientifique en deux écoles : les « symbolistes » et les « connexionnistes », se distinguant par leur approche[16]. Si l'on cherche à classer les techniques utilisées par les spécialistes de l'IA, il est possible de différencier les techniques d'IA « symbolique », applications de règles strictes, n'ayant de différence avec les logiciels que nous connaissons que la complexité des problèmes résolus, et les techniques d'IA « connexionniste », fondées sur la mise en œuvre de réseaux de neurones artificiels au fonctionnement opaque ainsi que sur des méthodes dites « d'apprentissage automatique »[17]. Ces deux champs de techniques seront étudiés dans le cadre de la thèse, bien que le second nécessite une attention plus importante du fait de sa différence plus marquée avec les logiciels informatiques traditionnels.

6. **IA faible et IA forte.** La littérature fait également état d'une deuxième distinction, entre l'IA dite « faible » et l'IA dite « forte ». La première renvoie aux systèmes et machines effectuant des tâches précises, quand bien même leur fonctionnement serait fondé sur des techniques d'apprentissage[18]. La seconde, l'IA « forte », renverrait à des machines douées de conscience et de raison, capables de résoudre n'importe quel type de problème et non pas seulement des tâches spécifiques[19]. Ces machines pourraient fonctionner avec un niveau d'autonomie presque total, ne nécessitant plus d'intervention humaine pour les guider. Bien que certains des plus réputés chercheurs en IA poursuivent toujours cet objectif[20], il relève

---

[15] CONSEIL DE L'EUROPE, *Glossaire*, 2020, disponible en ligne : <https://www.coe.int/fr/web/artificial-intelligence/glossary>, consulté le 11 février 2020.
[16] D. Cardon, J.P. Cointet, A. Mazières, « La revanche des neurones : L'invention des machines inductives et la controverse de l'intelligence artificielle », *Réseaux*, 2018, n° 211, 5, 173.
[17] *Ibid.*
[18] R. Kurzweil, *The Singularity is Near: When Humans Transcend Biology*, Viking Penguin, New York, 2005, 602 p.
[19] C. Pennachin, B. Goertzel, « Contemporary Approaches to Artificial General Intelligence », in *Artificial General Intelligence*, dir. B. Goertzel, C. Pennachin, Springer, Berlin, 2007, pp. 1–30.
[20] Les travaux actuels de Yann LeCun, figure de l'apprentissage automatique et considéré comme le « père » des réseaux de neurones, visent à donner du « bon sens » à la machine, ce qui, selon lui, est le dernier verrou avant d'atteinte une IA « forte ». Voir notamment son interview : S. Revello, « Yann LeCun: 'Les machines manquent

aujourd'hui de la science-fiction. Puisqu'une IA « forte » n'existe pas encore, et que les spécialistes eux-mêmes doutent qu'elle n'existe un jour[21], elle ne sera pas traitée dans la thèse. En effet, la totalité des applications actuelles relèvent de l'IA « faible »[22], réalisant une tâche précise : diagnostic médical, jeu d'échec, pilotage autonome, analyse d'images, et bien d'autres.

7. **Le mythe de l'auto-apprentissage.** Une dernière précision doit être apportée sur les notions « d'apprentissage automatique »[23] et de « systèmes apprenants ». La première renvoie à un ensemble de techniques permettant à un système logiciel d'apprendre à résoudre un problème précis (classification d'images, traitement du langage…) à partir de données d'exemple ainsi que de réduire son taux d'erreur par l'expérience[24]. Cet apprentissage sur des données peut se faire de trois façons. Premièrement, il peut être « supervisé », auquel cas l'humain fournit des données labellisées à l'algorithme de telle manière que ce dernier dispose des couples « données d'entrée-résultat à obtenir ». Après l'apprentissage, l'algorithme sera capable de prédire des résultats à partir de nouvelles données d'entrée. Deuxièmement, l'apprentissage peut se faire de façon « non supervisée ». Dans ce cas, les données d'entrée d'apprentissage ne sont pas labellisées, l'algorithme identifiera les points communs entre celles-ci. Cette méthode permet notamment de construire des algorithmes de classification ou d'association. Troisièmement, l'apprentissage peut se faire « par renforcement », cas où l'algorithme apprend de son expérience pour atteindre un objectif fixé par l'humain, *via* un système de récompense et de pénalité[25]. Chacune de ces méthodes vise à concevoir un modèle d'IA, c'est-à-dire un algorithme entraîné, capable d'opérer sur de nouvelles données. Dans la pratique, il est rare que les modèles évoluent après leur phase d'apprentissage[26], contrairement à ce que laisse penser

---

de bon sens' », *Le Temps*, 4 octobre 2018, disponible en ligne : <https://www.letemps.ch/economie/yann-lecun-machines-manquent-sens>, consulté le 11 mars 2022.

[21] R. Fjelland, « Why general artificial intelligence will not be realized », *Humanities Social Sciences Communications*, 2020, vol. 7, n°10.

[22] C. Pennachin, B. Goertzel, *op. cit.*

[23] Traduction française de l'expression *« Machine learning »*.

[24] S.J. Russel, P. Novig, *Artificial Intelligence : a modern approach*, Pearson, 2020, 4th ed., 1152 p.

[25] *Ibid.*

[26] Certains systèmes conçus à partir de techniques d'apprentissage automatique peuvent évoluer au cours de leur utilisation mais, dans la plupart des cas, après un « réentraînement » de l'algorithme. Les seuls cas d'algorithmes apprenants sans supervision et ayant été déployés publiquement se sont d'ailleurs soldés par des échecs retentissants : D. Victor, « Microsoft Created a Twitter Bot to Learn From Users. It Quickly Became a Racist Jerk », *New York Times*, 24 mars 2016, disponible en ligne :

une partie de la doctrine faisant état de systèmes « auto-apprenants »[27]. Afin d'ancrer notre analyse dans une réalité technique et scientifique, les systèmes apprenant de façon autonome, sans supervision humaine et continue tout au long de leur cycle de vie ne seront pas étudiés.

8. **Transition.** Le champ de l'IA en tant que domaine de recherche technique est particulièrement large. Son étude juridique nécessite un effort de délimitation.

B/ La nécessaire définition juridique de l'IA

9. **Définir juridiquement l'IA : une tâche malaisée.** Selon le Professeur Charles Eisenmann[28], une définition n'est féconde et suffisante que si elle est suffisamment précise, univoque et qu'elle permet de distinguer ce qui est défini de tout autre élément. Cela revient à *« délimiter, […] situer et opposer pour individualiser »*[29]. Il peut donc paraître surprenant de trouver une pluralité de définitions de l'IA dans le champ juridique, qu'elles émanent d'autorités normatives[30] ou de la doctrine académique[31]. Pourtant, cette pluralité peut être justifiée par plusieurs raisons, notamment le degré d'abstraction du mot à définir, l'absence de définition en droit positif ou encore des divergences théoriques chez les rédacteurs[32]. La

<https://www.nytimes.com/2016/03/25/technology/microsoft-created-a-twitter-bot-to-learn-from-users-it-quickly-became-a-racist-jerk.html>, consulté le 10 septembre 2019.

[27] Sur les limites à l'apprentissage automatique après la phase de conception et dans la durée, voir notamm. : G.I. Parisi, R. Kemker, J.L. Part, *et al.*, « Continual lifelong learning with neural networks : A review », *Neural Networks*, 2019, vol. 113, pp. 54-71 ; Z. Chen, B. Liu, « Lifelong Machine Learning », *Synthesis Lectures on Artificial Intelligence and Machine Learning*, 2nd ed., août 2018, vol. 12, n°3, pp. 1-207.
Sur la diffusion de l'idée d'algorithmes auto-apprenants au cours de leur utilisation, y compris dans la doctrine juridique, voir notamm. : J. G. Ganascia, *Le mythe de la singularité. Faut-il craindre l'intelligence artificielle ?*, Seuil, 2017, p. 47, évoquant la possibilité donnée aux algorithmes d'apprentissage automatique de *« se reconfigurer en réécrivant leurs programmes »*.

[28] C. Eisenmann, « Quelques problèmes de méthodologie des définitions et des classifications en science juridique », *Archives de philosophie du droit*, 1966, vol. 11, p. 27-43, spec. p. 26.

[29] *Ibid.*, p. 30.

[30] Voir notamm. la définition de l'IA du Conseil de l'Europe, Supra, note 16 ; ISO, *Norme ISO/IEC 2382 :2015, Technologies de l'information*, Vocabulaire, Partie 28 : Intelligence artificielle - Notions fondamentales et systèmes experts : « *La capacité d'une unité fonctionnelle à exécuter des fonctions généralement associées à l'intelligence humaine, telles que le raisonnement et l'apprentissage* ».

[31] S. Merabet, *Vers un droit de l'intelligence artificielle*, Thèse de doctorat en droit privé, Dalloz, coll. Nouvelles bibliothèque de thèses, 2020, vol. 197, 1ère ed., 592 p., 119 ; A. Bensamoun, G. Loiseau, « L'intelligence artificielle : faut-il légiférer », *Recueil Dalloz*, 2017, p. 581.

[32] V. Champeil-Desplats, *Méthodologies du droit et des sciences du droit*, Dalloz, 2016, coll. Méthodes du droit, 2ème ed., 514–517.

doctrine académique cherche plutôt à définir une « notion cadre », s'intéressant moins à ce qu'est l'IA qu'à rendre la notion compatible avec les évolutions technologiques en la matière[33]. La démarche aboutit dans la thèse du Professeur Samir Merabet à la définition suivante : *« l'IA consiste en un système informatique doué de capacité cognitive lui permettant d'effectuer des choix de manière autonome, qui ne sont pas déterminés par la personne qui l'a conçu ou qui en a l'usage »*[34]. Si l'intention d'identifier et de mettre en exergue les éléments fondamentaux qui caractérisent l'IA est louable, il est regrettable que les juristes se refusent à adopter une approche plus technique du sujet. En effet, une acception trop large risque de conduire le juriste à se déconnecter de la réalité en s'intéressant à des applications relevant de la science-fiction ou d'englober des logiciels classiques. Sur un sujet où les divergences de vue sont nombreuses, il est important d'éviter l'introduction dans la définition de termes eux-mêmes non définis[35], tels que la « capacité cognitive » d'un système informatique.

10. **Le choix du terme à définir.** En premier lieu, il convient de choisir avec attention le terme à définir. L'expression « IA » renvoie à un champ de recherche pluridisciplinaire visant à confier à des machines des tâches complexes auparavant réalisées par un être humain. Ce sens fait désormais consensus[36] et il serait malheureux de vouloir y accoler un sens nouveau. Dès lors, si l'IA est un domaine de recherche, il convient de cibler l'objet précis auquel on entend s'intéresser : les produits de la recherche dans ce domaine, à savoir les « systèmes d'IA ». Eux-mêmes font l'objet de plusieurs définitions par des autorités normatives[37]. Ces définitions sont particulièrement larges, risquant elles aussi d'englober des systèmes logiciels classiques, et font encore l'objet de débats et évolutions. Ainsi, plutôt que de définir par référence, nous utiliserons dans le cadre de la thèse une définition dite « stipulative », consistant à utiliser un terme

[33] A. Bensamoun, G. Loiseau, « L'intelligence artificielle : faut-il légiférer », *Recueil Dalloz*, 2017, p. 581.
[34] S. Merabet, *op. cit.*, 119.
[35] R. Guastini, *Teoria del diritto. Approccio metodologico*, Mucchi, 2013, coll. Piccole conferenze, p. 46.
[36] Voir notamm. la définition retenue dans le manuel d'IA de S. J. Russel et P. Novig, considéré comme la référence en la matière : S. J. Russel et P. Novig, *Artificial Intelligence : a modern approach*, Pearson, 2020, 4th ed., 1152 p., spec. Preface (vii) : « *we define AI as the study of agents that receive percepts from the environment and perform actions* »
[37] Voir notamm. OCDE, *Recommandation du Conseil sur l'intelligence artificielle*, Recueil des instruments juridiques de l'OCDE, 2019, p. 8 ; COMMISSION EUROPÉENNE, *Proposition de règlement du Parlement européen et du Conseil établissant des règles harmonisées concernant l'intelligence artificielle (législation sur l'intelligence artificielle) et modifiant certains actes législatifs de l'Union*, 21 avril 2021, 2021/0106 (COD), article 3.

préexistant (les « systèmes d'IA ») d'une façon plus précise que son usage commun[38]. Ne cherchant pas à établir une vérité générale, la définition dans le cadre de la thèse doit permettre, dans une logique opératoire, d'en délimiter précisément le sujet, c'est-à-dire cibler les systèmes traités et les différencier d'autres objets, tels que les logiciels classiques.

11. **L'identification des spécificités de l'objet à définir.** En deuxième lieu, il convient d'identifier les caractéristiques particulières des systèmes d'IA. Pour Samir Merabet, il s'agissait principalement de la capacité à réaliser des choix et de l'autonomie des systèmes informatiques[39]. Plusieurs précisions techniques peuvent être apportées pour délimiter la notion et la distinguer d'autres systèmes informatiques utilisés depuis des décennies. Le « plus petit dénominateur commun »[40] à tous les systèmes d'IA peut être trouvé dans leur capacité à générer des prédictions, recommandations ou des décisions, de façon plus ou moins autonome, et ce à partir de règles fixées par l'humain ou d'un processus d'apprentissage sur des données existantes[41]. Le croisement des travaux de groupes d'experts au niveau européen[42] et de la littérature scientifique sur l'IA[43] permet de mettre en évidence quatre caractéristiques qui peuvent être utilisées comme faisceau d'indices pour différencier un système d'IA d'un système informatique classique.

---

[38] V. Champeil-Desplats, *op. cit.*, 497.
[39] S. Merabet, *op. cit.*, 119.
[40] A. Bensamoun, G. Loiseau, « L'intégration de l'intelligence artificielle dans l'ordre juridique en droit commun : questions de temps », *Dalloz IP/IT*, 2017, 239.
[41] S. Merabet, *op. cit.*, 119 ; repris en partie dans la proposition de règlement sur l'IA du 21 avril 2021 par la Commission européene : COMMISSION EUROPÉENNE, *Proposition de règlement du Parlement européen et du Conseil établissant des règles harmonisées concernant l'intelligence artificielle (législation sur l'intelligence artificielle) et modifiant certains actes législatifs de l'Union*, 21 avril 2021, 2021/0106 (COD), article 3.
[42] EXPERT GROUP ON LIABILITY AND NEW TECHNOLOGIES, *Liability for AI and other emerging digital technologies*, 21 novembre 2019, spec. p. 32 et s., disponible en ligne : <https://op.europa.eu/en/publication-detail/-/publication/1c5e30be-1197-11ea-8c1f-01aa75ed71a1/language-en/format-PDF>, consulté le 18 janvier 2020 ; GROUPE D'EXPERTS DE HAUT NIVEAU EN IA, *A Definition of AI : Main capabilities and disciplines*, Rapport du groupe d'experts indépendants établis par la Commission européenne, 8 avril 2019, disponible en ligne : <https://digital-strategy.ec.europa.eu/en/library/definition-artificial-intelligence-main-capabilities-and-scientific-disciplines>, consulté le 18 janvier 2020.
[43] V. notamm. D. Cardon, J.-P. Cointet, A. Mazières. « La revanche des neurones : L'invention des machines inductives et la controverse de l'intelligence artificielle », *Réseaux*, 2018, n° 211, 5, 173 ; J. McCarthy, « What is artificial intelligence ? », 24 novembre 2004, disponible en ligne : <https://homes.di.unimi.it/borghese/Teaching/AdvancedIntelligentSystems/Old/IntelligentSystems_2008_2009/Old/IntelligentSystems_2005_2006/Documents/Symbolic/04_McCarthy_whatisai.pdf>, consulté le 22 mai 2021 ; S. J. Russel et P. Novig, *op. cit.*

12. **La complexité structurelle des systèmes d'IA.** La première caractéristique est leur complexité structurelle, car certains systèmes combinent une composante logicielle et une ou des composantes matérielles[44]. Cette combinaison n'est pas nouvelle mais les systèmes d'IA peuvent présenter un niveau de complexité plus important que de simples objets connectés. Le cas-échéant, les composantes matérielles du système d'IA lui permettent de « percevoir » son environnement[45] et « d'agir » sur lui[46]. Les interactions potentielles entre les couches logicielles et matérielles sont nombreuses, d'où la complexité de la structure globale.

13. **L'opacité du fonctionnement des systèmes d'IA.** La deuxième caractéristique des systèmes d'IA à laquelle la thèse entend s'intéresser est l'opacité du fonctionnement, définie comme l'impossibilité de comprendre et d'expliquer de façon rationnelle et formelle la logique de la prise de décision d'un système d'IA. Elle est inhérente aux techniques utilisées et prépondérante dans le cas de l'apprentissage automatique dit « profond »[47].

14. **La potentielle autonomie des systèmes d'IA.** Le troisième attribut des systèmes d'IA est leur potentielle autonomie, justement identifiée et étudiée par Samir Merabet qui en présente les différentes échelles[48]. L'autonomie est potentielle dans la mesure où tous les systèmes d'IA ne fonctionnent pas sans intervention humaine, notamment dans le cas de systèmes d'aide à la décision où le résultat fourni ne constitue qu'un des éléments pris en compte par l'humain pour prendre une décision.

15. **La dépendance à la donnée des systèmes d'IA.** La quatrième et dernière caractéristique pouvant définir les systèmes d'IA est leur dépendance à la donnée. Cette dépendance peut concerner toutes les phases du cycle de vie du système, de l'apprentissage au fonctionnement, en passant par le paramétrage et la validation[49]. Les données peuvent provenir de différentes sources. Elles peuvent être directement renseignées par un humain qui les aurait collectées et

[44] Note : un système d'IA peut ne pas être incorporé dans une enveloppe matérielle.
[45] Par exemple par des capteurs.
[46] Par exemple dans le cas d'un équipement piloté par la composante logicielle du système d'IA, tel qu'un véhicule autonome.
[47] Traduction de l'anglais « Deep learning », V. notamm. F. Pasquale, *The Black Box Society: The Secret Algorithms That Control Money and Information*, Harvard University Press, Cambridge, 2015, 260 p, spec. pp. 189-218.
[48] S. Merabet, *op. cit.*, 89 et s.
[49] EXPERT GROUP ON LIABILITY AND NEW TECHNOLOGIES, *op. cit.*, spec. p. 33.

préparées en amont, générées par des capteurs reliés au système ou communiquées par un autre système.

16. **La définition retenue dans la thèse.** À partir des caractéristiques identifiées, il est possible de circonscrire l'objet de la thèse : les « systèmes d'IA », définis comme *des systèmes logiciels, pris isolément ou incorporés à un produit, ayant la capacité de générer, à partir de données d'entrée, des prédictions, recommandations ou des décisions susceptibles d'influencer leur environnement. Ils sont conçus à partir de règles fixées par l'humain ou d'un processus d'apprentissage automatique sur des données. Selon les techniques employées, ces systèmes peuvent fonctionner à différents niveaux d'autonomie et de façon plus ou moins opaque*. Par souci de simplicité et de fluidité, les expressions « une IA » ou « l'IA » renverront respectivement, dans la suite des développements, à « un système d'IA » ou « les systèmes d'IA pris dans leur ensemble », sauf s'il est précisé autrement.

17. **Transition.** Les systèmes d'IA tels que nous venons de les définir ont déjà fait l'objet de nombreuses études dans la doctrine juridique. La présente thèse s'inscrit donc dans un contexte doctrinal plus large qu'il convient de présenter.

II – Le contexte doctrinal de la recherche : la réflexion sur l'intégration de l'IA dans l'ordre juridique

18. **Le rôle du juriste dans un contexte d'innovation technologique.** L'intégration d'un nouvel objet aux caractéristiques inédites dans le système juridique en place peut menacer son équilibre. Cet équilibre, fragile, dépend en grande partie de la capacité du juriste – et encore plus du chercheur en droit – à comprendre le nouvel objet auquel il fait face, son articulation avec les règles de droit existantes et, le cas-échéant, à penser leur adaptation.

19. **Des travaux précurseurs relatifs à l'informatique avancée.** L'histoire de l'IA n'étant que la continuité de celles de la cybernétique et de l'informatique, il convient de mentionner l'existence de travaux précurseurs, preuves que le sujet n'est finalement pas nouveau dans la sphère juridique. De premières références peuvent déjà être trouvées dans les travaux préparatoires de la loi Informatique et Libertés du 6 janvier 1978, où le rapporteur Jean Foyer fait le constat que *« l'informatique a permis d'accomplir des opérations qu'on avait cru*

jusqu'alors réservées à l'intelligence humaine »[50] en évoquant une *« machinerie, elle-même inintelligente »*[51]. La similitude de la formule avec la définition de l'IA en tant que champ de recherche est frappante. Du côté de la doctrine académique, c'est également à partir des années 1970 que l'on assiste à l'émergence d'ouvrages et articles traitant des liens entre les développements de l'informatique de l'époque, visant donc à reproduire l'intelligence humaine, et le Droit[52]. Un peu plus tard, ce sont les systèmes informatiques de prise de décision et les systèmes dits « experts » qui mobiliseront les juristes[53].

20. **Les premières références juridiques à l'IA.** En matière d'IA, la jurisprudence est particulièrement pauvre. On n'y trouvera que des références aux « robots » ou aux « algorithmes », références alternatives mais non suffisantes[54]. Nos principales sources sont donc à trouver dans la doctrine académique. Le premier ouvrage juridique français véritablement dédié à l'IA date de 1994[55], avec un temps de retard sur la doctrine anglo-saxonne elle-même foisonnante en la matière[56]. Après une période pauvre en publications, ce n'est qu'à partir de 2016 que l'on peut observer un regain d'intérêt de la doctrine juridique sur le sujet de l'IA. Bien qu'encore peu nombreuses, les thèses sur le droit de l'IA se multiplient également[57].

---

[50] ASSEMBLÉE NATIONALE, *Archives des débats parlementaires sur le projet de loi Informatique et Libertés du 4 octobre 1977*, publiées au JORF n°79 A.N. du 5 octobre 1977.
[51] *Ibid.*
[52] Voir notamm. : X. Linant de Bellefonds, *L'informatique et le Droit*, PUF, coll. Que sais-je ?, Paris, 1981, 127 p. ; J.P. Buffelan, « Le droit, l'informatique et la mathématique », *Journal de la société statistique de Paris*, 1974, tome 115, pp. 301-316 ; R. Laperrière, « L'informatique et les droits des personnes », *Cahiers de recherche sociologique*, 1993, n°21, pp. 53–77.
[53] D. Bourcier, *La décision artificielle*, PUF, coll. Les voies du droit, Paris, 1995, 240 p. ; M. Vivant, « Le droit des systèmes-experts », in *Congrès sur l'information et la documentation*, 1987, 7, pp. 159-163 ; A. Bertrand, « L'intelligence artificielle, la robotique, les systèmes experts et le droit », *Expertises*, 1987, n°96, pp. 219-224 ; M. Schauss, « Systèmes experts et droit », *Revue interdisciplinaire d'études juridiques*, 1987, vol. 18, n°1, pp. 101-114.
[54] Références étudiées notamm. par S. Merabet, *op. cit.*, spec. pts 52 et s.
[55] D. Boursier, P. Asset, C. Roquilly, *Le droit et l'intelligence artificielle : une Révolution de la connaissance juridique*, Romillat, Paris, 1994, 304 p.
[56] S. N. Lehman-Wilzig, « Frankenstein unbound: Towards a legal definition of artificial intelligence », *Futures*, 1981, vol. 13, n°6, pp. 442-457 ; P. McNally, S. Inayatullah, « The rights of robots: Technology, culture and law in the 21st century », *Futures*, 1988, Vol. 20, n°2, pp. 119-136 ; M. Gemignani, « Laying Down the Law to Robots », *San Diego Law Review*, 1984, vol. 21, 1045.
[57] S. Merabet, *op. cit.* ; G. Guegan, *L'élévation des robots à la vie juridique*, Thèse pour le doctorat en droit privé, Université Toulouse 1 Capitole, 2016, 368 p. ; J. Pouget, *La réparation du dommage impliquant une intelligence artificielle*, Thèse pour le doctorat en droit privé, Université d'Aix Marseille, 2019, 410 p.

L'ensemble de ces travaux visent à définir l'IA[58], la qualifier juridiquement[59], déterminer si les régimes juridiques actuels s'y appliquent[60] ou encore étudier l'opportunité de créer une règlementation *ad hoc*[61].

21. **L'opposition doctrinale sur les enjeux juridiques du développement de l'IA.** Certains considèrent que la « *ringardisation annoncée des régimes juridiques par le numérique* »[62] est en partie illusoire et que le corpus juridique existant est suffisant pour encadrer l'IA. D'autres, d'un courant doctrinal parfaitement opposé, reconnaissent les limites des règles existantes[63] et plaident pour leur adaptation, voire pour la création d'un nouveau « Droit »[64]. Si le débat sur l'opportunité d'édicter des règles spéciales pour l'IA est toujours ouvert malgré la prise de position des institutions européennes[65], la nécessité d'adapter certaines règles existantes commence à faire consensus[66].

22. **Contenu et nature des études juridiques sur l'IA.** Parmi les régimes juridiques dont l'application à l'IA fait débat, on trouve principalement ceux relatifs à la responsabilité civile,

---

[58] S. Merabet, *op. cit.*, spec. pts 22 et s.

[59] Sur l'opportunité de qualifier l'IA de sujet de droit et donc de lui reconnaître une personnalité juridique : D. Bourcier, « De l'intelligence artificielle à la personne virtuelle : émergence d'une entité juridique ? », *Droit et société*, 2001, vol. 49, n°3, pp. 847-871 ; A. Bensamoun, « Des robots et du droit... », *Dalloz IP/IT*, 2016, 281 ; J. R. Binet, « Personnalité juridique des robots : une voie à ne pas suivre », *Revue Droit de la famille*, LexisNexis, 2017, n°6, p. 2.

[60] Voir notamm. l'ouvrage collectif : A. Bensamoun, G. Loiseau, *Droit de l'intelligence artificielle*, LGDJ, coll. Les intégrales, 2019, 444 p.

[61] C. Castets-Renard, « Comment construire une intelligence artificielle responsable et inclusive ? », *Recueil Dalloz*, 2020, p. 225.

[62] V.-L. Benabou, « Un droit vivant. Manifeste pour des juristes incarnés et sensibles à l'heure de l'intelligence artificielle », *Mélanges Michel Vivant*, 2020, p. 715-730.

[63] A. Bensamoun, G. Loiseau. « L'intégration de l'intelligence artificielle dans l'ordre juridique en droit commun : questions de temps », Dalloz IP/IT, 2017, 239 ; Y. Meneceur, « DataJust face aux défis de l'intelligence artificielle », *JCP*, 2020, 1087 ; X. Henry, « Le renouvellement de la jurisprudence. À propos du site de Cerclab », *JCP*, 2020, 938.

[64] Y. Meneceur, *L'intelligence artificielle en procès : Plaidoyer pour une règlementation internationale et européenne*, Bruylant, Paris, 2020, 209 p.

[65] Voir par exemple la proposition de règlement européen sur l'IA par la Commission européenne le 21 avril 2021 : COMMISSION EUROPÉENNE, *Proposition de règlement du Parlement européen et du Conseil établissant des règles harmonisées concernant l'intelligence artificielle (législation sur l'intelligence artificielle) et modifiant certains actes législatifs de l'Union*, 21 avril 2021, 2021/0106 (COD)..

[66] A. Bensamoun, G. Loiseau, « L'intégration de l'intelligence artificielle dans l'ordre juridique en droit commun : questions de temps », *op. cit.* ; COUR D'APPEL DE PARIS, *La réforme du droit français de la responsabilité civile et les relations économiques*, Rapport, avril 2019, spec. p. 107 et s., disponible en ligne : <http://www.justice.gouv.fr/art_pix/Rapport_CA_PARIS_reforme_responsabilite_civile.pdf>, consulté le 20 mai 2021.

à la protection des données personnelles et à la propriété intellectuelle[67]. Les débats sur la propriété intellectuelle portent exclusivement sur la protection des créations de l'IA et non du système en lui-même[68]. Les applications de l'IA étudiées dans la thèse n'ayant pas vocation à « créer », cette problématique ne sera pas adressée dans les développements. On notera également un parent pauvre de la doctrine juridique en matière d'IA : les approches sectorielles ou études de cas. Seules les applications de l'IA dans les secteurs de la santé[69], de la banque[70], de la finance[71] et des assurances[72] ont fait l'objet de publications juridiques. Pourtant, ces études sectorielles sont utiles pour ancrer la pensée juridique dans des considérations pratiques et, le cas-échéant, pour intégrer ces dernières dans une réflexion plus globale sur l'opportunité de créer de nouvelles règles *ad hoc*. Sur un sujet aussi technique que celui de l'IA, il apparaît opportun de compléter les approches abstraites, majoritaires en doctrine, par l'approche sectorielle permettant un raisonnement inductif, partant d'observations réelles pour aboutir à des théories de portée générale[73].

23. **Transition.** Notre recherche s'inscrit dans cette démarche en s'intéressant aux enjeux juridiques liées à l'utilisation de l'IA dans le secteur de l'électricité.

[67] G. Courtois, J. S. Mariez, « Intelligence artificielle : quels objets de droit pour quel encadrement contractuel ? », *RLDA*, septembre 2019, n°151, p. 22, 6781.
[68] J. Larrieu, « Robot et propriété intellectuelle », *Dalloz IP/IT*, 2016, 291 ; A. Bensamoun, « Intelligence artificielle et propriété intellectuelle », in *Droit de l'intelligence artificielle*, dir. A. Bensamoun, G. Loiseau, LGDJ, 2019, n°397 et s. ; I. Randrianirina, « Plaidoyer pour un nouveau droit de propriété intellectuelle sur les productions générées par intelligence artificielle », *Recueil Dalloz*, 2021, p. 91.
[69] D. Gruson, « Enjeux juridiques de l'intelligence artificielle en santé : le stable et le mouvant », *Revues des Juristes de Sciences Po*, juin 2016, n° 21, 16.
[70] N. Martial-Braz, « L'apport de l'intelligence artificielle à la banque », *Revue de Droit bancaire et financier*, novembre 2019, n°6, comm. 78.
[71] T. de Ravel d'Esclapon, « La gouvernance des algorithmes dans le secteur financier : le point de vue de l'ACPR », *Revue de Droit bancaire et financier*, juillet 2020, n° 4, étude 12 ; N. Martial-Braz, « L'intelligence artificielle bientôt régulée. L'incidence du AI Act dans le secteur financier », *Revue de Droit bancaire et financier*, mai 2021, n°3, comm. 81.
[72] L. Grynbaum, « IA et assurance », *RLDA*, septembre 2019, n°151, p. 22, 6782.
[73] C. Perelman, L. Obrechts-Tyteca, *Traité de l'argumentation*, Editions de l'université de Bruxelles, 2008, 6ème ed., 740 p. ; voir aussi M.L. Mathieu, *Logique et raisonnement juridique*, PUF, Paris, 2015, coll. Themis 2ème ed., 446 p.

III - Le sujet de la recherche : les enjeux juridiques liés à l'utilisation des systèmes d'IA dans le secteur de l'électricité

24. **La diversité des cas d'usage.** Les applications concrètes de l'IA sont nombreuses et de natures très diverses. Les briques technologiques sous-jacentes aux systèmes vont du traitement automatique du langage (agents conversationnels, reconnaissance vocale, extraction de données dans des documents) à la reconnaissance d'images, en passant par l'aide à la décision (systèmes de recommandation ou de prévision)[74]. L'utilisation de ces briques ou leur combinaison peuvent conduire à des applications telles que des systèmes de pilotage automatique, de maintenance prédictive, ou encore de trading algorithmique[75]. De ces nouvelles applications, la société civile et les industriels attendent légitimement de potentiels bénéfices tels qu'une réduction des coûts, une augmentation de la sécurité *via* la suppression de l'aléa humain ou plus globalement l'ouverture de nouvelles opportunités d'affaires[76]. Il n'est donc pas surprenant de constater que de nombreux secteurs d'activité y prêtent une grande attention, et en particulier dans les grandes industries : l'automobile, la santé, la finance ou encore l'énergie[77].

25. **Plan.** La présente thèse, réalisée dans le cadre d'une Convention Industrielle de Formation par la Recherche (CIFRE) et financée par EDF S.A., propose une approche sectorielle des enjeux juridiques soulevés par l'utilisation de systèmes d'IA, à travers le prisme du secteur de l'électricité. Les cas d'usage de systèmes d'IA spécifiques à ce secteur n'ont pour l'heure fait l'objet d'aucune étude juridique en France[78]. Pourtant, les énergéticiens sont particulièrement actifs sur le sujet et ces applications se multiplient. Leur étude nécessite en

---

[74] GROUPE D'EXPERTS DE HAUT NIVEAU EN IA, *A Definition of AI : Main capabilities and disciplines*, *op. cit.*, p. 2 et s.

[75] D. Gronier, *L'intelligence artificielle en action : Santé, environnement, énergie... ce que l'IA change concrètement*, Eyrolles, 2020, pref. C. Villani, 220 p.

[76] C. Villani, *Donner un sens à l'intelligence artificielle*, Rapport dans le cadre d'une mission parlementaire du 8 septembre 2017 au 8 mars 2018 confiée par le Premier Ministre Edouard Philippe, *La Documentation Française*, 8 mars 2018, p. 184 et s.

[77] Pour une analyse sectorielle de l'impact économique du développement de l'IA, voir DGE, *Intelligence artificielle : État de l'art et perspectives pour la France*, rapport commandité par le Pôle interministériel de prospective et d'anticipation des mutations économiques, février 2019, 333 p. ; D. Gronier, *op. cit.*

[78] Au niveau international, on trouvera quelques publications traitant des conséquences juridiques de l'utilisation de systèmes d'IA dans l'énergie. Voir par exemple : A.L. Stein, « Artificial Intelligence and Climate Change », *Yale Journal on Regulation*, 2020, 37, n°890 ; M. Mylrea, « Smart energy-internet-of-things opportunities require smart treatment of legal, privacy and cybersecurity challenges », *The Journal of World Energy Law & Business*, avril 2017, vol. 10, n°2, pp. 147–158.

premier lieu de comprendre les activités constituant l'industrie électrique et leurs enjeux (**A**) afin de pouvoir, en deuxième lieu, présenter une typologie des cas d'usage de systèmes d'IA dans ce secteur (**B**). À partir de cette typologie, fruit d'un travail de terrain en entreprise constitué de recherches documentaires, d'observations et d'entretiens, il sera possible de mettre en lumière les principaux enjeux juridiques soulevés par l'IA dans le secteur de l'électricité (**C**).

A/ Les enjeux du secteur de l'électricité dans un contexte de transition écologique

26. **Présentation du secteur de l'électricité et de ses composantes.** L'importance de l'approvisionnement en électricité dans un pays n'est plus à démontrer. Le préambule du contrat de service public entre l'État français et EDF S.A. de 2005 stipule ainsi que *« l'électricité est au cœur de la vie quotidienne, elle est indispensable à toute économie développée. Bien de première nécessité, énergie stratégique, l'électricité n'est pas un bien comme les autres, ce qui confère aux entreprises œuvrant dans ce secteur un niveau élevé de responsabilité. L'électricité concourt en effet à la cohésion sociale, au développement équilibré du territoire, à la recherche et au progrès technologique »*[79]. Elle intervient dans la production de la quasi-totalité des biens et services et elle est aussi, en grande quantité, consommée directement par les ménages. Le secteur de l'électricité peut être défini comme l'ensemble des activités nécessaires à l'approvisionnement en électricité d'un pays. Il est subdivisé en quatre fonctions : la production, le transport, la distribution et la fourniture[80]. Ces activités font l'objet de règles

[79] Contrat de service public entre l'État français et EDF, 2005, disponible en ligne : <https://cpdp.debatpublic.fr/cpdp-epr/docs/pdf/dossier_mo/contrat-service-public-etat-edf-oct2005.pdf>, consulté le 30 mars 2022.

[80] F. Steiner, « L'industrie de l'électricité : réglementation, structure du marché et performances », *Revue économique de l'OCDE*, 2001, vol. nº32, n°1, pp. 159–201.

spécifiques issues du Code de l'énergie[81], résultant pour la majorité de la transposition de différentes directives européennes relatives à la régulation du secteur de l'électricité[82].

27. **La production d'électricité.** La production consiste en la transformation d'une forme d'énergie en électricité. Elle regroupe l'ensemble des différents moyens de produire de l'électricité à partir de gaz naturel ou de charbon (centrales dites « thermiques »), de l'énergie nucléaire (centrales nucléaires de production d'électricité ou « CNPE »), de l'énergie hydraulique (barrages hydroélectriques), de combustibles renouvelables (centrales dites « biomasse »), ainsi qu'à partir des énergies éolienne et solaire[83]. L'activité de production est ouverte à la concurrence depuis la loi du 10 février 2000 relative à la modernisation et au développement du service public de l'électricité[84], tout en étant encadrée par des règles générales ou spécifiques au moyen de production concerné, figurant aux articles L311-1 à L315-9 du Code de l'énergie.

28. **Le transport d'électricité.** Le transport recouvre l'acheminement de l'électricité sur le réseau à haute ou à très haute tension depuis les moyens de production vers les réseaux de distribution[85]. Cette activité comporte également la gestion du réseau connectant l'ensemble des moyens de production décentralisés de façon à maintenir l'équilibre entre l'électricité injectée sur le réseau et l'électricité finalement consommée. Un déséquilibre entre ces deux valeurs conduirait à des situations de surtension pouvant endommager les infrastructures ou de sous-tension causant des coupures de courant. L'activité de transport est régulée conformément

---

[81] Code de l'énergie, article L111-1 : *« Les secteurs de l'électricité et du gaz distinguent, notamment, quatre activités obéissant à des règles d'organisation et soumises à des obligations différentes. Les activités d'exploitation des réseaux publics de transport et de distribution d'électricité ainsi que d'exploitation des réseaux de transport et des réseaux publics de distribution de gaz naturel sont régulées conformément aux dispositions du présent livre. Les activités de production et de vente aux consommateurs finals ou fourniture s'exercent au sein de marchés concurrentiels sous réserve des obligations de service public énoncées au présent livre et des dispositions des livres III et IV »*.

[82] V. notamm. *Directive (UE) 2019/944 du Parlement européen et du Conseil du 5 juin 2019 concernant des règles communes pour le marché intérieur de l'électricité et modifiant la directive 2012/27/UE*, publiée au JOUE n°l158/125 du 14 juin 2019.

[83] *Ibid.* ; MINISTÈRE DE LA TRANSITION ÉCOLOGIQUE, *La production d'électricité*, site officiel, 17 février 2017, disponible en ligne : <https://www.ecologie.gouv.fr/production-delectricite>, consulté le 30 mars 2022.

[84] *Loi n°2000-108 du 10 février 2000 relative à la modernisation et au développement du service public de l'électricité*, publiée au JORF n°35 du 11 février 2000, texte n°1 ; Code de l'énergie, article L111-1.

[85] *Directive (UE) 2019/944 du Parlement européen et du Conseil du 5 juin 2019 concernant des règles communes pour le marché intérieur de l'électricité et modifiant la directive 2012/27/UE*, publiée au JOUE n°l158/125 du 14 juin 2019, article 2, pt 34.

aux dispositions du Livre premier du Code de l'énergie[86]. À ce titre, elle est exercée en France par RTE qui est donc responsable *« de l'exploitation, de la maintenance et du développement du réseau de transport […], de ses interconnexions avec d'autres réseaux et chargé de garantir la capacité à long terme du réseau à satisfaire une demande raisonnable de transport d'électricité »*[87].

29. **La distribution d'électricité.** Cette activité désigne l'acheminement de l'électricité sur des réseaux de distribution à haute, à moyenne et à basse tension aux fins de fourniture à des consommateurs finals[88]. Contrairement au transport, la distribution vise à acheminer de l'électricité en moins grande quantité et sur de courtes distances. Cette activité fait également partie du domaine « régulé », exercé par un gestionnaire désigné par les autorités publiques[89]. En France, Enedis (ex-ERDF) est le principal gestionnaire du réseau public d'électricité. C'est à ce titre que l'entreprise est en charge des compteurs électriques et a déployé, aux fins de modernisation, les compteurs communicants Linky. Elle est soumise à d'importantes règles d'indépendance et de non-discrimination envers l'ensemble des fournisseurs d'électricité malgré sa dépendance capitalistique à EDF S.A.

30. **La fourniture d'électricité.** Enfin, la fourniture désigne la vente d'électricité aux consommateurs finals. Y sont inclues toutes les activités de commercialisation de biens et services relatifs à l'électricité, par exemple les services de conseil en économie d'énergie. Ces activités de « commercialisation » font partie du domaine « dérégulé » et sont donc ouvertes à la concurrence. En tant qu'acteur historique, EDF S.A. est en position dominante sur le marché de l'électricité. Plusieurs mécanismes ont été créés par la loi pour favoriser le développement

[86] Voir les articles L111-2 à L111-50 du Code de l'énergie pour les dispositions d'ordre général, relatives notamment aux modalités de désignation du gestionnaire de réseau de transport et aux règles d'indépendance, et les articles L321-1 à L321-19 pour les dispositions spécifiques au gestionnaire du réseau de transport d'électricité, relatives notamment à ses missions telles que l'exploitation et la maintenance du réseau.

[87] *Directive (UE) 2019/944 du Parlement européen et du Conseil du 5 juin 2019 concernant des règles communes pour le marché intérieur de l'électricité et modifiant la directive 2012/27/UE*, publiée au JOUE n°l158/125 du 14 juin 2019, article 2, pt 35 ; transposé à l'article L321-6 du Code de l'énergie.

[88] *Directive (UE) 2019/944 du Parlement européen et du Conseil du 5 juin 2019 concernant des règles communes pour le marché intérieur de l'électricité et modifiant la directive 2012/27/UE*, publiée au JOUE n°l158/125 du 14 juin 2019, article 2, pt 28.

[89] Voir les dispositions générales sur l'organisation du gestionnaire du réseau public de distribution d'électricité aux articles L111-51 à L111-66 du Code de l'énergie, et les dispositions spécifiques relatives aux missions du gestionnaire aux articles L322-1 à L322-12.

de la concurrence, tels que l'Accès Régulé à l'Electricité Nucléaire Historique[90] permettant aux fournisseurs alternatifs d'accéder, à un prix plafonné, à l'électricité produite par les centrales nucléaires historiques d'EDF, dans la limite de 100 TWh par an, soit environ 25 % de la production.

31. **Les enjeux du secteur de l'électricité.** L'ensemble des activités présentées dans les paragraphes précédents constituent le secteur de l'électricité, dont il faut comprendre les grands enjeux avant de pouvoir apprécier en quoi l'utilisation de systèmes d'IA peut lui être utile.

32. **La place centrale de l'électricité dans l'atteinte de la neutralité carbone.** Tout d'abord, le secteur de l'électricité relève d'une importance majeure dans un contexte de transition écologique. L'atteinte des objectifs climatiques fixés par l'Accord de Paris[91] et la nécessaire décarbonation de l'économie le mettent doublement à contribution : d'une part, l'électrification d'industries fortement émettrices de CO2 (tels que le passage des véhicules thermiques à l'électrique) va conduire à une hausse de la demande, et, d'autre part, les moyens de produire de l'électricité doivent évoluer pour abandonner le recours aux énergies fossiles[92]. Dans le cadre de sa Stratégie Nationale Bas Carbone (SNBC)[93] et de la Programmation Pluriannuelle de l'Energie (PPE)[94], la France s'est fixée comme objectif d'atteindre 40% de part d'énergies renouvelables en 2030, contre 22% en 2021.

---

[90] Mécanisme créé par la *loi n° 2010-1488 du 7 décembre 2010 portant nouvelle organisation du marché de l'électricité*, publiée au JORF n°0284 du 8 décembre 2010, dite loi « NOME », et codifié aux articles R336-1 et suivants du Code de l'énergie.

[91] Accord de Paris adopté le 12 décembre 2015, signé par la France à New York le 22 avril 2016 ; *Décret n° 2016-1504 du 8 novembre 2016 portant publication de l'accord de Paris adopté le 12 décembre 2015, signé par la France à New York le 22 avril 2016*, publié au JORF n°0262 du 10 novembre 2016, texte n°1.

[92] En France, le « mix énergétique » (répartition des sources d'énergie utilisées pour produire de l'électricité) est déjà largement décarboné avec près de 92% de la production assurée par des sources n'émettant pas de gaz à effet de serre. Sur 2021, on trouve 69% de nucléaire, 12% d'hydraulique, 7% d'éolien, 3% de solaire, contre 7% de thermique fossile et 2% de thermique renouvelable et déchets (RTE, *Bilan électrique 2021*, site officiel, 25 février 2022, disponible en ligne : <https://www.rte-france.com/actualites/bilan-electrique-2021>, consulté le 31 mars 2022.

[93] *Stratégie nationale bas carbone*, version révisée et complète, mars 2020, disponible en ligne : <https://www.ecologie.gouv.fr/sites/default/files/2020-03-25_MTES_SNBC2.pdf>, consultée le 31 mars 2022 ; adoptée par le *Décret n° 2020-457 du 21 avril 2020 relatif aux budgets carbone nationaux et à la stratégie nationale bas-carbone*, publié au JORF n°0099 du 23 avril 2020, texte n°4.

[94] *Programmation pluriannuelle de l'énergie*, mars 2020, disponible en ligne : <https://www.ecologie.gouv.fr/sites/default/files/20200422%20Programmation%20pluriannuelle%20de%20l%27e%CC%81nergie.pdf>, consultée le 31 mars 2022 ; adoptée par le *Décret n° 2020-456 du 21 avril 2020 relatif à la programmation pluriannuelle de l'énergie*, publié au JORF n°0099 du 23 avril 2020, texte n°3.

33. **La décentralisation du réseau électrique.** En outre, la diversification du mix énergétique va conduire à une décentralisation du réseau électrique, avec une multiplication des sites de production, laquelle aura d'importantes conséquences pour le gestionnaire du réseau de transport chargé d'équilibrer production et consommation d'électricité.

34. **La difficile intégration sur le réseau de moyens de production non pilotables et intermittents.** De plus, l'intégration de multiples moyens de productions non pilotables et intermittents (dans le cas du solaire et de l'éolien) fragilise le réseau et nécessite de prévoir plus finement leur production pour adapter en conséquence la production des sources pilotables (dont le nucléaire et l'hydraulique). Le réseau est d'autant plus en tension que les points de consommation d'électricité se diversifient avec l'électrification des usages, se traduisant par exemple par la multiplication des bornes de recharge de véhicules électriques.

35. **Des infrastructures exposées à de nombreux risques.** Enfin, les infrastructures liées à la production, au transport et à la distribution de l'électricité doivent toujours répondre du plus haut niveau de sécurité au vu des conséquences dramatiques que pourrait générer un défaut d'alimentation du pays. C'est la raison pour laquelle une partie des activités de ce secteur sont qualifiées « activités d'importance vitale » regroupant les activités *« essentielles et difficilement substituables ou remplaçables […] visant à produire et à distribuer des biens ou des services indispensables »*, en l'espèce l'approvisionnement en énergie électrique[95]. Les activités de transport et de production d'électricité sont régies à ce titre par un important corpus règlementaire.

36. **Résumé des enjeux rencontrés par le secteur de l'électricité.** Ainsi, le contexte de transition écologique, la décentralisation de la production, le recours croissant à des modes de production intermittents et non pilotables, l'électrification des usages et les impératifs de sécurité constituent les enjeux majeurs auxquels est confronté le secteur de l'électricité. Pour y

[95] *Arrêté du 11 août 2016 fixant les règles de sécurité et les modalités de déclaration des systèmes d'information d'importance vitale et des incidents de sécurité relatives au sous-secteur d'activités d'importance vitale « Approvisionnement en énergie électrique » et pris en application des articles R. 1332-41-1, R. 1332-41-2 et R. 1332-41-10 du code de la défense*, publié au JORF n°0197 du 25 août 2016, texte n°4.

répondre, le système électrique – entendu comme l'ensemble des infrastructures nécessaires à l'approvisionnement en électricité – connaît actuellement de profondes mutations.

37. **Transition.** Parmi elles, l'utilisation des technologies numériques, et notamment le recours à des systèmes d'IA, pourrait contribuer à la transition vers un système électrique plus fiable, sûr, sobre et décarboné[96].

B/ Les utilisations souhaitables de systèmes d'IA dans le secteur de l'électricité

38. **Un secteur en pleine transition numérique.** L'utilisation des nouvelles technologies numériques dans le secteur de l'électricité est déjà particulièrement répandue, si bien que l'on parle depuis quelques années de sa « transition numérique »[97]. D'abord impulsée par un besoin de moderniser des infrastructures de réseau vieillissantes et dépassées[98], elle est aujourd'hui accrue par le contexte de transition écologique. En effet, l'intégration des énergies renouvelables dans le réseau ou l'impératif de sobriété énergétique impliquent à la fois un besoin de flexibilité des réseaux et d'optimisation du pilotage des consommations. La réponse à chacun de ces besoins peut s'appuyer sur des technologies et algorithmes intelligents[99]. Certains acteurs du secteur vont jusqu'à considérer que la transition numérique du secteur de l'énergie est une condition à la transition énergétique[100]. La « mise en données » du secteur de l'énergie[101] ouvre des possibilités de développement de systèmes d'IA autour des enjeux propres au secteur. En effet, l'installation des compteurs Linky, le développement d'objets

---

[96] F. Choné, « L'énergéticien du XXIe siècle : le numérique au service du consommateur et de la transition énergétique, *Annales des Mines - Responsabilité et environnement*, 2017, vol. 87, n°3, 43.
[97] H. Ferreboeuf, « Pour une sobriété numérique », *Futuribles*, 2019, vol. 429, n°2, 15.
[98] INSTITUT MONTAIGNE, *Transition énergétique : faisons jouer nos réseaux*, 2019, disponible en ligne : <https://www.institutmontaigne.org/ressources/pdfs/publications/transition-energetique-faisons-jouer-nos-reseaux-rapport.pdf>, consulté le 5 février 2020.
[99] COLOMBUS CONSULTING, *Data & IA : Quels enjeux et bénéfices concrets pour le secteur de l'énergie ?*, rapport, 2019, disponible en ligne : <https://colombus-consulting.com/nos-publications/data-ia-quels-enjeux-et-benefices-concrets-pour-le-secteur-energie/>, consulté le 5 février 2020.
[100] F. Choné, *op. cit.*
[101] COMMISSION DE RÉGULATION DE L'ÉNERGIE (CRE), *Synthèse de l'étude sur les perspectives stratégiques de l'énergie*, rapport, 2018, disponible en ligne : <http://fichiers.cre.fr/Etude-perspectives-strategiques/1SyntheseGenerale/Perspectives_Strategiques_du_secteur_de_l_energie_Synthese_generale_FR.pdf >, consulté le 5 février 2020.

connectés pour les foyers ou encore la modernisation du réseau de distribution[102] créent autant de sources de données mobilisables pour entraîner des algorithmes et développer des systèmes d'IA.

39. **Les cas d'usage étudiés dans la thèse.** La première étape de la présente étude a été de recenser les principaux cas d'usage de systèmes d'IA dans le secteur de l'électricité, à partir de recherches documentaires, d'observations de terrain et d'entretiens avec des experts. Il ressort de cette analyse que les systèmes d'IA peuvent être utilisés dans chacune des composantes du secteur de l'électricité et y servir, à chaque fois, leur mutation dans un contexte de transition écologique.

40. **L'IA dans la production d'électricité.** Dans la production, le recours à des systèmes d'IA permet notamment de concevoir des applications de maintenance prédictive ou préventive[103]. À titre d'exemple, la technologie de Metroscope, filiale à 100% du Groupe EDF, permet de modéliser un jumeau numérique d'une centrale nucléaire, à partir duquel l'IA peut comparer le fonctionnement réel du site avec le fonctionnement optimal, réaliser un diagnostic et fournir à l'exploitant humain des recommandations sur les actions de maintenance à mener[104]. Dans un autre domaine, des systèmes de reconnaissance d'images peuvent être utilisés pour analyser des photos d'infrastructures et identifier d'éventuels défauts, comme des fissures sur des barrages hydroélectriques[105]. Les systèmes d'IA sont également particulièrement performants en matière de prévision de la production photovoltaïque ou éolienne[106], ce qui

[102] INSTITUT MONTAIGNE, *op. cit.*

[103] H. M. Hashemian, « State-of-the-Art Predictive Maintenance Techniques », *IEEE Transactions on Instrumentation and Measurement*, Janvier 2011, vol. 60, n°1, pp. 226–236.

[104] Voir notamment la filiale du Groupe EDF à 100% « Metroscope » : METROSCOPE, Site officiel, disponible en ligne : <https://metroscope.tech/>, consulté le 1er avril 2022 ; EDF, « EDF lance Metroscope, la solution d'intelligence artificielle au service de l'excellence opérationnelle de ses clients industriels », *Communiqué de presse*, 29 mars 2018, disponible en ligne : <https://www.edf.fr/sites/groupe/files/contrib/groupe-edf/espaces-dedies/espace-medias/cp/2018/cp-20180328-metroscope-vf.pdf>, consulté le 1er avril 2022.

[105] C. Bernstone, A. Heyden, « Image analysis for monitoring of crack growth in hydropower concrete structures », *Measurement*, juillet 2009, vol. 42, n°6, pp. 878–893.
En matière de production hydroélectrique, voir aussi les applications d'IA visant à planifier plus efficacement les arrêts pour maintenance : M. Bulut, E. Özcan, « A new approach to determine maintenance periods of the most critical hydroelectric power plant equipment », *Reliability Engineering & System Safety*, 2021, vol. 205, 107238.

[106] S. A. Kalogirou, « Artificial neural networks in renewable energy systems applications: a review », *Renewable and Sustainable Energy Reviews*, décembre 2001, vol. 5, n°4, pp. 373–401 ; C. Chen, S. Duan, T. Cai, B. Liu, « Online 24-h solar power forecasting based on weather type classification using artificial neural network », *Solar Energy*, novembre 2011, vol. 85, n°11, pp. 2856–2870.

permet la création de services de « centrale virtuelle » agrégeant plusieurs sources de production électrique comme le propose Agregio[107]. Enfin, même si de nombreuses applications sont explorées dans le domaine nucléaire[108], il ressort de notre analyse que les fonctions critiques ne sont pour le moment pas concernées et que l'utilisation de systèmes d'IA se résume à des systèmes d'aide à la décision ne se substituant jamais à la décision humaine.

41. **L'IA dans les réseaux de transport et de distribution d'électricité.** Dans les réseaux de transport et de distribution, on retrouve les mêmes catégories de cas d'usage de systèmes d'IA : maintenance prédictive[109], aide à la décision, ou encore prédiction des niveaux de production ou des besoins de consommation afin d'équilibrer le réseau[110]. Le système électrique doit évoluer pour pouvoir intégrer à la fois les nouveaux moyens de production non pilotables et les nouveaux usages de l'électricité tels que la mobilité électrique, tout en préservant l'équilibre du réseau. Pour ce faire, s'est développé le concept de « réseau intelligent » consistant en *« un réseau bidirectionnel qui connecte producteurs et consommateurs au moyen des nouvelles technologies »*[111]. La décentralisation de la production et de la consommation d'électricité nécessite que les réseaux de transport et de distribution soient plus flexibles, là où auparavant il fallait prévoir la demande énergétique et ajuster la production des centrales au jour le jour[112]. Les technologies numériques peuvent y aider en fournissant des données plus précises sur les besoins de consommation ou sur la capacité de production des énergies

[107] Voir la présentation de la solution de centrale virtuelle fondé sur un système d'IA par Agregio, filiale à 100% du Groupe EDF : AGREGIO, « Savoir-faire », *Site officiel*, disponible en ligne : <https://www.agregio.com/savoir-faire/>, consulté le 1er avril 2022 ; T. Charrier, « Agregio : du post-it à la construction d'un vertical », *Capgemini (blog)*, disponible en ligne : <https://www.capgemini.com/fr-fr/cas-client/agregio-cloud/>, consulté le 1er avril 2022.

[108] M. Gomez-Fernandez, K. Higley, A. Tokuhiro, *et al.*, « Status of Research and Development of Learning-Based Approaches in Nuclear Science and Engineering: A Review », *Nuclear Engineering and Design*, avril 2020, vol. 359, 110479.

[109] M.A. Mahmoud, N.R. Nasir, M. Gurunathan, *et al.*, « The Current State of the Art in Research on Predictive Maintenance in Smart Grid Distribution Network: Fault's Types, Causes, and Prediction Methods – A Systematic Review », *Energies*, 2021, vol. 14, 5078.

[110] V. S. B. Kurukuru, A. Haque, M. A. Khan, *et al.*, « A Review on Artificial Intelligence Applications for Grid-Connected Solar Photovoltaic Systems », *Energies*, 2021, vol. 14, n°15, 4690 ; M. Q. Raza, A. Khosravi, « A review on artificial intelligence based load demand forecasting techniques for smart grid and buildings », *Renewable and Sustainable Energy Reviews*, 2015, vol. 50, pp. 1352–1372,

[111] S. Harvey, « Des réseaux intelligents, stables et fiables : le rôle des réseaux intelligents et de l'électronucléaire dans les systèmes énergétiques bas carbone », *IAEA Bulletin*, septembre 2020, vol. 61, n°3.

[112] INSTITUT MONTAIGNE, *Transition énergétique : faisons jouer nos réseaux*, 2019, disponible en ligne : <https://www.institutmontaigne.org/ressources/pdfs/publications/transition-energetique-faisons-jouer-nos-reseaux-rapport.pdf>, consulté le 5 février 2020.

renouvelables, en aidant les opérateurs du réseau à traiter de plus grandes quantités d'information, ou encore en détectant en amont d'éventuelles vulnérabilités des réseaux[113]. Les utilisations possibles de systèmes d'IA dans ce contexte sont très nombreuses : du pilotage automatique de la demande en électricité[114] à la prédiction de l'apparition de pannes techniques[115], en passant par l'optimisation de la charge de véhicules électriques[116]. Sur ce dernier point, la mobilité électrique n'est pas seulement un facteur de nouvelles contraintes sur le réseau, elle est aussi une opportunité pour contribuer à l'équilibrage en temps réel, grâce à des bornes de recharge bidirectionnelles (stockage et déstockage) pilotées par des systèmes intelligents garantissant à la fois l'équilibrage du réseau et le respect des besoins des utilisateurs. Cette technologie, appelée le « Vehicle-to-Grid » ou « V2G », a été certifiée pour la première fois par le gestionnaire du réseau de transport d'électricité français RTE en février 2022[117].

42. **L'IA dans la fourniture d'électricité et les services associés.** Enfin, dans l'activité de commercialisation et les services liés à l'électricité, les utilisations possibles de systèmes d'IA sont également nombreuses. Elles couvrent la relation entre les clients et leur fournisseur d'énergie, avec notamment la mise en œuvre d'assistants virtuels automatisant la réponse à leurs demandes courantes ou réclamations[118], de solutions d'analyse de données de consommation aux fins d'économies d'énergie[119], ou encore de pilotage de la consommation,

[113] S.D. Ramchurn, P. Vytelingum, A. Rogers, N.R. Jennings, « Putting the 'smarts' into the smart grid: a grand challenge for artificial intelligence », *Communications of the ACM*, 2012, vol. 55, n°4, pp. 86–97.
[114] K.G. Di Santo, S.G. Di Santo, R.M. Monaro, *et al.*, « Active demand side management for households in smart grids using optimization and artificial intelligence », *Measurement*, 2018, vol. 115, pp. 152–161.
[115] C. Rudin, D. Waltz, R. N. Anderson, « Machine Learning for the New York City Power Grid », *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2012, vol. 34, n°2, pp. 328–345.
[116] S.D. Ramchurn, P. Vytelingum, A. Rogers, N.R. Jennings, *op. cit. ;* E. S. Rigas, S. D. Ramchurn, N. Bassiliades, « Managing Electric Vehicles in the Smart Grid Using Artificial Intelligence: A Survey », *IEEE Transactions on Intelligent Transportation Systems*, août 2015, vol. 16, n°4, pp. 1619–1635.
[117] RTE, DREEV, « Pour la première fois en France des véhicules électriques pourront participer à l'équilibrage en temps-réel du système électrique », *Communiqué de presse*, 1er février 2022, disponible en ligne : <https://assets.rte-france.com/prod/public/2022-02/CP_vehicules%20electriques_RTE_Dreev_V2G.pdf>, consulté le 3 avril 2022.
[118] EDF, « Pour EDF l'IA doit être au service de l'humain », *Le Parisien*, 26 juin 2020, disponible en ligne : <https://www.leparisien.fr/societe/pour-edf-l-ia-doit-etre-au-service-de-l-humain-26-06-2020-8337975.php>, consulté le 3 avril 2022.
[119] ENDESA, « Artificial intelligence to improve our services », *Site officiel (blog)*, disponible en ligne : <https://www.endesa.com/en/projects/a201904-artificial-intelligence-improve-services.html>, consulté le 17 décembre 2019.

plus ou moins automatisé[120]. La plupart de ces applications reposent sur l'exploitation des données des clients, notamment des données de consommation générées par des compteurs communicants appartenant au gestionnaire du réseau de distribution ou des objets connectés tels que des thermostats ou chaudières intelligents.

43. **Transition.** L'ensemble des utilisations de systèmes d'IA dans le secteur de l'électricité, qu'elles soient dans la production, la gestion des réseaux ou la commercialisation, peuvent contribuer à rendre l'approvisionnement électrique plus fiable, sûr, sobre et décarboné[121]. Il paraît donc normal que les politiques publiques visent à promouvoir leur développement, que ce soit au niveau national[122] ou européen[123]. Toutefois, l'utilisation croissante de systèmes d'IA n'est pas sans risques pour les droits et libertés des individus.

C/ Les risques liés à l'utilisation de systèmes d'IA dans le secteur de l'électricité

44. **La caractérisation du risque généré par l'utilisation de l'IA.** Le risque peut être défini comme la « *combinaison de la probabilité d'un dommage et de sa gravité* »[124] ou comme la « *possibilité de survenance d'un dommage résultant d'une exposition aux effets d'un phénomène dangereux* »[125]. Les caractéristiques particulières des systèmes d'IA peuvent générer de nouveaux risques pour les individus, ou aggraver la probabilité de survenance de

[120] L. De Matharel, « EDF dégaine Sowee, un hub pour piloter la smart home et une marque IoT », *Le Journal du Net*, 13 octobre 2016, disponible en ligne : < https://www.journaldunet.com/economie/energie/1186520-edf-sowee-station-connectee-smart-home/>, consulté le 3 avril 2022 ; SOWEE, « Avec Sowee & Alexa, votre maison vous obéit au doigt et à la voix ! », *Site officiel (blog)*, 27 juin 2018, disponible en ligne : <https://www.sowee.fr/conseils/autour-de-sowee/avec-sowee-alexa-votre-maison-vous-obeit-au-doigt-et-a-la-voix/>, consulté le 3 avril 2022.

[121] Pour une synthèse des cas d'usage de l'IA dans le secteur de l'électricité, voir notamm. Eurelectric, *AI Insights : The Power Sector in a Post-Digital Age*, rapport, 26 novembre 2020, disponible en ligne : <https://www.eurelectric.org/media/5016/ai-insights-final-report-26112020.pdf>, consulté le 16 décembre 2020.

[122] DGE, *Intelligence artificielle : État de l'art et perspectives pour la France*, rapport commandité par le Pôle interministériel de prospective et d'anticipation des mutations économiques, février 2019, spec. p. 223 et s.

[123] COMMISSION EUROPÉENNE, *Action plan on the digitalisation of the energy sector*, 22 juillet 2021, Ares(2021)4720847.

[124] Norme ISO/CEI 51:2014, *Aspects liés à la sécurité – Principes directeurs pour les inclure dans les normes*.

[125] *Circulaire du 10 mai 2010 récapitulant les règles méthodologiques applicables aux études de dangers, à l'appréciation de la démarche de réduction du risque à la source et aux plans de prévention des risques technologiques (PPRT) dans les installations classées en application de la loi du 30 juillet 2003*, publiée au BO du MEEDDM n°2010/12 du 10 juillet 2010.

ceux préexistants. La présence d'erreurs dans les jeux de données utilisés lors de la conception, la modification de l'environnement dans lequel évolue le système ou encore les potentielles manipulations informatiques sont autant de situations pouvant conduire à la production de résultats erronés par les systèmes d'IA[126]. Selon l'utilisation concernée, ces résultats erronés peuvent causer des préjudices de différente nature. Le préjudice peut être matériel (le système déclenche ou recommande une action endommageant son environnement), corporel (le système déclenche ou recommande une action blessant physiquement un ou plusieurs individus), immatériel (le système déclenche ou recommande une action portant atteinte à la vie privée d'un ou plusieurs ou individus), ou encore financier (le système déclenche ou recommande une action conduisant à des pertes financières importantes). La littérature scientifique a également pu mettre en évidence des risques sociaux liés à l'utilisation de systèmes d'aide à la décision reproduisant des biais discriminatoires[127] et d'autres risques pour les droits et libertés fondamentaux[128].

45. **La caractérisation des risques générés par l'utilisation de l'IA dans le secteur de l'électricité.** Dans le secteur de l'électricité, les principaux risques liés à l'utilisation de systèmes d'IA relèvent d'abord de la **sécurité physique et informatique des infrastructures[129]**. En effet, si un défaut du système venait affecter le fonctionnement de centrales de production d'électricité, les conséquences matérielles et humaines pourraient être désastreuses. En tant qu'infrastructures critiques, les centrales et les réseaux sont déjà sujets à d'importants risques de sécurité, tant liés à la sûreté du fonctionnement des installations qu'aux potentielles attaques terroristes ou informatiques. L'intégration de systèmes d'IA en leur sein

[126] M. Brundage, S. Avin, J. Clark, *et al.*, *The malicious use of artificial intelligence : forecasting, prevention, mitigation*, rapport, février 2018, 101 p., disponible en ligne : <https://arxiv.org/ftp/arxiv/papers/1802/1802.07228.pdf>, consulté le 14 novembre 2019.
[127] C. O'Neil, *Weapons of Math Destruction : How Big Data Increases Inequality and Threatens Democracy*, Crown Random House, 2016, 272 p. ; A. Chander, « The Racist Algorithm ? », *Michigan Law Review*, 2017, 115 ; S. Barocas, A. Selbst, « Big Data's Disparate Impact », *California Law Review*, 2016, vol. 104, 671.
[128] CONSEIL D'ÉTAT, *Le numérique et les droits fondamentaux*, Etude annuelle 2014, La Documentation française, 64, 447 p. ; CONSEIL DE L'EUROPE, *Algorithmes et droits humains*, Etude menée par le comité d'experts sur les intermédiaires d'internet MSI-NET, 2017, DGI (2017)12, disponible en ligne : <https ://rm.coe.int/algorithms-and-human-rights-fr/1680795681>, consulté le 8 mars 2020.
[129] V. notamm. les exemples pris dans : EXPERT GROUP ON LIABILITY AND NEW TECHNOLOGIES, *Liability for AI and other emerging digital technologies*, 21 novembre 2019, spec. p. 32 et s., disponible en ligne : <https://op.europa.eu/en/publication-detail/-/publication/1c5e30be-1197-11ea-8c1f-01aa75ed71a1/language-en/format-PDF>, consulté le 18 janvier 2020.

peut multiplier les facteurs de risques en y ajoutant des composants opaques et autonomes. Ensuite, le développement de l'IA dans le pilotage du réseau ou les services énergétiques peut nécessiter l'utilisation de grandes quantités de données, dont certaines à caractère personnel. À cet égard, le déploiement d'objets connectés collectant des données, y compris dans les foyers, et leur traitement « en nuage » peuvent générer des **risques pour la sécurité des données personnelles et, ce faisant, pour la vie privée des individus**[130]. Enfin, et bien que les systèmes d'IA puissent être utilisés dans le secteur de l'électricité pour contribuer à la transition énergétique, leur propre empreinte environnementale ne doit pas être négligée[131]. En effet, leur fonctionnement nécessite la mobilisation d'une importante puissance de calcul, requérant donc une grande quantité d'énergie, la collecte et le stockage de nombreuses données ainsi que la fabrication d'infrastructures dédiées (centres de données, puces électroniques, capteurs connectés…). À titre d'exemple, les émissions liées aux besoins en électricité nécessaires à l'apprentissage de certains modèles de langage naturel peuvent représenter jusqu'à cinq fois les émissions d'une voiture thermique sur l'ensemble de sa durée de vie[132]. En raison de leur fonctionnement énergivore, et malgré leur potentiel pour aider les entreprises à réaliser des économies d'énergie[133], les systèmes d'IA peuvent donc, paradoxalement, **aggraver les risques environnementaux**.

46. **Transition.** L'ensemble de ces risques justifient l'attention portée par les législateurs et les juristes à l'IA. Les enjeux juridiques du développement de l'IA dans le secteur de l'électricité sont nombreux et posent notamment la question de l'opportunité de réguler cette technologie qui semble promettre autant de bénéfices qu'elle n'apporte de maux.

---

[130] A. Debet, « Intelligence artificielle et données à caractère personnel », in *Droit de l'intelligence artificielle*, dir. A. Bensamoun, G. Loiseau, LGDJ, coll. Les intégrales, 2019, p. 269.
[131] P. Dhar, « The Carbon Impact of Artificial Intelligence », *Nature Machine Intelligence*, 2020, vol. 2, n°8, 423.
[132] E. Strubel, « Energy and Policy Considerations for Deep Learning in NLP », *57th Annual meeting of the Association for Computational Linguistics*, 5 juin 2019, p. 1.
[133] R. Evans, J. Gao, « DeepMind AI reduces Google data centre cooling bill by 40% », *DeepMind website (blog)*, 20 juillet 2016, disponible en ligne : <https://www.deepmind.com/blog/deepmind-ai-reduces-google-data-centre-cooling-bill-by-40>, consulté le 12 février 2019.

IV - La problématisation de la recherche

47. **La nécessaire mise en balance des risques et bénéfices liés à l'utilisation de l'IA dans le secteur de l'électricité.** Comme on a pu le voir, il existe des utilisations souhaitables de l'IA dans le secteur de l'électricité, celles utiles à sa modernisation dans un contexte de transition écologique, et des utilisations non souhaitables, générant de nouveaux risques ou aggravant ceux existants. Toutefois, la frontière entre ces deux catégories n'est pas étanche : certaines applications peuvent être indispensables à la transition énergétique, par exemple celles relatives à l'équilibrage du réseau électrique, tout en portant en elles-mêmes des risques non négligeables, liés à la numérisation d'infrastructures critiques ou à la collecte invasive de données sur les individus. Il convient donc de trouver des solutions pour s'assurer que les systèmes d'IA développés dans le secteur de l'électricité présentent **une balance bénéfices-risques positive**. Ces solutions peuvent d'abord être techniques et consister en la découverte de nouvelles méthodes plus sûres[134]. Elles peuvent également être d'ordre économique avec la mise en place de dispositifs financiers incitatifs pour encourager le développement de systèmes d'IA souhaitables ou désincitatifs pour les applications peu vertueuses. Les solutions peuvent enfin être juridiques, par exemple en imposant des règles contraignantes sur la conception ou l'utilisation de systèmes d'IA afin d'en minimiser les risques, en mettant en place un système de surveillance par une autorité de contrôle, en allégeant les fardeaux règlementaires sur les cas d'usage considérés comme les plus vertueux, ou encore en renforçant les droits des individus afin qu'ils puissent exercer un contrôle sur la technologie. La présente étude traitera de ce dernier volet, en tâchant de répondre à la problématique suivante : **par quels moyens peut-on construire un cadre juridique conciliant promotion de l'innovation et prévention des risques liés à l'utilisation de l'IA dans le secteur de l'électricité ?** La question ainsi posée vise à interroger tant les adaptations potentielles du corpus existant que les moyens juridiques mobilisables pour bâtir un cadre spécifique à l'IA afin que le droit dans son ensemble remplisse l'objectif légitime de mise en balance des risques et bénéfices de la technologie.

[134] Voir par exemple l'ensemble des techniques numériques mobilisables pour protéger plus efficacement la vie privée, appelées *« privacy enhancing technologies »* ou *« P.E.T. »* : J. Heurix, P. Zimmermann, T. Neubauer, S. Fenz, « A taxonomy for privacy enhancing technologies », *Computers & Security*, 2015, vol. 53, pp. 1–17.

48. **Méthodologies descriptive et prescriptive.** Penser l'encadrement juridique d'un objet de droit nouveau comme le sont les systèmes d'IA nécessite de suivre une méthodologie rigoureuse. Selon la Professeure Véronique Champeil-Desplats dans son ouvrage *Méthodologie du droit et des sciences du droit*, deux grands types de postures sont habituellement distinguées : *« une posture descriptive qui se donne pour objet de répondre à la question « comment se comporte de fait » le juriste ; et une posture prescriptive qui tente de répondre à la question « comment doit se comporter » le juriste »*[135]. Le terme « juriste » renvoyant tant aux autorités normatives, habilitées à produire des normes juridiques (le législateur, les juges, l'administration), qu'à ceux qui analysent et commentent cette production. La présente étude nécessitera de mêler les deux méthodologies.

49. **Une approche descriptive.** D'une part, une approche descriptive permettra d'identifier les normes juridiques applicables aux systèmes d'IA et analyser les effets de leur application. Cette analyse pourra mettre en lumière d'éventuelles lacunes ou des contraintes juridiques disproportionnées au regard des bénéfices attendus de la technologie.

50. **Une approche prescriptive.** D'autre part, l'étude complètera systématiquement ces constats par une approche prescriptive, en proposant des solutions visant au perfectionnement du système juridique. Ces propositions pourront être justifiées, selon le cas, par des arguments *a completudine*, ayant pour but la résorption des lacunes du corpus existant[136], *a coherentia*, visant à sa cohérence[137], ou encore *ab exemplo*, s'appuyant sur l'analogie avec les moyens juridiques employés dans des situations similaires[138].

51. **Une recherche appliquée.** L'ambition de la thèse est de donner une visée pratique aux propositions défendues. Ces dernières pourront être utiles aux entreprises développant des systèmes d'IA afin de comprendre les problématiques posées par le cadre juridique actuel, mais aussi aux autorités normatives qui y trouveront des pistes de réforme, des analyses critiques des

[135] V. Champeil-Desplats, *op. cit.*, 17 ; citant également : N. Bobbio, « Metodo », in *Contributi ad un dizionario giuridico*, Giapichelli editore, Turino, 1994, pp. 171–173 ; U. Scarpelli, « La natura della metodologica giuridica », *Rivista internazionale de filosofia del Diritto*, 1956, XXXIII, pp. 247–255.
[136] G. Tarello, « Sur la spécificité du raisonnement juridique », *Archives de philosophie du droit et de philosophie sociale*, 1972, n°7, p. 105.
[137] *Ibid.*
[138] *Ibid.*

projets de règlementation en cours, ainsi que des propositions de nouvelles règles *ad hoc*. Cette visée pratique est également justifiée par le contexte de la réalisation de cette étude, à savoir le partenariat entre une entreprise privée, EDF, et une université dans le cadre d'une convention CIFRE. Sur un sujet tel que l'IA où la compréhension de la technique est cruciale pour pouvoir questionner les discours de promotion ou de critique qui l'accompagnent[139], le travail en entreprise a constitué un réel atout, en ce qu'il a permis d'ancrer la réflexion théorique dans une réalité pratique[140].

52. **Une recherche transposable.** La réflexion menée sur l'encadrement juridique des systèmes d'IA, bien qu'abordée sous le prisme du secteur de l'énergie électrique, est transposable à d'autres secteurs hautement régulés. En effet, la méthode employée, visant à questionner les normes juridiques applicables à la lumière des apports potentiels de la technologie, peut être répliquée dans d'autres domaines tels que la finance, l'automobile ou l'aéronautique. De plus, certaines problématiques traitées ne sont pas spécifiques au secteur étudié et peuvent donc donner lieu à des propositions transversales.

V – La démonstration proposée

53. **La diversité du corpus juridique étudié.** Contrairement à l'idée commune selon laquelle les nouvelles technologies, et en particulier les systèmes d'IA, évoluent dans un vide juridique, le fait est que leur conception, leur distribution et leur utilisation ne sont pas exemptées du respect du droit existant. Ce droit est composé, d'une part, des règles juridiques de portée générale, sans lien avec le secteur d'activité. C'est le cas par exemple des régimes de responsabilité civile et pénale, de propriété intellectuelle, de protection des droits et libertés fondamentaux, ou de protection des données à caractère personnel. D'autre part, dans le cadre de notre étude, l'environnement juridique dans lequel naissent et évoluent les systèmes d'IA est

[139] F. Rivière, *op. cit.*

[140] La littérature comprend peu de travaux « sectoriels » en dehors de quelques recherches dans les domaines de la santé (voir par exemple : D. Gruson, « Enjeux juridiques de l'intelligence artificielle en santé : le stable et le mouvant », *Revues des Juristes de Sciences Po*, juin 2021, n° 21, 16) ou de la finance (voir par exemple : T. de Ravel d'Esclapon, « La gouvernance des algorithmes dans le secteur financier : le point de vue de l'ACPR », Revue de Droit bancaire et financier, juillet 2020, n° 4, étude 12).

également composé de l'ensemble des règles sectorielles, spécifiques au secteur de l'électricité. Ces dernières sont majoritairement d'origine règlementaire et peuvent émaner d'autorités de régulation sectorielles telles que la Commission de régulation de l'énergie (CRE) pour le fonctionnement du marché ou l'Autorité de sûreté nucléaire (ASN) pour la gestion des centrales nucléaires par exemple. On trouvera également de nombreuses règles spécifiques, y compris d'origine législative, dans les codes de l'énergie, de l'environnement ou de la défense.

54. **Une étude du corpus existant.** L'application des normes juridiques existantes génère des obstacles à l'utilisation des systèmes d'IA dans le secteur de l'électricité. À titre d'exemple, l'impératif de sécurité du système électrique a justifié une règlementation particulièrement stricte du secteur, ne laissant aucune possibilité de recours à de nouvelles techniques, fussent-elles plus sûres. De la même manière, le régime de la protection des données à caractère personnel tel qu'il existe aujourd'hui ne permet pas aux énergéticiens de tirer parti de l'IA pour proposer des applications vertueuses, notamment d'optimisation énergétique. Si la contrainte est la plupart du temps justifiée, la présente étude entend questionner sa proportionnalité au regard des bénéfices potentiels du recours aux systèmes d'IA dans le secteur de l'électricité. De plus, l'application de certains régimes juridiques anciens, construits dans un contexte technologique différent, peut être source d'insécurité juridique pour les entreprises souhaitant développer ou utiliser des systèmes d'IA. L'ensemble de ces constats nous poussent à croire que **le cadre juridique actuel doit être adapté** (**Partie 1**). Après avoir identifié les régimes juridiques dont l'application génère des obstacles à l'utilisation de systèmes d'IA dans le secteur de l'électricité, les propositions formulées dans la première Partie viseront à clarifier les dispositions à l'application incertaine et à réformer les contraintes règlementaires disproportionnées par rapport au risque réel.

55. **Une proposition de corpus à construire.** Si le corpus existant couvre déjà une grande partie des risques liés à l'utilisation de systèmes d'IA, il semble aujourd'hui insuffisant pour garantir un développement vertueux de la technologie. La question de l'empreinte environnementale de l'IA, considérée comme l'un des *« angles mort de la doctrine*

environnementaliste » par Gilles Martin[141], en est une illustration. Ainsi, outre l'adaptation de l'existant, **la création d'un cadre juridique spécifique aux systèmes d'IA**, conciliant promotion des utilisations souhaitables et prévention des risques générés par les utilisations dangereuses, **apparaît donc indispensable** (**Partie 2**). Les développements de la deuxième Partie présenteront les motifs justifiant la création de ce cadre, les objectifs qu'il doit poursuivre et les moyens juridiques pour les atteindre, sur le fond et sur la forme, en s'inspirant notamment de ceux employés pour réguler des situations analogues. À la lumière de ces réflexions théoriques, pourront être analysés les projets de règlementation initiés à travers le monde, à commencer par la proposition de règlement européen sur l'IA[142], afin de formuler, le cas-échéant, des propositions d'amendements.

[141] G.J. Martin, « Les angles morts de la doctrine juridique environnementaliste », *Revue juridique de l'environnement*, Lavoisier, 2020/1, 45, nº 67-80.
[142] COMMISSION EUROPÉENNE, *Proposition de règlement du Parlement européen et du Conseil établissant des règles harmonisées concernant l'intelligence artificielle (législation sur l'intelligence artificielle) et modifiant certains actes législatifs de l'Union*, 21 avril 2021, 2021/0106 (COD)..

Partie 1 : *De lege lata*, la nécessaire adaptation d'un droit inefficace

56. **Les limites de l'application par défaut du droit actuel aux systèmes d'IA.** Le corpus législatif actuel, s'il a su s'adapter par le passé à de nouveaux objets de droit[143], semble aujourd'hui présenter des lacunes lorsqu'il est appliqué aux systèmes d'IA. En effet, ces derniers présentent plusieurs spécificités par rapport aux logiciels traditionnels – la complexité structurelle, la dépendance à la donnée, la potentielle autonomie et l'opacité du fonctionnement[144] – pouvant gêner l'application de la règle de droit.

57. **Les limites de l'application des régimes de droit commun à l'IA.** Le constat se vérifie notamment dans l'étude des régimes de droit commun[145] tels que le droit de la responsabilité ou le droit des données à caractère personnel. Par exemple, les régimes de responsabilité civile reposent en grande partie sur des concepts tels que la faute ou le lien de causalité dont l'application à des systèmes logiciels opaques et fonctionnant avec un certain niveau d'autonomie peut être rendue très compliquée. Quel est le statut juridique de l'IA ? Peut-elle être fautive ? Un humain peut-il être tenu responsable d'un apprentissage réalisé de façon autonome par la machine ? Comment attribuer la responsabilité à un acteur sans être en mesure d'identifier l'origine du dommage causé par une IA dont le fonctionnement est opaque ? Autant de questions que les opérateurs de systèmes d'IA dans le secteur de l'électricité se posent avant d'y avoir recours. L'absence de réponse claire génère une insécurité juridique freinant l'adoption de l'IA. Cette insécurité peut être levée, notamment, par une adaptation à la marge du corpus existant.

58. **Les limites de l'application des règlementations sectorielles à l'IA.** De la même manière, en plus des régimes de droit commun, certaines règles juridiques spécifiques au secteur de l'électricité peuvent contraindre le développement de l'IA. L'électricité étant un bien de première nécessité, il semble normal que ce secteur fasse l'objet d'une régulation très stricte. Il est également logique que certaines des règles qui la composent restreignent l'utilisation de

---

[143] Pour un historique de l'adaptation du droit civil aux révolutions industrielles et numériques, voir S. Savatier, *Les métamorphoses économiques et sociales du droit civil d'aujourd'hui*, Panorama des mutations, 3e éd., 1964.
[144] Sur les attributs spécifiques des systèmes d'IA : voir Supra, 11-16.
[145] On entend ici par « régimes de droit commun » tout corpus de règles qui n'est pas spécifique à un secteur d'activité précis.

l'IA afin de garantir la sécurité d'approvisionnement ou la sûreté des infrastructures critiques du secteur. Pour autant, il est essentiel de préserver la capacité d'innovation de ses acteurs au vu des bénéfices potentiels que l'IA peut apporter, notamment en termes de performance et de sécurité. Pour ce faire, un assouplissement de certaines règles sectorielles peut également s'avérer nécessaire.

59. **Plan.** Aux fins de clarté, la présente Partie traitera de ces deux problématiques de façon successive. Nous nous intéresserons en premier lieu aux régimes de droit commun qui, s'ils sont pleinement applicables aux systèmes d'IA, présentent des lacunes ou imprécisions lorsqu'ils sont confrontés aux spécificités de ces nouveaux objets de droit et nécessitent à ce titre d'être adaptés (**Titre 1**). En second lieu, puisque notre thèse s'inscrit dans un secteur précis, seront étudiées les principales règles juridiques susceptibles de s'appliquer aux systèmes d'IA dans le secteur de l'électricité. Certaines apparaissent très contraignantes et sont susceptibles de générer un frein disproportionné à l'adoption de la technologie (**Titre 2**).

Titre 1 : Une insécurité juridique résultant de l'application des régimes de droit commun

60. **L'application indifférenciée des textes européens aux systèmes d'IA.** Les premiers articles de doctrine juridique sur l'IA avaient principalement pour objectifs de comprendre les récentes avancées technologiques ainsi que d'analyser comment le Droit pouvait les appréhender[146]. Ces premiers travaux, complétés par ceux visant à déterminer la qualification de l'IA, en tant que sujet ou objet de droit[147], ont sans doute inspiré les institutions européennes dans les années qui ont suivi. Ainsi, le Conseil de l'Europe a publié en 2017 l'étude « *Algorithmes et droits humains* »[148] mettant en lumière la menace que pourrait représenter l'utilisation de l'IA pour les droits fondamentaux, notamment à l'égard des droits à la dignité, à la non-discrimination, au respect de la vie privée ou à un recours juridictionnel effectif[149]. Il convient de préciser sur ce point que les règlementations européennes en matière de droits de l'homme ont vocation à pleinement s'appliquer aux systèmes d'IA. À titre d'exemple, les directives européennes sur la non-discrimination[150], sur l'égalité en matière d'emploi et de travail[151] ou encore le règlement général sur la protection des données (RGPD)[152] s'appliquent sans distinction aux systèmes d'IA. Toujours en droit européen, les textes relatifs à la sécurité des produits ont également vocation à s'appliquer aux risques générés par les systèmes d'IA.

61. **L'application indifférenciée des législations nationales aux systèmes d'IA.** En droit interne, il en va de même pour les régimes de droit commun tels que le droit des contrats, de la

---

[146] Voir notamm. X. Labbée, « L'homme augmenté », *Recueil Dalloz*, 2012, 2323 ; G. Loiseau, « Des robots et des hommes », *Recueil Dalloz*, 2015, 2369 ; A. Bensoussan, « Droit des robots : science-fiction ou anticipation ? », Recueil Dalloz, 2015, 1640 ; A. Bensamoun, G. Loiseau. « L'intégration de l'intelligence artificielle dans l'ordre juridique en droit commun : questions de temps », Dalloz IP/IT, 2017, 239.

[147] G. Loiseau, « Des robots et des hommes », *op. cit.* ; A. Mendoza-Caminade, « Le droit confronté à l'intelligence artificielle des robots : vers l'émergence de nouveaux concepts juridiques ? », *Recueil Dalloz*, 2016, 445.

[148] CONSEIL DE L'EUROPE, *Algorithmes et droits humains*, Etude menée par le comité d'experts sur les intermédiaires d'internet MSI-NET, 2017, DGI (2017)12, disponible en ligne : <https ://rm.coe.int/algorithms-and-human-rights-fr/1680795681>, consulté le 8 mars 2020.

[149] *Ibid*, p. 11 et s.

[150] *Directive 2000/43/CE du Conseil du 29 juin 2000 sur l'égalité de traitement entre les personnes sans distinction de race ou d'origine ethnique*, publiées au JOUE n°L180 du 19 juillet 2000.

[151] *Directive 2000/78/CE du Conseil du 27 novembre 2000 sur l'égalité de traitement en matière d'emploi et de travail*, publiée au JOUE n°L303 du 2 décembre 2000.

[152] *Règlement (UE) 2016/679 du Parlement européen et du Conseil du 27 avril 2016 relatif à la protection des personnes physiques à l'égard du traitement des données à caractère personnel et à la libre circulation de ces données, et abrogeant la directive 95/46/CE (règlement général sur la protection des données)*, publié au JOUE n°L119/1 le 4 mai 2016, ci-après le « RGPD ».

responsabilité civile ou de la propriété intellectuelle[153]. À l'aune de ces analyses, force est de constater que le corpus juridique de droit commun s'applique de façon indifférenciée aux systèmes d'IA qui ne sont en réalité qu'un outil mobilisé par l'homme.

62. **Les limites de l'application indifférenciée du corpus existant aux systèmes d'IA.** Toutefois, l'application par défaut de ces règles de droit, conçues dans un autre contexte technologique, peut être troublée par les spécificités des systèmes d'IA présentées précédemment[154]. D'abord, leur mise en œuvre peut être gênée par l'opacité d'algorithmes que le juge, même éclairé par des experts, peinera à comprendre[155]. Vérifier le respect de la législation en vigueur devient alors très difficile et la possibilité de contester les décisions prises à l'aide de l'IA risque d'être compromise. Ensuite, l'effectivité des règles peut être remise en cause s'il subsiste une trop grande incertitude quant à leur applicabilité. Enfin, les régimes préexistants peuvent être incomplets s'ils échouent à couvrir les nouveaux risques générés par l'utilisation de l'IA.

63. **Délimitation de l'étude et méthodologie suivie.** Les attributs spécifiques des systèmes d'IA peuvent donc remettre en cause le corpus législatif existant. Le présent Titre n'a pas pour objectif de recenser tous les régimes juridiques susceptibles de s'appliquer à l'IA. Dans le cadre de la thèse, réalisée en convention CIFRE avec EDF S.A., une vingtaine d'entretiens semi-directifs ont été réalisés auprès de professionnels. Les personnes interrogées souhaitent pour la majorité garder l'anonymat et le contenu des entretiens peut révéler des informations confidentielles, c'est pourquoi seules des tendances générales pourront être évoquées dans nos développements. Les profils interrogés sont divers : juristes du Groupe EDF, dirigeants des filières informatiques, ingénieurs-chercheurs, dirigeants de start-ups innovantes, ou encore experts en politiques publiques européennes. Chaque entretien a été mené selon une méthode définie en amont. Chaque professionnel s'est vu posé une série de trois questions générales sur la définition de l'IA, la présentation de son activité en lien avec l'IA et les principaux freins juridiques identifiés dans la pratique avant de laisser place à une discussion libre. Les réponses

[153] Pour une étude de l'IA à travers ses interactions avec plusieurs régimes de droit commun, voir notamm. A. Bensamoun, G. Loiseau, *Droit de l'intelligence artificielle*, LGDJ, coll. Les intégrales, 2019, 444p.
[154] Voir Supra, 11-16.
[155] E. Barbin, « Le contrôle juridictionnel de l'outil numérique d'aide à la décision administrative », *RFDA*, 2021, 491.

des sondés ont permis d'orienter de nombreuses parties de la présente thèse afin de l'ancrer dans une réalité pratique. Deux thématiques sont revenues dans la majorité des entretiens menés dans le cadre de la thèse et sont perçues par les acteurs du secteur de l'électricité comme les deux principales sources d'incertitudes juridiques en matière d'IA : la responsabilité et la protection des données à caractère personnel. Pour apporter des réponses concrètes aux professionnels du secteur de l'énergie, les développements qui suivent se concentreront sur ces deux régimes de droit commun.

64. **La responsabilité du fait de l'IA.** D'une part, l'applicabilité des régimes de responsabilité aux dommages causés par un système d'IA a déjà fait l'objet de plusieurs analyses[156]. Pourtant, l'absence de jurisprudence en la matière inquiète les entreprises développant des systèmes d'IA qui peinent à identifier quel régime de responsabilité serait applicable en cas de préjudice. De plus, ces dernières sont également préoccupées par l'éventuelle répartition des responsabilités qui n'est clarifiée ni par la lettre des textes en vigueur, ni par la jurisprudence. Un effort de clarification des régimes de responsabilité apparaît nécessaire (**Chapitre 1**) dans un souci de sécurité juridique.

65. **La protection des données personnelles traitées par l'IA.** D'autre part, si l'applicabilité des règles liées au traitement des données personnelles n'est pas remise en cause par les spécificités des technologies intelligentes, leur application fait tout de même naître de fortes contraintes. Outre les coûts importants liés à la mise en conformité, la règlementation en matière de données à caractère personnel présente des incompatibilités avec les caractéristiques des systèmes d'IA. À titre d'exemple, ces derniers nécessitent le traitement de grandes quantités de données pour leur conception et leur fonctionnement, tandis que le RGPD impose un principe de minimisation des données collectées. À bien des égards, les textes applicables pénalisent l'innovation. Un rééquilibrage des contraintes liées à l'exploitation des données personnelles nous apparaît cependant possible pour préserver la capacité d'innovation des entreprises (**Chapitre 2**).

---

[156] J. Pouget, *La réparation du dommage impliquant une intelligence artificielle*, Thèse pour le doctorat en droit privé, Université d'Aix Marseille, 2019, 410 p. ; J.S. Borghetti, « Civil Liability for Artificiel Intelligence : What Should Its Basis Be ? », *Revue des juristes de Sciences Po*, 2019, n°17, pp. 76-84.

Chapitre 1 : Une clarification nécessaire des régimes de responsabilité

66. **La perception des enjeux de responsabilité par les acteurs de l'IA dans le secteur de l'électricité.** Les entretiens réalisés dans le cadre de la thèse ont montré que les questions de responsabilité sont parmi les principales préoccupations des personnes sondées sur les enjeux juridiques du développement de l'IA. Le constat se vérifie d'ailleurs tant auprès de profils juridiques que techniques, et quel que soit le niveau hiérarchique. En effet, les mythes et les préjugés autour des capacités de l'IA, les scénarios de science-fiction et les diverses études sur les dilemmes moraux auxquels pourraient être confrontés des systèmes autonomes[157] ont contribué à jeter le flou sur les conséquences de la réalisation d'un dommage causé par l'action d'un système d'IA et en particulier sur la réponse à la question « Qui est responsable ? ». Sur ce point, on observe toutefois une différence notable entre les points de vue des experts techniques et des juristes. Les premiers s'interrogent sérieusement sur la désignation du responsable d'un dommage causé par un système autonome. Ils craignent, d'une part, que ces dommages échappent aux régimes de responsabilité civile et ne soient pas réparés, constituant un frein moral à l'adoption de l'IA. Les experts techniques craignent, d'autre part, que la répartition des responsabilités entre les différents acteurs de la chaîne d'approvisionnement de l'IA soit inéquitable, constituant cette fois un frein économique. Les juristes, quant à eux, cherchent à clarifier l'articulation des différents régimes existant mais peinent à en identifier les conséquences pratiques. Selon eux, la couverture contractuelle des risques entre les différents fournisseurs et utilisateurs est à même de couvrir, dans l'état actuel de la technique, les risques générés par l'utilisation de l'IA. Une minorité parmi les juristes interrogés ajoute à cet argument que les régimes actuels de responsabilité civile extracontractuelle sont parfaitement capables d'appréhender des préjudices causés par des systèmes d'IA, notamment à travers les régimes de responsabilité sans faute.

67. **Les potentiels dommages causés par des systèmes d'IA dans le secteur de l'électricité.** Il est naturel de questionner l'applicabilité des régimes de responsabilité lorsque l'on s'intéresse aux utilisations de l'IA dans le secteur de l'électricité. En effet, outre sa fonction normative, le droit de la responsabilité a pour fonction principale de réparer les dommages subis

[157] E. Awad, S. Dsouza, R. Kim, *et al.*, « The Moral Machine experiment », *Nature*, 2018, 563, 59-64.

par les individus. Or, il est possible que les applications de l'IA dans le secteur de l'électricité causent des préjudices, et ce, malgré toutes les précautions prises dans leur conception. À titre d'exemple, une mauvaise recommandation produite par un système de maintenance prédictive pourrait conduire à un incident de sécurité dans une centrale de production d'électricité. La réalisation automatique de certaines opérations dans le circuit primaire pourrait causer des dommages corporels au personnel sur place (rejet de matière radioactive dans l'environnement, fuite de gaz, chute de matériel…) ou des préjudices économiques pour les populations si la centrale venait à être arrêtée par mesure de précaution, pouvant alors causer des coupures de courant. C'est la raison pour laquelle la décision reste aujourd'hui complètement humaine, l'IA n'intervenant qu'à titre d'aide à la décision dans les systèmes informatiques de la production d'électricité. Toutefois, il est normal pour le concepteur du système d'IA ayant causé cette situation, l'exploitant de la centrale ou encore les fournisseurs des composants physiques ayant été actionnés par l'IA, de s'interroger sur la répartition des responsabilités si un tel événement venait à se produire[158].

68. **L'importance d'une juste répartition des responsabilités en cas de dommage.** La réponse à cette question passe par l'étude de l'applicabilité des régimes de responsabilité aux cas où l'opération d'un système d'IA serait la cause d'un dommage. Cette étude doit d'une part identifier quel régime de responsabilité, civile ou pénale, contractuelle ou délictuelle, pour faute ou sans faute, doit s'appliquer à chaque situation. Elle doit d'autre part permettre aux acteurs de la chaîne d'approvisionnement de l'IA d'anticiper la répartition éventuelle des responsabilités qui doit être la plus claire et la plus juste possible, sans quoi une réticence à l'investissement pourrait être créée. Or un tel frein à l'innovation, né d'une incertitude juridique, pourrait être facilement levé pour libérer le potentiel de l'IA dans le secteur de l'électricité soit simplement grâce à une meilleure compréhension des différents régimes, soit, dans certains cas, par une clarification législative ou jurisprudentielle.

[158] S. Migayron, « Pratique contentieuse : Intelligence artificielle : qui sera responsable ? », *Communication commerce électronique*, avril 2018, n°4.

69. **Objectifs du Chapitre.** Au vu de l'abondance de la doctrine sur le sujet des enjeux de responsabilité liés au développement de l'IA[159], l'ensemble des enjeux et propositions d'adaptation du corpus existant ne seront pas étudiés. Le présent Chapitre vise à présenter, à partir d'exemples concrets issus du secteur de l'électricité, les forces et faiblesses de chaque régime de responsabilité s'ils étaient appliqués à un dommage causé par l'action d'un système d'IA. Les solutions consensuelles en doctrine y seront présentées, critiquées et, le cas-échéant, complétées par nos propres propositions d'adaptation.

70. **Plan.** Les régimes de responsabilité en droit français semblent en grande partie suffisamment adaptables pour appréhender les dommages causés par l'action de systèmes d'IA. La responsabilité pénale sera rapidement étudiée, mais notre analyse se concentrera sur la responsabilité civile. En cas de préjudice causé par un système d'IA, la désignation du ou des responsables dépendra de l'origine du dommage. Ce dernier peut être intentionnel ou non, causé par une négligence, un défaut dans la conception du système ou dans l'utilisation qui en est faite. Le défaut de conception peut être lié à une mauvaise programmation de la composante logicielle du système, à l'utilisation de données biaisées lors de la phase d'apprentissage, à la mauvaise anticipation des conditions réelles d'utilisation ou encore à un sabotage intentionnel par un développeur. La diversité de ces hypothèses complexifie la désignation du responsable en cas de dommage causé par l'action d'un système d'IA. En effet, le régime de responsabilité applicable dépendra des circonstances de la réalisation du préjudice et de son origine. Dès lors, il n'est pas possible de fournir une réponse unique à la question « Qui est responsable ? ». L'objectif du présent Chapitre est de fournir aux entreprises du secteur de l'électricité une synthèse des régimes de responsabilité applicables lorsque l'action d'un système d'IA causerait un dommage. En effet, il n'existe pas de vide juridique : plusieurs régimes de responsabilité peuvent être applicables suivant les circonstances de réalisation du dommage (**Section 1**). Les

[159] L'idée que le développement de l'IA puisse remettre en cause le bon fonctionnement du droit de la responsabilité, civile et pénale, n'est pas nouvelle et a fait l'objet d'une doctrine abondante. Elle est évoquée dès 2012 par Xavier Labbée (X. Labbée, « L'homme augmenté », *Recueil Dalloz*, 2012, 2323), puis cristallise les débats autour de l'IA de 2015 à nos jours (V. notamm. A. Bensoussan, « Droit des robots : science-fiction ou anticipation ? », Recueil Dalloz, 2015, 1640 ; G. Courtois, « Robots intelligents et responsabilité : quels régimes, quelles perspectives ? », *Dalloz IP/IT*, 2016, 287 ; C. Coulon, « Du robot en droit de la responsabilité civile : à propos des dommages causés par les choses intelligentes », *RCA*, 2016, Étude 6 ; J. Pouget, *La réparation du dommage impliquant une intelligence artificielle*, Thèse pour le doctorat en droit, Université d'Aix-Marseille, 2019, 410 p.).

régimes dits « sans faute » semblent les plus adaptés, bien que persistent certains doutes quant à leur applicabilité à tous les types de systèmes d'IA, notamment les plus autonomes (**Section 2**). Ces constats soulèvent la question de l'opportunité d'une réforme des régimes de responsabilité, qui nous apparaît souhaitable (**Section 3**).

Section 1 : L'applicabilité limitée des régimes de responsabilité pour faute

71. **Plan.** Suivant la nature du dommage causé par un système d'IA, son origine, les circonstances de sa réalisation et les acteurs impliqués, plusieurs régimes de responsabilité pour faute peuvent trouver à s'appliquer (**§1**). Toutefois, les spécificités de l'IA, notamment son autonomie et son opacité, peuvent rendre difficile la démonstration de l'existence d'une faute et ainsi remettre en cause l'applicabilité de tous ces régimes (**§2**).

§1 : La diversité des régimes de responsabilité pour faute potentiellement applicables

72. **Plan.** Le présent paragraphe adopte une acception large de la notion de faute, incluant les fautes pénales, délictuelles et contractuelles. D'abord, la responsabilité pénale des fournisseurs ou utilisateurs de systèmes d'IA peut être engagée dans certaines circonstances (**A**). Ensuite, le développement et l'utilisation d'un système d'IA se font généralement dans le cadre d'une relation contractuelle entre les différentes parties prenantes. Dès lors, le contrat apparaît comme un moyen privilégié pour organiser une répartition consensuelle des responsabilités mais peut également permettre l'engagement de la responsabilité des co-contractants (**B**). Enfin, un manquement des concepteurs à leur devoir de sécurité peut justifier la mobilisation du régime de la responsabilité civile délictuelle (**C**).

A/ La responsabilité pénale des fournisseurs et utilisateurs de systèmes d'IA

73. **Les cas d'engagement de la responsabilité pénale des acteurs du cycle de vie d'un système d'IA.** Plusieurs situations dans le secteur de l'électricité pourraient conduire à l'engagement de la responsabilité pénale des entreprises utilisant des systèmes d'IA. Par exemple, il est possible qu'un système autonome de pilotage d'équipements électriques dans le

foyer présente un dysfonctionnement causant une surtension électrique et l'électrification d'un utilisateur. Si un tel défaut engendrait un décès ou des blessures pour la victime, les délits d'homicide ou de blessures involontaires pourraient alors être constitués[160]. Le dommage étant causé au moyen d'un système d'IA dont les buts et fonctionnalités ont été programmées par un humain, c'est bien à l'humain responsable du défaut de sécurité qu'incombera la responsabilité pénale des faits. De plus, en cas de dommage causé par un défaut de sécurité de l'intelligence artificielle, le producteur[161], le fabricant, l'importateur, le distributeur ou toute autre partie impliquée dans la chaîne d'approvisionnement de l'IA pourraient être condamnés pénalement pour délit de tromperie[162]. En effet, le défaut de sécurité pourrait être considéré comme une tromperie du contractant sur l'aptitude à l'emploi ou les risques inhérents à l'utilisation du produit. Enfin, d'autres délits en droit pénal pourraient trouver à s'appliquer à l'IA, tels que le délit de presse ou la manipulation de cours dans le secteur financier[163]. Ainsi, un défaut de sécurité ou une utilisation frauduleuse d'un système d'IA peut conduire à la réalisation de dommages et engager la responsabilité pénale des personnes impliquées dans sa conception et sa distribution.

74. **La subjectivité du droit pénal face à l'autonomie de la machine.** Il convient ici de rappeler que tous les délits sont intentionnels[164], sauf disposition contraire, notamment dans les cas évoqués précédemment des délits d'homicide ou blessures involontaires. Dès lors, certains considèrent que le droit pénal est incompatible avec l'IA, arguant de son absence de

---

[160] Code pénal, articles 221-6 à 221-7 pour le délit d'homicide involontaire et Code pénal, article 222-19 pour le délit de blessures involontaires.

[161] Selon l'article 1245-5 du Code civil, « *est producteur, lorsqu'il agit à titre professionnel, le fabricant d'un produit fini, le producteur d'une matière première, le fabricant d'une partie composante. Est assimilée à un producteur pour l'application du présent chapitre toute personne agissant à titre professionnel :*
1. Qui se présente comme producteur en apposant sur le produit son nom, sa marque ou un autre signe distinctif ;
2. Qui importe un produit dans la Communauté européenne en vue d'une vente, d'une location, avec ou sans promesse de vente, ou de toute autre forme de distribution. […] ».

[162] Code de la consommation, article L441-1. Aux termes de cet article, la tromperie peut porter sur « les risques inhérents à l'utilisation du produit » ainsi que « les contrôles effectués ». En particulier, sont visés les tromperies sur la conformité à des normes (Crim. 10 avr. 1997, n°96-82.183 P, *JCP*, 1997, IV, 1780 ; Crim. 2 avr. 2007, *Dr. pénal*, 2008, 1, obs. L. Jean.), la dangerosité (Crim. 22 juin 1994, n°93-83.900 P), ou sur l'absence de contrôles effectués (Paris, 23 avr. 1992, *CCC*, 1992, n°216, obs. G. Raymond ; Crim. 29 mai 1996, *Gaz. Pal.*, 1996, 2, 152, obs. J.P. Doucet).

[163] Pour une étude des délits de presse et de manipulation de cours, V. S. Merabet, *op. cit.*, p. 430, 459 et s.

[164] Y. Mayaud, *Droit pénal général*, PUF, 2018, 6e éd., n°232.

personnalité juridique et de sa capacité à réaliser des choix indépendamment des personnes qui l'ont créé ou en ont l'usage[165]. Toutefois, cette analyse repose sur le fait que l'IA réalise des choix autonomes[166]. Une vision qui nous apparaît erronée au vu de la réalité technique de l'IA aujourd'hui[167]. À partir du moment où le comportement d'une IA est explicable, interprétable ou *a minima* traçable, l'objectivité et l'autonomie de la machine ne seront plus un problème puisque l'on pourra remonter à l'origine du défaut ayant causé le dommage. Il nous semble plus pertinent de concentrer nos efforts sur les moyens pour imposer cette transparence, plutôt que sur la transformation du droit pénal. En effet, ce dernier semble disposé à traiter des situations où l'IA, y compris celle agissant avec un haut niveau d'autonomie, serait impliquée dans la commission d'infractions en imposant de remonter à l'humain dont la négligence a conduit à la réalisation du dommage.

B/ La responsabilité contractuelle pour une répartition consensuelle des responsabilités

75. **L'aménagement contractuel des responsabilités entre les personnes impliquées dans la conception d'un système d'IA.** Le régime de la responsabilité contractuelle pourra être appliqué en cas de dommage causé par l'IA entre le fournisseur et l'utilisateur s'ils sont liés par un contrat[168]. En réalité, cette voie est plus large car elle peut être étendue à toute personne qui intervient dans la chaîne d'approvisionnement du système, de l'élaboration à sa vente (le vendeur, l'éditeur, le concepteur, l'intégrateur, le sous-traitant…). Lorsque plusieurs acteurs interviennent dans la conception, ils pourront invoquer, le cas échéant, les clauses de garantie qu'ils auront pris le soin d'insérer dans les contrats les liant aux autres intervenants dans la chaine de responsabilités. La relation contractuelle entre les différents professionnels impliqués

[165] S. Merabet, *op. cit.*, p. 441, spec. 476.
[166] *Ibid.*, 59.
[167] Voir Supra, 5-7.
[168] Code civil, article 1217 ; Civ. 2ème, 15 mars 2018, nº 16-15.791.

dans la conception du système aura dès lors une importance majeure et requerra une attention particulière[169].

76. **La garantie des vices cachés du système d'IA et la garantie contractuelle.** Si l'IA présentait un défaut susceptible de compromettre l'utilisation que l'acheteur souhaitait en faire, ce dernier pourrait invoquer la garantie des vices cachés sur le fondement de l'article 1641 du Code civil. Toutefois, plusieurs conditions devraient être satisfaites : le système devrait présenter un vice d'une certaine gravité, nécessairement caché au moment de la vente, antérieur ou concomitant à celle-ci. Dans ce cas, le délai d'action est de deux ans à compter de la découverte du vice[170]. De plus, les directives européennes 2019/770[171] et 2019/771[172] du 20 mai 2019 sont venues préciser les règles applicables aux contenus et services numériques ainsi qu'aux objets connectés. La large définition adoptée dans ces textes laisse supposer que leurs dispositions seront pleinement applicables aux systèmes d'IA qu'ils soient immatériels ou incorporés dans un objet[173]. Elles confirment notamment le délai de deux ans et prévoient les remèdes en cas d'action d'un consommateur sur le fondement de la garantie légale de conformité : la mise en conformité du contenu numérique, la réduction du prix ou la résolution du contrat[174].

77. **Les limites de l'aménagement contractuel des responsabilités.** Toutefois, il est possible de délimiter ou d'exclure la garantie contre les vices à laquelle le vendeur s'engage.

---

[169] Sur les précautions à prendre lors de la contractualisation entre acteurs de la chaîne d'approvisionnement d'un système d'IA (développeur, commanditaire, éditeur,…), V. notamm. N. Quoy, A. Boullet, « Pratique contractuelle. L'encadrement contractuel de l'intelligence artificielle », *Communication Commerce électronique*, février 2020, n°2, 3 ; G. Courtois, « Robots intelligents et responsabilité : quels régimes, quelles perspectives ? », *Dalloz IP/IT*, 2016, 287.

[170] Code civil, article 1641 et s.

[171] *Directive (UE) 2019/770 du Parlement européen et du Conseil du 20 mai 2019 relative à certains aspects concernant les contrats de fourniture de contenus numériques et de services numériques*, publiée au JOUE n°L136/1 du 22 mai 2019.

[172] *Directive (UE) 2019/771 du Parlement européen et du Conseil du 20 mai 2019 relative à certains aspects concernant les contrats de vente de biens, modifiant le règlement (UE) 2017/2394 et la directive 2009/22/CE et abrogeant la directive 1999/44/CE*, publiée au JOUE n°L136/28 du 22 mai 2019.

[173] G. Loiseau, « Un droit de l'intelligence artificielle à l'échelle européenne en construction : analyse des deux récentes propositions de directives adaptant les règles de la responsabilité civile au développement de l'IA », *Le Club des Juristes (blog)*, 14 décembre 2022, disponible en ligne : <https://blog.leclubdesjuristes.com/un-droit-de-lintelligence-artificielle-a-lechelle-europeenne-en-construction-analyse-des-deux-recentes-propositions-de-directives-adaptant-les-regles-de-la-responsabilite-civile-au/>, consulté le 15 janvier 2023.

[174] C. Hélaine, « Adaptation de la garantie légale de conformité pour les biens et les contenus et services numériques », *Dalloz Actualité*, 5 octobre 2021.

Généralement le vendeur se limite à remplacer la chose viciée ou à la remettre en l'état et instaure un délai pendant lequel l'acquéreur peut en faire la demande. Il convient de rappeler que face au vendeur professionnel, l'acquéreur particulier ou professionnel (sauf si ce dernier est de la même spécialité que le vendeur) pourra toujours invoquer la garantie légale des vices si elle lui est plus favorable[175]. Ainsi, si les défauts étaient insusceptibles d'être délimités ou complétement éliminés par la clause de garantie contractuelle, ce serait la garantie légale qui aurait vocation à s'appliquer[176]. L'encadrement contractuel des responsabilités apparaît aujourd'hui comme le moyen le plus adéquat pour garantir la sécurité juridique de l'achat ou de la vente de systèmes d'IA, mais elle requiert une expertise juridique particulière, consciente des enjeux techniques et des risques que peuvent générer de tels systèmes pour les délimiter au mieux.

C/ La responsabilité civile délictuelle pour faute en cas de manquement au devoir de sécurité

78. **La faute pour manquement au devoir de sécurité.** En cas de dommage causé par un système d'IA, il serait tout d'abord envisageable d'invoquer la responsabilité extracontractuelle pour faute du fournisseur du système d'IA. Il est effectivement possible de caractériser la faute du professionnel ayant commercialisé le système d'IA sur le fondement de l'article 1240 du Code civil pour manquement à son devoir de sécurité, indépendamment de ses obligations contractuelles[177]. Ce raisonnement ne vaut qu'en présence d'un dommage causé par un vice dans la chose vendue, ce qui, en l'état actuel de la technique, couvrirait la majorité des dommages envisageables. Notons également que le fait de mentionner dans un contrat le risque qu'un défaut de sécurité survienne au cours de l'utilisation du système ne dédouanera pas le

[175] Cette faculté devant expressément être mentionnée dans le contrat conclue entre le vendeur professionnel et le consommateur (art. L.217-15, alinéa 4 du Code de la consommation).
[176] P. Le Tourneau, « Effets de la vente », in *Droit de la responsabilité et des contrats*, *op. cit.*, 3363.431 et s. ; O. Barret, P. Brun, « Vente : effets – Garanties contre les vices cachés », *Répertoire de droit civil*, 2020 655-658 ; Com. 30 mai 1967, Gaz. Pal., 1967, 2, 79.
[177] Civ. 3ème, 5 décembre 1972, D. 1973, 401, note J. Mazeaud : En matière de vente, le seul fait pour le fabricant de mettre sur le marché une chose viciée est constitutif d'une faute délictuelle envers les tiers ; Civ. 1ère, 17 janvier 1995, n°93-13.075 ; V. aussi l'obligation générale de sécurité à la charge des professionnels prévue à l'article L 421-3 du Code de la consommation.

fournisseur de ses responsabilités, qui relèvent ici d'un fondement délictuel. Ce constat ne vaut que dans le cas où la victime serait en mesure d'apporter la preuve de la faute, du préjudice, ainsi que du lien de causalité entre les deux, conformément aux dispositions de l'article 1240 du Code civil.

79. **L'ineffectivité de la responsabilité pour faute en dehors du cas du défaut de sécurité.** Dans le cas où le système d'IA causerait un dommage dans son fonctionnement normal et malgré la pleine diligence des concepteurs et fabricants du système, la faute pourrait ne pas suffire pour identifier un responsable et enclencher le mécanisme de réparation[178]. En effet, si son fonctionnement est « normal », le comportement de l'algorithme pourrait difficilement être considéré comme « fautif », et ce, qu'on le compare au comportement d'un humain ou au comportement d'algorithmes équivalents[179]. Dès lors, certains auteurs s'accordent sur l'absence d'effectivité de la responsabilité pour faute lorsqu'elle est confrontée à un dommage causé par le comportement normal d'une IA[180]. Ce régime, bien qu'il soit à même de couvrir les cas de manquement au devoir de sécurité, ne semble pas être applicable à tous les autres cas où des dommages pourraient être causés par un système d'IA dans le cadre d'un comportement normal.

80. **Transition.** En effet, l'existence d'une faute, qu'elle soit pénale, délictuelle ou contractuelle, permettrait de mobiliser les régimes de responsabilité décrits dans le présent paragraphe. Toutefois, cette notion ne semble pas adaptée aux spécificités de l'IA, si bien qu'il serait très difficile, pour la victime d'un préjudice causé par un système autonome, d'apporter la preuve d'un comportement fautif.

---

[178] P. Jourdain, « Faute délictuelle et manquement contractuel : des relations complexes. Illustration à travers les fautes délictuelles de l'entrepreneur et du mandataire », *Revue Trimestrielle de Droit civil*, 1995, 895.

[179] Sur les difficultés d'apprécier les contours du comportement potentiellement « fautif » d'un système algorithmique, voir J.S. Borghetti, « Civil Liability for Artificiel Intelligence : What Should Its Basis Be ? », *Revue des juristes de Sciences Po*, 2019, n°17, pp. 76-84.

[180] M. Bacache, « Intelligence artificielle et droits de la responsabilité et des assurances », in *Droit de l'intelligence artificielle*, dir. A. Bensamoun, G. Loiseau, *LGDJ*, coll. Les intégrales, 2019, p. 69 et s. ; S. Merabet, *op. cit.*, p. 485, 523 et s. ; G. Loiseau, M. Bourgeois, « Du robot en droit à un droit des robots », *JCP*, 2014, 1231.

§2 : La difficile démonstration de l'existence d'une faute

81. **L'appréciation de la faute par le juge en cas de préjudice causé par un système d'IA.** La faute est un standard objectif apprécié par le juge *in abstracto*[181], en comparant le comportement de l'agent ayant causé un préjudice avec le comportement qu'aurait adopté un autre agent de la même nature dans les mêmes circonstances. Toutefois, dans le cas de dommages causés par le fonctionnement d'une IA, faut-il analyser le comportement du professionnel opérant le système ou bien le comportement du système lui-même par rapport au comportement attendu d'un système « raisonnable » ? Le deuxième cas ne nous semble pas pertinent à deux égards. D'une part, il conduirait à rendre le logiciel lui-même responsable des dommages, ce qui n'est pas possible à défaut de lui reconnaître la personnalité juridique. D'autre part, si l'analyse conduisait à la caractérisation d'un écart entre le comportement attendu du système et le comportement réel ayant conduit à un préjudice, alors cet égard pourrait vraisemblablement permettre la qualification d'un défaut du produit. Une telle qualification permettrait la réparation du dommage au titre du régime de la responsabilité des produits défectueux prévu par l'article 1245-17 du Code civile. Il convient donc de se concentrer sur l'analyse du comportement du professionnel opérant le système d'IA. Nous avons déjà évoqué la possibilité de qualifier une faute en cas de manquement au devoir de sécurité. Malgré l'absence de jurisprudence en la matière, d'autres fondements pour caractériser une faute sont envisageables.

82. **Vers une interprétation élargie de la faute ?** Outre le devoir de sécurité, trois fondements supplémentaires semblent mobilisables pour engager la responsabilité d'un professionnel opérant un système d'IA[182]. Premièrement, le développement de sources de droit souple en matière d'IA pourrait conduire le juge à consacrer de nouvelles normes éthiques dans sa jurisprudence. Ainsi, une entreprise ayant publiquement adhéré à des principes éthiques en matière d'IA, tels que ceux contenus dans les lignes directrices pour une IA digne de confiance[183], pourrait voir sa responsabilité engagée si le juge venait à considérer, par exemple,

[181] M. Bacache, « Intelligence artificielle et droits de la responsabilité et des assurances », in *Droit de l'intelligence artificielle*, dir. A. Bensamoun, G. Loiseau, LGDJ, coll. Les intégrales, 1ère ed., 2019, 113.
[182] *Ibid*.
[183] GROUPE D'EXPERTS DE HAUT NIVEAU EN IA, *Lignes directrices du 8 avril 2019 pour une IA digne de confiance*, 8 avril 2019.

que le non-respect d'un principe de contrôle humain ou de non-discrimination avait conduit à des préjudices. Les entreprises doivent donc prêter attention au développement de la *soft law* en matière d'IA puisque ces textes pourraient par la suite devenir de véritables sources de responsabilité. Deuxièmement, certains considèrent que la faute de précaution, consacrée par le juge[184], pourrait conduire à une nouvelle source de responsabilité pour les personnes privées[185]. Son application aux dommages physiques ou matériels causés par le fonctionnement de systèmes d'IA ne nous semble pas opportun en raison de l'existence de régimes de responsabilité plus adaptés, notamment ceux décorrélés de la notion de faute. Les contours flous du principe de précaution rendraient encore plus difficile la caractérisation d'une faute. Troisièmement, le devoir de vigilance, consacré par la jurisprudence civile[186] et constitutionnelle[187] ainsi que par la loi[188], impose notamment aux entreprises de prendre des mesures pour prévenir les risques d'atteinte aux droits fondamentaux. Dès lors, le défaut de mise en œuvre de mesures suffisantes pour garantir l'absence de biais discriminatoires ou de traitements intrusifs de données à caractère personnel par un système d'IA pourrait potentiellement être caractérisé comme une faute de vigilance. L'ensemble des fautes exposées dans le présent paragraphe ne sont qu'hypothétiques, en l'absence de jurisprudence en la matière. Toutefois, les fondements existent, encore faut-il pouvoir les mobiliser en prouvant l'existence d'une faute.

83. **L'impossible preuve de la faute par la victime.** L'engagement de la responsabilité d'une entreprise opérant un système d'IA se heurte à une difficulté majeure. La victime d'un dommage ne disposera probablement pas des éléments de preuve nécessaires pour démontrer l'existence des fautes décrites précédemment. En effet, comment une victime pourrait-elle démontrer qu'une entreprise a manqué à son devoir de vigilance ou de sécurité dans la conception d'un système d'IA ? Pour ce faire, elle aurait besoin d'informations détaillant les

[184] Cass. 3ème civ., 3 mars 2010, n°08-19108.
[185] M. Boutonnet, *Le principe de précaution en droit de la responsabilité*, LGDJ, 2005 ; G. Viney, « Le principe de précaution et la responsabilité civile des personnes privée », *Recueil Dalloz*, 2007, 1542 ; cités dans M. Bacache, *op. cit.*, 112.
[186] Cass. 1ère civ., 7 mars 2006, n°04-16179.
[187] Cons. constit., 8 avril 2011, n°2011-116 QPC.
[188] Loi n°2017-399 du 27 mars 2017 relative au devoir de vigilance des sociétés mères et des entreprises donneuses d'ordre.

modalités de conception et de fonctionnement du système. Or, d'une part, ces éléments de preuve ne sont pas suffisamment accessibles pour la victime et, d'autre part, l'entreprise elle-même ne sera pas forcément en mesure de détailler la façon dont le système a été conçu. En effet, dans le cas du recours à des techniques d'apprentissage profond (*Deep learning*), même des experts ne seraient pas en mesure d'expliquer précisément les raisons qui ont conduit le logiciel à produire un résultat précis (en l'espèce, le résultat ayant conduit à la réalisation d'un dommage). Les difficultés liées à la preuve sont donc proportionnelles aux degrés de complexité, d'autonomie et d'opacité du système d'IA. Or, à défaut de pouvoir prouver l'existence d'une faute ou le lien de causalité entre un fait fautif et son préjudice, la victime ne pourrait obtenir réparation.

84. **Transition.** Les régimes de responsabilité fondés sur la notion de faute connaissent plusieurs lacunes lorsqu'ils sont appliqués à des dommages causés par des systèmes d'IA. Plus que la notion de faute en elle-même, c'est avant tout la question de la preuve qui conduirait à l'absence de réparation de dommages causés par une IA. Dans les situations où les logiciels sont les plus complexes et autonomes, les régimes de responsabilité sans faute semblent contenir des dispositions plus appropriées.

Section 2 : L'applicabilité discutée des régimes de responsabilité sans faute

85. **La pertinence des régimes de responsabilité sans faute.** La complexité, l'autonomie et l'opacité du fonctionnement de certains systèmes d'IA peuvent rendre difficiles la détermination de l'origine du dommage, la caractérisation de la faute ou l'identification du responsable. C'est la raison pour laquelle les régimes de responsabilité dits « sans faute » apparaissent comme les plus adaptés aux préjudices causés par des systèmes d'IA.

86. **Plan.** Premièrement, le régime de l'article 1242 du Code civil appelé communément « responsabilité du fait des choses » peut être appliqué à l'IA qui relève de cette qualification (**§1**). Les modalités de son application présentent néanmoins quelques incertitudes théoriques notamment au regard de la notion de « garde » de la chose. Incertitudes théoriques seulement car le concept semble suffisamment malléable pour y être transposé. Deuxièmement, le Code civil organise également la responsabilité du fabricant de produits défectueux. La qualification de l'IA en tant que produit, de la défectuosité du système et les nombreuses exceptions pouvant être mobilisées pour y échapper ont été sujets à d'importants débats en doctrine (**§2**).

§1 : La responsabilité du fait des choses

87. **La responsabilité du fait des choses intelligentes.** La potentielle autonomie des systèmes d'IA ne nous semble pas remettre en cause de façon insurmontable les concepts fondant la responsabilité du fait des choses. L'article 1242 du Code civil prévoit que l'on est responsable du dommage causé par les choses que l'on a sous sa garde. En application de l'adage selon lequel il convient de ne pas distinguer là où la loi ne distingue pas, les systèmes d'IA, en tant que choses[189], entrent bien dans le champ d'application. Seule l'applicabilité de la notion de garde est encore sujette à débat[190].

[189] Sur la qualification des systèmes d'IA en tant que chose : voir Infra, 95.

[190] S. Migayron, « Pratique contentieuse : Intelligence artificielle : qui sera responsable ? », *Communication commerce électronique*, avril 2018, n°4 ; COUR D'APPEL DE PARIS, *La réforme du droit français de la responsabilité civile et les relations économiques*, Rapport du groupe de travail établi avec l'Université de Versailles Saint-Quentin-En-Yvelines, avril 2019, p. 107 et s.

88. **La garde de la chose intelligente conçue et/ou supervisée par l'homme.** La garde est un pouvoir de fait exercé sur la chose et ayant pour objet l'usage, le contrôle et la direction de cette chose[191]. Il nous semble que cette notion est applicable à l'IA puisqu'elle est une chose sur laquelle l'utilisateur a un contrôle (fusse-t-il à distance) dans la majorité des cas. La différence majeure est que la garde sera avant tout intellectuelle dans la mesure où ce sont les instructions programmées par l'utilisateur qui permettront de mettre en œuvre les critères d'usage, de direction et de contrôle de l'IA : la garde physique du XXème siècle et de la révolution industrielle devient aujourd'hui une garde intellectuelle. Le gardien sera donc responsable des dommages causés par l'IA dont il a la garde. Les cas où l'IA serait pleinement autonome, définirait elle-même ses propres règles de conduite et ne serait ni dirigée ni contrôlée, même à distance et en différé, par un humain nous semblent bien loin de la réalité[192] et ne feront donc pas l'objet de plus amples développements ici.

89. **L'absence de rôle actif du gardien dans la réalisation du dommage.** Par ailleurs, il est à noter que l'IA peut avoir comme particularité d'être en mouvement et de pouvoir entrer en contact avec la victime. Le rôle actif du gardien dans la réalisation du dommage n'est pas requis en l'état de la jurisprudence, puisque la « garde » n'est qu'un « pouvoir de fait »[193]. Néanmoins, ce dernier doit présenter une certaine « autonomie »[194]. Si ce pouvoir de fait et cette autonomie du gardien ne pouvait être qualifiés, la distinction classique entre « la garde de la structure » et « la garde du comportement »[195] pourrait être ressuscitée. Cette distinction est utilisée lorsqu'un dommage est causé par une chose mue par un dynamisme propre et dangereux[196]. L'IA pourrait être perçue comme telle. Dans ce cas, la garde du comportement serait attribuée à l'utilisateur et la garde de la structure au fabricant ou fournisseur faisant peser sur lui un risque qu'il devrait prendre en compte dès la conception du système. Le « gardien » du système pourra donc être,

---

[191] Civ. 2ème, 10 février 1982, n° 81-40.495, *JCP*, 1983, II, 20069, obs. A. Cœuret.
[192] Voir Supra,5-7.
[193] J. Julien et P. Le Tourneau, « Responsabilité générale du fait des choses » in *Droit de la responsabilité et des contrats*, dir. P. Le Tourneau, Dalloz action, 2021, 2221.151 et s.
[194] Civ. 3ème, 20 octobre 1971, n°70-13.035, n°505, D. 1972, 414, obs. C. Lapoyade-Deschamps.
[195] Civ. 2ème, 5 janvier 1956, GAJC, t. II, 13e éd., 2015, n°205, D. 1957, 261, note R. Rodière; JCP 1956. II. 9095, note R. Savatier ; voir aussi P. Dupichot, *La garde de la structure et la garde du comportement dans la responsabilité civile*, thèse pour le doctorat en droit privé, Université Paris XII, 1984 ; A. Tunc, « Garde du comportement et garde de la structure dans la responsabilité du fait des choses inanimées », *JCP* 1957. I. 1384.
[196] Pour des illustrations de choses entrant dans cette qualification, voir J. Julien et P. Le Tourneau, *op. cit.*, 2221-220 – 227.

selon les cas, le fournisseur, le fabricant ou l'utilisateur. Néanmoins, puisque selon l'adage « le spécial déroge au général », il convient de s'intéresser au régime spécial de la responsabilité du fait de produits défectueux, qui pourrait également trouver à s'appliquer dans le cas d'un dommage causé par une IA, indépendamment de la présence d'un superviseur ou gardien humain.

§2 : La responsabilité du fait des produits défectueux

90. **La responsabilité du producteur du fait d'une IA défectueuse.** En application de l'article 1245-17 du Code civil, le producteur est responsable du dommage causé par un défaut de son produit, qu'il soit ou non lié par un contrat avec la victime. La notion de « produit » est entendue largement comme tout bien meuble[197]. À priori, rien n'empêche donc d'y inclure un système d'IA qui, on l'a vu, est bien une chose malgré sa nature composite (logiciel, potentielle incarnation dans une enveloppe physique, capacité d'action sur des composants matériels). Toutefois, l'extension de la notion de produit au domaine de l'immatériel et l'opportunité de la graver dans le marbre de la loi ont fait débat[198]. En effet, les textes n'excluent pas expressément le logiciel de la responsabilité du fait des produits défectueux. Il n'est donc pas forcément nécessaire d'amender le corpus existant pour appliquer ce régime aux logiciels fondés sur des techniques d'IA. Le Ministère de la Justice a cependant précisé que *« les seuls dommages dont ladite loi assure la réparation sont les atteintes physiques et les dommages matériels causés aux biens »* [199]. L'application du régime des produits défectueux aux logiciels ne viserait donc, en l'état actuel des textes, que les situations où ceux-ci seraient à l'origine d'une atteinte directe à la sécurité physique des personnes et des biens.

---

[197] Code civil, article 1245-5 : « *Est un produit tout bien meuble, même s'il est incorporé dans un immeuble, y compris les produits du sol, de l'élevage, de la chasse et le pêche. L'électricité est considérée comme un produit* » ; V. aussi M. Poumarède, P. Le Tourneau, « Domaine de la responsabilité », in *Droit de la responsabilité et des contrats, op. cit.*, 6312.21.

[198] S. Migayron, *op. cit.* ; COUR D'APPEL DE PARIS, *La réforme du droit français de la responsabilité civile et les relations économiques*, Rapport du groupe de travail établi avec l'Université de Versailles Saint-Quentin-En-Yvelines, avril 2019, spec. p.108 : « *[…] Une extension expresse à l'article 1243 du projet de réforme du champ d'application de cette responsabilité aux choses incorporelles […] ne serait pas neutre au regard notamment de son application potentielle à l'information* ».

[199] Réponse ministérielle n°15677 sur l'applicabilité aux logiciels de la *loi n°98-389 du 19 juin 1998 relative à la responsabilité du fait des produits défectueux*, publiée au JO du 24 août 1998, p. 4728.

91. **La défectuosité de l'IA comme critère d'attribution de la responsabilité.** La notion de « produit défectueux » désigne le produit qui n'offre pas la sécurité à laquelle on peut légitimement s'attendre[200]. La définition de la défectuosité apparaît donc à première vue suffisamment large pour englober les comportements et actions non anticipés par le concepteur d'un système d'IA, qui s'apparenteraient à des défauts remettant en cause la sécurité du produit vis-à-vis des tiers. Caractériser la défectuosité s'avèrera parfois une tâche très compliquée : une IA peut très bien générer un dommage voire commettre une « faute » ponctuelle tout en ayant un fonctionnement efficace et utile dans le reste de son utilisation. Sans oublier les dommages que pourrait causer une IA résultant de son fonctionnement normal, sans défectuosité, simplement en conséquence des données d'entrée dans le système. On s'aperçoit alors que lorsque l'on s'intéresse au détail des conditions juridiques d'engagement de la responsabilité sur ce fondement, l'application de ces critères n'est pas si évidente et peut conduire à des incertitudes. Toutefois, le fait que l'analyse se concentre sur « la sécurité à laquelle on peut légitimement s'attendre » et non sur un défaut de conformité permet d'argumenter en faveur de l'inclusion des dommages causés malgré un fonctionnement « sans défaut » de l'algorithme.

92. **Des limites de la responsabilité du fait des produits défectueux face au dommage causé par un système d'IA.** Il convient de noter que plusieurs causes d'exonération pourraient être invoquées par le producteur : l'absence de commercialisation du produit (en cas de vol par exemple), l'absence de défaut d'origine au moment de la commercialisation et la démonstration que le défaut est dû à une usure excessive en raison d'une mauvaise utilisation par l'utilisateur (justifiant l'attention à porter aux notices d'utilisation), ou encore le risque de développement[201]. En particulier, ce dernier cas d'exonération risque d'avoir une application trop étendue[202] puisqu'il suffirait au producteur de démontrer qu'il ne pouvait pas déceler le défaut au moment où le système a été mis en circulation, puisqu'il serait par exemple dû à l'apprentissage de l'IA après sa commercialisation (pour rappel, l'apprentissage post-conception ne concerne pas qu'une part infime des solutions d'IA actuellement sur le marché).

---

[200] Code civil, article 1245-3 : « *Un produit est défectueux au sens du présent chapitre lorsqu'il n'offre pas la sécurité à laquelle on peut légitimement s'attendre* » ; V. aussi P. Le Tourneau, M. Poumarède, « Régime de la responsabilité », in *Droit de la responsabilité et des contrats*, *op. cit.*, 6313.31.
[201] J. Julien, P. Le Tourneau, *op. cit.*, 2221.251 et s.
[202] G. Loiseau, M. Bourgeois, *op. cit.*

Outre les cas d'exonération, on peut résumer les limites de ce régime en citant les travaux de Laurène Mazeau, présentant les objections auxquelles le recours à l'article 1245 du Code civil pourrait se heurter : l'identification de la défectuosité au moment de la commercialisation, l'établissement du lien de causalité entre le dommage et le défaut et plus globalement l'application de ce régime aux biens immatériels[203]. Ces obstacles ne reflètent pas complètement la réalité technique des systèmes d'IA. D'abord, contrairement aux croyances populaires autour de l'IA, la majorité des systèmes n'apprennent que dans le cadre de leur conception. Une fois le modèle d'inférence construit, il est généralement testé et validé avant son utilisation. Il n'évoluera généralement qu'après une phase de réapprentissage ou des mises à jour. Les cas d'auto-apprentissage non supervisé sont extrêmement marginaux. Ensuite, il nous semble que si les systèmes d'IA étaient suffisamment explicables et leur conception documentée (ce qui est l'objectif de la règlementation européenne en construction[204]), la recherche de l'origine du défaut et la preuve de la défectuosité seraient grandement facilitée. Enfin, les incertitudes résiduelles peuvent être levées soit par une clarification jurisprudentielle (par exemple son applicabilité aux « produits » immatériels), soit par une intervention législative.

93. **Transition.** Toutefois, devant le constat des lacunes des régimes de responsabilité existants, il est pertinent de s'interroger sur l'opportunité de les réformer pour clarifier les modalités de leur application à l'IA.

[203] L. Mazeau, « Intelligence artificielle et responsabilité civile : le cas des logiciels d'aide à la décision en matière médicale », *Revue pratique de la prospective et de l'innovation*, avril 2018, n°1, 6.
[204] COMMISSION EUROPÉENNE, *Proposition de règlement du Parlement européen et du Conseil établissant des règles harmonisées concernant l'intelligence artificielle (législation sur l'intelligence artificielle) et modifiant certains actes législatifs de l'Union*, 21 avril 2021, 2021/0106 (COD).

Section 3 : Une réforme souhaitable des régimes de responsabilité

94. **Plan.** Quelles sont les solutions juridiques pour assurer la juste réparation des préjudices causés par l'action de systèmes d'IA ? D'une part, certains ont souhaité doter l'IA de la personnalité juridique. Cette solution aurait permis d'appliquer les principes de la responsabilité à l'IA de façon indifférenciée et donc de ne pas réformer les règles existantes. Toutefois, cette piste ne nous semble pas opportune (**§1**). D'autre part, il est possible de réviser les régimes de responsabilité ou d'en créer de nouveaux adaptés aux spécificités de l'IA, une voie empruntée par les institutions européennes (**§2**).

§1 : L'inopportune reconnaissance de la personnalité juridique de l'IA

95. **L'inadéquation de la qualification de l'IA comme sujet de droit pour résoudre les questions de responsabilité.** Pour répondre aux incertitudes relatives à l'attribution de la responsabilité en cas de dommage causé par l'action d'un système d'IA, certains ont proposé la reconnaissance de la personnalité juridique aux robots les plus avancés. Cette proposition reviendrait à qualifier ces systèmes de sujets de droit plutôt que d'objets de droit, une solution vivement critiquée en doctrine. En effet, un houleux débat a opposé les subjectivistes, plaidant pour la reconnaissance d'une personnalité juridique du robot et de l'IA[205] et les objectivistes, arguant qu'une telle mesure troublerait fortement les catégories juridiques existantes[206]. Les arguments des seconds semblent aujourd'hui emporter la conviction et faire consensus. L'excellente analyse du Professeur Samir Merabet sur la question en fait la synthèse en ces termes : « *S'il existe un intérêt théorique à qualifier l'intelligence artificielle de sujet de droit, en pratique, il s'avère que cette hypothèse ne résout aucun des problèmes juridiques provoqués par l'intelligence artificielle. Au contraire, elle est de nature à susciter des interrogations plus*

[205] A. Bensoussan, J. Bensoussan, *Droit des robots*, Larcier, 2015, p. 47 ; une proposition reprise par le Parlement européen dans une résolution de 2017 : *Résolution 2015/2103(INL) du Parlement européen du 16 février 2017 contenant des recommandations à la Commission concernant des règles de droit civil sur la robotique*, publiée au JOUE n°C252/239 le 18 juillet 2018.

[206] A. Bensamoun, G. Loiseau, « L'intégration de l'intelligence artificielle dans l'ordre juridique en droit commun : questions de temps », *Dalloz IP/IT*, 2017, 239.

importantes. Si des alternatives à la personnalité ont été proposées en doctrine, elles ne semblent pas plus adaptées à l'objet de la présente étude. Par conséquent, il convient d'écarter définitivement cette voie »[207]. Peu d'auteurs ont depuis remis en cause cette position. Le Parlement européen, qui s'était prononcé en faveur d'une qualification subjective de l'IA en 2017, semble finalement s'être rangé du côté des objectivistes dans sa recommandation de 2020 sur un régime de responsabilité civile pour l'IA[208]. Dès lors, il semble peu pertinent d'envisager cette solution, tant les arguments en sa faveur sont faibles.

96. **Transition.** Cette proposition étant mise de côté, il convient d'étudier les autres solutions pour répondre aux limites des régimes de responsabilité présentées dans la Section précédente. À ce titre, les institutions européennes ont adopté une voie radicale : celle de l'adaptation des textes en vigueur et de la création d'un régime de responsabilité *ad hoc*.

§2 : Une réforme initiée au niveau européen

97. **Un courant doctrinal en faveur de la réforme des régimes de responsabilité.** Plusieurs auteurs ont étudié les limites des régimes existants et démontré que des concepts clés de la responsabilité pouvaient être remis en cause par les spécificités des systèmes d'IA[209]. En particulier, ce sont son immatérialité, sa complexité structurelle, son opacité et sa potentielle autonomie qui préoccupent le civiliste. Il peut en effet s'avérer difficile de qualifier des notions comme la faute, le défaut ou le produit lorsqu'une IA est concernée[210]. Les auteurs arguent du « défaut d'objectivité » des régimes de responsabilité, créés pour la subjectivité humaine et

[207] S. Merabet, *op. cit.*, spec. p.144, 147 ; sur la qualification de l'IA en tant que chose et non en tant que personne, et l'inopportunité de lui consacrer une personnalité juridique V. A. Bensamoun, G. Loiseau. « L'intégration de l'intelligence artificielle dans l'ordre juridique en droit commun : questions de temps », *op. cit.* ; G. Loiseau, *Le droit des personnes*, Ellipses, 2016, p. 74 ; J. R. Binet, « Personnalité juridique des robots : une voie à ne pas suivre », *Revue Droit de la famille*, LexisNexis, 2017, n°6, p. 2.

[208] *Résolution 2020/2014(INL) du Parlement européen du 20 octobre 2020 portant recommandations à la Commission sur un régime de responsabilité civile pour l'intelligence artificielle*.

[209] J.S. Borghetti, « Civil Liability for Artificiel Intelligence : What Should Its Basis Be ? », *Revue des juristes de Sciences Po*, 2019, n°17, pp. 76-84 ; J. Pouget, *op. cit.*

[210] Pour une étude des limites des régimes de responsabilité face à l'IA, V. notamm. J.S. Borghetti, « L'accident généré par l'intelligence artificielle autonome », *Actes du colloque du Master 2 Droit privé général et du laboratoire de droit civil*, in *JCPG*, 2017, n° spécial ; L. Mazeau, « Intelligence artificielle et responsabilité civile : le cas des logiciels d'aide à la décision en matière médicale », *Revue pratique de la prospective et de l'innovation*, Avril 2018, n°1, 6.

donc inadaptés à la machine[211]. Considérant que « *La chose intelligente est si innovante qu'elle bouleverse les instruments juridiques ancestraux* [de la responsabilité] »[212], ces derniers cherchent alors à théoriser un droit de la responsabilité propre à l'IA[213]. Une perspective embrassée par le Parlement européen qui l'envisageait déjà en 2017[214] et a eu l'occasion de renouveler sa volonté de voir créé un nouveau régime de responsabilité dans sa résolution d'octobre 2020 sur un régime de responsabilité civile pour l'IA[215]. Cette volonté a été concrétisée par la Commission européenne en septembre 2022 par la publication de ses propositions de directives européennes relatives à l'adaptation des règles en matière de responsabilité civile extracontractuelle au domaine de l'intelligence artificielle[216] et à la responsabilité des produits défectueux[217]. Ces deux propositions ont initié le processus législatif mais n'ont pas encore été adoptées.

98. **Le futur régime européen de responsabilité civile extracontractuelle en matière d'IA.** Le premier texte proposé par la Commission européenne est relatif aux règles en matière de responsabilité civile extracontractuelle. Elle fait le choix, en rupture avec la position antérieure du Parlement européen, d'un régime de responsabilité pour faute et instaure plusieurs mécanismes de présomptions afin d'alléger la charge de la preuve. Ce choix est surprenant puisque cet allègement est justement la raison d'être des régimes de responsabilité sans faute. Premièrement, la proposition contient une disposition permettant à un demandeur d'obtenir la divulgation forcée d'éléments de preuve par le défendeur concernant un système d'IA soupçonné d'avoir causé un dommage[218]. Toutefois, le juge destinataire de la demande devra

---

[211] S. Merabet, *op. cit.*, p. 451, 477 et s.
[212] *Ibid.*, p. 486, 513.
[213] *Ibid.* ; V. aussi J.S. Borghetti, « Civil Liability for Artificiel Intelligence : What Should Its Basis Be ? », *Revue des juristes de Sciences Po*, 2019, n°17, pp. 76-84 ; J. Pouget, *op. cit.*
[214] *Résolution 2015/2103(INL) du Parlement européen du 16 février 2017 contenant des recommandations à la Commission concernant des règles de droit civil sur la robotique*, publiée au JOUE n°C252/239 le 18 juillet 2018.
[215] *Résolution* 2020/2014(INL) *du Parlement européen du 20 octobre 2020 portant recommandations à la Commission sur un régime de responsabilité civile pour l'intelligence artificielle*.
[216] *Proposition 2022/0303(COD) du 28 septembre 2022 de Directive du Parlement européen et du Conseil relative à l'adaptation des règles en matière de responsabilité civile extracontractuelle au domaine de l'intelligence artificielle, (Directive sur la responsabilité en matière d'IA)*, COM(2022) 496 final, ci-après la « Directive sur la responsabilité en matière d'IA ».
[217] *Proposition 2022/0302 (COD) du 28 septembre 2022 de Directive du Parlement européen et du Conseil relative à la responsabilité des produits défectueux*, COM(2022) 495 final, ci-après la « Directive amendée sur la responsabilité des produits défectueux ».
[218] Directive sur la responsabilité en matière d'IA, article 3.

apprécier *« la plausibilité d'une action en réparation »*[219] sur la base des éléments de faits présentés par le demandeur. Une telle disposition soulagerait effectivement la charge de la preuve pour les victimes, sous réserve que ces dernières soient déjà en possession de premiers éléments suffisants pour convaincre le juge de donner droit à la demande[220]. Le défendeur qui ne divulguerait pas les informations pertinentes à la suite d'une injonction serait présumé fautif à son devoir de vigilance (présomption simple). Deuxièmement, la directive proposée crée une présomption réfragable d'un lien de causalité en cas de faute démontrée ou présumée au titre des dispositions susmentionnées[221]. Cette présomption est néanmoins aussitôt limitée puisqu'elle sera renversée dans le cas où le défendeur démontrerait que la victime pourrait prouver le lien de causalité en accédant à une expertise ou à des éléments de preuve suffisants[222]. L'effectivité de la présomption s'en trouve entachée tant il semble aisé de refuser à la victime son bénéfice en lui imposant de recourir à une coûteuse expertise[223]. Troisièmement, la proposition clarifie les acteurs contre lesquels une victime peut engager une action en responsabilité. Son article 4 distingue en effet entre les actions à l'encontre des fournisseurs de systèmes d'IA et contre les actions contre les utilisateurs de systèmes d'IA[224]. Il détaille les faits générateurs de responsabilité pour chacun, ce qui est bienvenu. On notera à ce titre que la responsabilité de l'utilisateur d'un système d'IA peut voir sa responsabilité présumée s'il ne s'est pas conformé aux instructions d'usage ou s'il a exposé le système à des données d'entrée non pertinentes[225]. En somme, la proposition contient des dispositions très utiles censées faciliter les actions en réparation pour les victimes, notamment en leur permettant d'accéder plus facilement à des éléments de preuve. Toutefois, les nombreuses exceptions aux présomptions prévues dans le texte nous font craindre que son effectivité soit en pratique très limitée. Le même constat peut être fait au sujet du deuxième texte proposé par la Commission

[219] Directive sur la responsabilité en matière d'IA, article 3, 1, al. 2.
[220] M. Bernelin, « Intelligence artificielle : une proposition de directive sur la responsabilité civile extracontractuelle », *Dalloz actualité*, 22 novembre 2022.
[221] Directive sur la responsabilité en matière d'IA, article 4, 1.
[222] *Ibid.*, article 4, 4.
[223] M. Bernelin, *op. cit.*
[224] Directive sur la responsabilité en matière d'IA, article 4, §2 et §3.
[225] *Ibid.*, §3, a) et b).

européenne en septembre 2022, portant sur le régime de la responsabilité des produits défectueux.

99. **L'amendement du régime européen de la responsabilité des produits défectueux.** La seconde proposition de la Commission amende la directive relative à la responsabilité (sans faute) des produits défectueux. Tout d'abord, la directive élargit la notion de produit aux logiciels et le champ des dommages réparables aux dommages causés aux biens ainsi qu'aux pertes de données[226]. Ensuite, le texte précise la notion de défectuosité pour l'appliquer aux systèmes d'IA fondés sur des techniques d'apprentissage automatique. En effet, il est précisé que l'appréciation de la défectuosité doit tenir compte de *« l'effet sur le produit de toute capacité à poursuivre son apprentissage après le déploiement »*[227]. Le fait qu'une IA apprenne au cours de son cycle de vie ne serait ainsi plus une cause d'exonération de responsabilité. Enfin, la proposition de révision contient un allègement de la charge de la preuve en présence de produits complexes et permettrait aux demandeurs d'obtenir la divulgation d'informations de la part du fabricant sur le même modèle que la Directive sur la responsabilité en matière d'IA[228]. Les dispositions des deux textes risquent de se heurter aux mêmes difficultés de mise en œuvre en raison, notamment, de leurs trop nombreuses exceptions permettant de renverser les présomptions de faute ou de lien de causalité. En somme, leur effectivité sera conditionnée à l'interprétation des juges : quel est le standard de preuve requis pour donner droit les demandes de divulgation d'informations ? Dans quelle mesure les juges permettront-ils aux fournisseurs de renverser la charge de la preuve ? Les juges et les victimes disposeront-elles des compétences nécessaires à l'analyse des informations divulguées pour démontrer l'existence d'une faute ? Il conviendra, pour les entreprises développant des systèmes d'IA, de

[226] Directive amendée sur la responsabilité des produits défectueux, article 4 ; S. Lemarchand, J. Dauzier, A. Pons, *et al.*, « La Commission européenne publie deux propositions de directive en matière de responsabilité applicables aux produits basés sur l'intelligence artificielle », Site du cabinet DLA PIPER (blog), disponible en ligne : <https://www.dlapiper.com/fr-fr/insights/publications/2022/10/european-commission-publishes-two-proposals-for-a-directive>, consulté le 15 janvier 2023.
[227] Directive amendée sur la responsabilité des produits défectueux, article 6, 1, c).
[228] G. Loiseau, « Un droit de l'intelligence artificielle à l'échelle européenne en construction : analyse des deux récentes propositions de directives adaptant les règles de la responsabilité civile au développement de l'IA », *Le Club des Juristes (blog)*, 14 décembre 2022, disponible en ligne : <https://blog.leclubdesjuristes.com/un-droit-de-lintelligence-artificielle-a-lechelle-europeenne-en-construction-analyse-des-deux-recentes-propositions-de-directives-adaptant-les-regles-de-la-responsabilite-civile-au/>, consulté le 15 janvier 2023.

suivre les évolutions de ces textes ainsi que les premières jurisprudences d'application qui seront sans nul doute riches d'enseignements.

100. **Conclusion du Chapitre 1 sur la nécessaire clarification des régimes de responsabilité.** À partir d'une littérature déjà abondante sur le sujet, le présent Chapitre a permis de montrer que les régimes de responsabilité sont en grande partie applicables aux dommages causés par un défaut dans un système d'IA ou par son fonctionnement normal[229]. En revanche, il persiste un certain nombre d'incertitudes, notamment quant à la répartition des responsabilités, qui peuvent faire naître des réticences dans la chaîne d'approvisionnement de l'IA. C'est la raison pour laquelle une attention particulière doit être portée à la rédaction des contrats lors de l'achat, le développement ou la fourniture de systèmes basés sur de l'IA. En effet, la responsabilité contractuelle semble être, à ce jour, le meilleur moyen de maîtriser la chaîne des responsabilités et d'aboutir à une répartition consensuelle, sous réserve que les dommages couverts ne soient pas des dommages corporels. Les incertitudes soulevées ont conduit à l'émergence de réflexions sur l'opportunité de la création d'une responsabilité du fait de la chose intelligente, défendue par plusieurs auteurs dans la doctrine et par le Parlement européen[230]. Pourtant, cette idée n'emporte pas la conviction de tous[231]. La majorité des incertitudes nées de l'application des régimes de responsabilité civile à l'IA sont dues à l'opacité des algorithmes, complexifiant l'identification de l'origine d'un potentiel défaut, ou à la potentielle autonomie de l'IA, entravant l'application de notions telles que la « faute » ou la « garde ». Avant de modifier des règles de responsabilité qui ont fait leur preuve ou d'en créer de nouvelles, il convient selon nous de se concentrer sur les usages de l'IA qui échapperaient à leur application : l'utilisation d'algorithmes opaques, sans aucune traçabilité ni explicabilité du

[229] D. Galbois-Lehalle, « Responsabilité civile pour l'intelligence artificielle selon Bruxelles : une initiative à saluer, des dispositions à améliorer », *Recueil Dalloz*, 2021, p. 87 ; CONSEIL D'ÉTAT, *Révision de la loi de bioéthique : quelles options pour demain ?*, Etude à la demande du Premier Ministre par la section du rapport et des études, 28 juin 2018, spec. pp. 192-210.
[230] V. notamm. S. Merabet, *op. cit.*, spec. p. 473 et s. ; J. Pouget, *op. cit.* ; *Résolution 2020/2014(INL) du Parlement européen du 20 octobre 2020 portant recommandations à la Commission sur un régime de responsabilité civile pour l'intelligence artificielle*.
[231] D. Galbois-Lehalle, « Responsabilité civile pour l'intelligence artificielle selon Bruxelles : une initiative à saluer, des dispositions à améliorer », *Recueil Dalloz*, 2021, p. 87.

comportement, ou complètement autonomes et auto-apprenants, sans supervision humaine. Toutefois, cela ne semble pas être la voie empruntée par la Commission européenne qui a publié en septembre 2022 deux propositions de directives européennes relatives à l'adaptation des règles en matière de responsabilité civile extracontractuelle au domaine de l'intelligence artificielle[232] et à la responsabilité des produits défectueux[233]. Ce faisant, elle espère faciliter la réparation des préjudices causés par des systèmes d'IA mais l'effectivité des textes proposés reste très discutable. Le projet de réforme va désormais faire l'objet de longues discussions au sein des institutions de l'Union européenne. Il conviendra de s'assurer que ces directives, dans leurs évolutions futures, ne viennent pas troubler plus que nécessaire les régimes de responsabilité français, en grande partie adaptables à l'IA.

101. **Transition.** Les régimes de responsabilité ne sont pas les seules règles de droit dont l'application est source d'incertitudes ou d'incohérences. En effet, l'opacité et l'absence de supervision humaine ont également des conséquences sur l'application des règles de protection des données à caractère personnel. Ces dernières méritent donc également d'être étudiées.

[232] *Proposition 2022/0303(COD) du 28 septembre 2022 de Directive du Parlement européen et du Conseil relative à l'adaptation des règles en matière de responsabilité civile extracontractuelle au domaine de l'intelligence artificielle, (Directive sur la responsabilité en matière d'IA)*, COM(2022) 496 final.
[233] *Proposition 2022/0302 (COD) du 28 septembre 2022 de Directive du Parlement européen et du Conseil relative à la responsabilité des produits défectueux*, COM(2022) 495 final.

Chapitre 2 : Une protection excessive des données à caractère personnel

102. **La protection de la vie privée à l'ère des technologies numériques.** L'histoire de la protection de la vie privée est profondément liée à l'histoire des technologies. La notion de « vie privée » a évolué en tant que concept normatif en parallèle du développement des nouvelles technologies de l'information et de la communication depuis les débuts de l'ère moderne[234]. Au début du XIXème siècle, une combinaison de changements sociétaux et d'avancées technologiques ont donné naissance à de nouvelles menaces pour la vie privée. Les technologies innovantes de l'époque, notamment les communications télégraphiques et les caméras portables, ont suscité des préoccupations croissantes en matière de protection de la vie privée[235], inspirant l'article « The right to privacy » de Samuel Warren en 1890[236], considéré comme l'un des pères fondateurs du droit à la vie privée outre-Atlantique[237]. Plus tard, dans la deuxième moitié du XXème siècle, le développement et l'utilisation massive de technologies numériques ont accéléré la collecte et l'utilisation de données personnelles. C'est ce progrès technique et les risques qu'il engendre qui ont conduit à la naissance des législations modernes sur la protection des données personnelles, au début des années 1970. En France, c'est en 1978 que le législateur a adopté la loi Informatique et Libertés[238], instaurant des gardes fous à l'utilisation des technologies numériques pour collecter et exploiter des données personnelles et offrant des droits spécifiques aux personnes concernées. Si elle a jusqu'à aujourd'hui parfaitement suffi pour répondre aux menaces que génèrent les technologies numériques pour la vie privée, elle fait face à un nouveau défi : le développement de l'IA et les risques éthiques, moraux ou sociétaux associés. L'Union européenne, quant à elle, a fait le choix d'un texte technologiquement neutre avec l'adoption du RGPD en 2016[239], se substituant aux précédentes

[234] I.R. Kramer, « The Birth of Privacy Law: A Century Since Warren and Brandeis », *Cath. U. Law Review*, 1990, 39, 703.
[235] U. Gasser, « Recoding Privacy Law: Reflections on the Future Relationship among Law, Technology, and Privacy », *Harvard Law Review* : *Law, privacy & technology commentary series*, 2016, 130, n°2, 10.
[236] S.D. Warren, L.D. Brandeis, « The Right to Privacy », *Harvard Law Review*, 1890, vol. 4, N°5, 193-220.
[237] I. R. Kramer, *op. cit.*
[238] *Loi n° 78-17 du 6 janvier 1978 relative à l'informatique, aux fichiers et aux libertés, dit « Loi Informatique et Libertés » ou « LIL »*, publiée au JORF du 7 janvier 1978.
[239] *Règlement* (UE) 2016/679 *du Parlement européen et du Conseil du 27 avril 2016, relatif à la protection des personnes physiques à l'égard du traitement des données à caractère personnel et à la libre circulation de ces données*, publié au JOUE n°L119/1 du 4 mai 2016.

directives sur le sujet[240]. La neutralité technologique se traduit concrètement par la création d'obligations générales, ne visant aucune technique en particulier. La Commission européenne, dans son évaluation du RGPD deux ans après son entrée en vigueur, souligne que cette approche, basée sur des principes, est conçue pour couvrir les nouvelles technologies au fur et à mesure de leur développement[241]. Il est donc considéré par beaucoup comme un outil essentiel et flexible pour garantir que ce progrès technique soit conforme aux droits fondamentaux. Mais quid des risques spécifiques à l'IA ?

103. **La protection des données personnelles et l'IA.** La conception et le fonctionnement des systèmes intelligents reposent, pour la majorité, sur le traitement de données. Ils nécessitent alors d'importants stocks de données, lesquelles peuvent être en tout ou partie à caractère personnel. Dès lors, les règles relatives à leur collecte et à leur traitement devront être respectées. Cela est d'autant plus légitime que l'IA peut générer des risques pour la vie privée des individus. En effet, souvent les stocks de données utilisés ont été collectés pour une finalité autre que l'apprentissage de la machine : quid du respect des droits des personnes concernées, notamment du droit d'opposition ? Aussi, comment garantir la confidentialité des données personnelles utilisées pour construire un modèle d'IA, lorsque l'on sait qu'il peut être victime d'attaques adverses visant à révéler les données d'apprentissage ? Enfin, est-il acceptable que nos émotions soient analysées, voire inférées, sans notre consentement ? La protection de la vie privée des individus, à travers les lois sur les données personnelles, est un garde-fou essentiel pour protéger nos droits face aux avancées technologiques, devenant alors une condition essentielle à la préservation de notre libre arbitre[242].

104. **Plan.** Le présent Chapitre a pour objectif de démontrer, dans un premier temps, que les principales règlementations en matière de protection des données à caractère personnel, à savoir la loi Informatique et Libertés en France et le RGPD au niveau européen, sont pleinement applicables aux systèmes d'IA (**Section 1**). En effet, ces régimes encadrent tous les traitements

---

[240] *Directive* 95/46/CE *du Parlement et du Conseil du 24 octobre 1995 relative à la protection des personnes physiques à l'égard du traitement des données à caractère personnel et à la libre circulation de ces données*, publiée au JOUE n°L281 du 23 novembre 1995.
[241] COMMISSION EUROPÉENNE, *Data protection as a pillar of citizens' empowerment and the EU's approach to the digital transition – two years of application of the GDPR*, document de travail accompagnant la communication de la Commission au Parlement et au Conseil, 24 juin 2020, COM(2020) 264 final.
[242] J.E. Cohen, « What privacy is for », *Harvard Law Review*, 2013, 126, n°7, 1904–1933.

de données à caractère personnel, sans égard à la technologie utilisée. Des dispositions spécifiques relatives au profilage ou à la prise de décision automatisée, souvent réalisés au moyen de techniques d'IA, viennent compléter cet encadrement. En revanche, l'application de ces règles et l'actualité jurisprudentielle de la Cour de Justice de l'Union européenne, en particulier sur le transfert de données à l'international, créent de nombreuses contraintes pour les entreprises souhaitant développer de tels systèmes. Tant la pratique en entreprise que la doctrine semblent confirmer que la protection des données génère des contraintes importantes limitant l'exploitation du plein potentiel de l'IA (**Section 2**). L'étude de ces contraintes révèle que, si la majorité des contraintes sont justifiées, certaines sont disproportionnées au vu des bénéfices attendus de la technologie et des moyens existants pour limiter le risque.

Section 1 : Un encadrement effectif des systèmes d'IA par le régime juridique de la protection des données

105. **L'application indifférenciée des principes de la protection des données aux traitements réalisés au moyen d'un système d'IA.** Dans son Livre blanc du 19 février 2020[243], la Commission européenne rappelle, au titre de ses options en matière de régulation de l'IA, que le cadre règlementaire préexistant est applicable à l'IA. À ce titre, elle vise plus spécifiquement la règlementation en matière de protection des données à caractère personnel contenue dans le RGPD, appliqué en France à travers la Loi Informatique et Libertés de 1978. Ces textes ont vocation à s'appliquer aux systèmes d'IA[244] dès lors qu'ils sont utilisés pour collecter ou traiter des données à caractère personnel, relatives à une personne physique identifiée ou identifiable[245]. En effet, avec le phénomène du Big data[246] et l'essor des techniques d'IA, « *les données sont collectées et traitées grâce à des supercalculateurs, capables d'obtenir des résultats que les moyens classiques de gestion de base de données ne permettent pas d'atteindre* »[247]. Toutefois, ces collectes et traitements n'échappent pas aux principes du RGPD et ne sauraient méconnaître les droits des personnes sur les données les concernant. Ainsi, tout traitement de données personnelles au moyen d'un système d'IA, qui semble rejoindre la notion de « traitement automatisé de données » dans les textes, doit respecter les principes contenus à l'article 5(1) du RGPD. Dès lors, devront être assurés les principes de licéité, loyauté et transparence, de limitation des finalités, de minimisation des données, d'exactitude, de limitation de la durée de conservation, ou encore de sécurité des données[248]. De la même façon,

---

[243] COMMISSION EUROPÉENNE, *Livre blanc du 19 février 2020 sur l'Intelligence Artificielle – Une approche européenne axée sur l'excellence et la confiance*, 19 février 2020, COM(2020) 65, p. 10.

[244] Sur le partage de ce constat par les institutions européennes, voir notamment : C. Crichton, « Publication par la Commission de son Livre blanc sur l'intelligence artificielle », *Dalloz Actualité*, 28 février 2020, disponible en ligne : <https://www.dalloz-actualite.fr/flash/publication-par-commission-de-son-livre-blanc-sur-l-intelligence-artificielle>, consulté le 11 juillet 2021 ; C. Crichton, « Intelligence artificielle : avis du CEPD sur le Livre blanc de la Commission », *Dalloz Actualité*, 17 juillet 2020, disponible en ligne : <https://www.dalloz-actualite.fr/flash/intelligence-artificielle-avis-du-cepd-sur-livre-blanc-de-commission>, consulté le 11 juillet 2021.

[245] Sur les champs d'application matériel et territorial du RGPD, voir N. Martial-Braz, « Le champ d'application du RGPD », in *Le règlement général sur la protection des données : Aspects institutionnels et matériels*, dir. A. Bensamoun, B. Bertrand, Mare & Martin, 2020, p. 20.

[246] J.S. Bergé, D. Le Métayer, « Phénomène de masse et droit des données », *Commerce Communication Electronique*, décembre 2018, Etudes, n°20.

[247] M. Quéméner, *Le droit face à la disruption numérique*, Gualino, Lextenson, 2018, p. 22.

[248] RGPD, article 5(1).

toute personne mettant en œuvre un tel traitement devra garantir l'effectivité des droits des personnes concernées, à savoir les droits à l'information, d'opposition, de recueil du consentement (le cas-échéant), d'accès et de rectification, ainsi qu'à la portabilité de ses données[249]. Voulues neutres technologiquement, ces obligations générales auront vocation à s'appliquer pleinement aux systèmes d'IA ainsi qu'à la collecte et aux traitements des données nécessaires à la conception. Leur respect pourra être rendu plus difficile par les spécificités de l'IA, en particulier en raison de l'opacité de certaines techniques. Ces difficultés ne remettent pas en cause l'applicabilité des règles, ce qui peut créer des contraintes au développement de certains traitements au moyen d'un système d'IA, lesquelles seront étudiées dans la Section 2. Avant cela, il convient de noter qu'au-delà des obligations générales contenues dans le RGPD et la LIL, pleinement applicable aux phases « amont » de construction d'une base de données d'entraînement et de conception du système d'IA, certaines dispositions viennent encadrer plus spécifiquement certains traitements algorithmiques, en pratique souvent réalisés au moyen de systèmes d'IA : le profilage et la prise de décision automatisée.

106. **Des dispositions encadrant plus spécifiquement les systèmes d'IA.** L'IA n'est ni définie ni directement visée par le RGPD. En revanche, plusieurs articles prévoient des mesures spécifiques au « profilage » (**§1**) et à la « prise de décision automatisée » (**§2**), deux types de traitement souvent réalisés au moyen de techniques d'IA.

§1 : Un encadrement souple pour le profilage réalisé au moyen d'un système d'IA

107. **Profilage et IA.** Le profilage est défini à l'article 4 §4 du RGPD, repris en son considérant 71, comme « *toute forme de traitement automatisé de données à caractère personnel visant à évaluer les aspects personnels relatifs à une personne physique, notamment pour analyser ou prédire des aspects concernant le rendement au travail de la personne concernée, sa situation économique, sa santé, ses préférences ou centres d'intérêt personnels, sa fiabilité ou son*

[249] RGPD, Chapitre III « Droits de la personne concernée », articles 12 – 23.

comportement, ou sa localisation et ses déplacements »[250]. Cette définition est très proche de celle utilisée par la Commission européenne pour définir l'IA : « *un logiciel qui est développé au moyen d'une ou plusieurs des techniques et approches énumérées à l'annexe I et qui peut […] générer des résultats tels que des contenus, des prédictions, des recommandations ou des décisions […]* »[251]. Si l'on croise ces deux définitions, on comprend que l'objectif du profilage (« analyser ou prédire des aspects personnels ») pourra être réalisé en pratique par l'IA (« générant des prédictions, des recommandations »). Les applications sont nombreuses et peuvent aller du « *suivi de personnes sur internet* », à la classification des profils, en passant par la prédiction des « *préférences, comportements, ou dispositions d'esprit* » des personnes concernées[252]. Dans le secteur de l'électricité, des systèmes de profilage basés sur de l'IA sont notamment utilisés dans la commercialisation et les services énergétiques. À titre d'exemple, l'IA peut être utilisée pour extraire des données relatives à un client à partir de ses échanges avec son fournisseur par mail ou en vocal avec son conseiller[253]. Les données collectées peuvent être traitées afin de déduire d'autres informations sur le client, telles que son humeur, la probabilité qu'il soit prêt à résilier son contrat et bien d'autres. L'ensemble de ces informations peuvent être renseignées sur son profil, lequel sera utilisé par le conseiller client afin de proposer la meilleure offre ou solution compte tenu des nombreux éléments de contexte[254]. D'autres exemples pourraient être pris dans le secteur de l'électricité, tels que l'utilisation du profil énergétique des clients, ou l'identification de nouveaux prospects sur la base de prédictions d'événements, tels qu'un déménagement ou un besoin de travaux en rénovation.

---

[250] Une définition embrassée par la doctrine, V. notamm. A. Debet, J. Massot, N. Metallinos, *La protection des données à caratère personnel en droit français et européen*, Lextenso Editions, 2015, n° 855 ; J.B. Duclercq, « Le droit public à l'ère des algorithmes », *Revue de Droit public*, 2017, n°5, 1401.

[251] COMMISSION EUROPÉENNE, *Proposition de règlement du Parlement européen et du Conseil établissant des règles harmonisées concernant l'intelligence artificielle (législation sur l'intelligence artificielle) et modifiant certains actes législatifs de l'Union*, 21 avril 2021, 2021/0106 (COD), article 3 (1).

[252] RGPD, Considérant 23.

[253] Voir sur le sujet deux exemples d'utilisation de l'IA à des fins de profilage dans la commercialisation de l'électricité : IBERDROLA, « Digital transformation and innovation plans », *Site official d'Iberdrola (blog)*, 2019, disponible en ligne : <https://www.iberdrola.com/about-us/utility-of-the-future/digital-transformation>, consulté le 12 juillet 2021 ; ENDESA, « Artificial intelligence to improve our services », *Site officiel d'Endesa (blog)*, 2019, disponible en ligne : <https://www.endesa.com/en/projects/a201904-artificial-intelligence-improve-services.html>, consulté le 12 juillet 2021.

[254] EDF, « Pour EDF l'IA doit être au service de l'humain », *Le Parisien (blog)*, 2020, disponible en ligne : <https://www.leparisien.fr/societe/pour-edf-l-ia-doit-etre-au-service-de-l-humain-26-06-2020-8337975.php>, consulté le 12 juillet 2021.

Fort heureusement, de telles pratiques sont rigoureusement encadrées par le RGPD, eu égard aux risques qu'elles engendrent pour les personnes concernées.

108. **Une pratique générant des risques pour les droits des personnes concernées.** Le profilage, en ce qu'il consiste généralement en la collecte de données sur une personne ou un groupe de personne et leur centralisation dans un système informatique unique en vue de leur analyse ultérieure, constitue finalement une sorte de « fichage » informatique. Les risques résident principalement dans les utilisations ultérieures des résultats du profilage. En effet, les profils ainsi réalisés peuvent être mobilisés pour déployer de vastes campagnes de communication[255] ou encore pour personnaliser des contenus tels que des sites internet[256] avec pour finalité d'influer sur le comportement des individus, atteignant au principe de l'autonomie individuelle[257]. Aussi, la performance des systèmes d'IA dans l'exploitation de ces « profils » est telle que les entreprises sont naturellement incitées à collecter un maximum de données[258], parfois au mépris des législations en vigueur. Dès lors, des entreprises mal intentionnées peuvent être amenées à collecter frauduleusement des données à des fins de profilage[259] ou à croiser des données afin de déduire certaines informations personnelles sans les obtenir directement auprès de la personne concernée, et souvent sans l'en informer. Ces risques pour la vie privée des individus[260] ont été pris au sérieux par les rédacteurs du RGPD puisque le texte

[255] J. Hinds, E.J. Williams, A.N. Joinson, « 'It wouldn't happen to me' : Privacy concerns and perspectives following the Cambridge Analytica scandal », *International Journal of Human-Computer Studies*, 2020, 143, 102498 ; I. Manokha, « Le scandale Cambridge Analytica contextualisé : le capital de plateforme, la surveillance et les données comme nouvelle « marchandise fictive » », *Cultures & Conflits*, 2018, 109, 39-59.

[256] G. Guebels, *Le phénomène des bulles de filtres sur Internet : Le moteur de recherche Google nous oriente-t-il à notre insu à cause de son algorithme de personnalisation ?*, Mémoire de recherche, dir. L. Groetaers, Faculté des sciences économiques, sociales, politiques et de communication, Université catholique de Louvain, 2018, disponible en ligne : <https://dial.uclouvain.be/memoire/ucl/object/thesis:16354>, consulté le 2 août 2021.

[257] K. Yeung, A. Howes, G. Pogrebna, « AI Governance by Human Rights-Centred Design, Deliberation and Oversight: An End to Ethics Washing », in *The Oxford Handbook of AI Ethics*, dir. M. Dubber, F. Pasquale, Oxford University Press, 2019 ; S. Sankaran, C. Zhang, M. Gutierrez Lopez, *et al.*, « Respecting Human Autonomy through Human-Centered AI », *in Proceedings of the 11th Nordic Conference on Human-Computer Interaction: Shaping Experiences, Shaping Society (NordiCHI '20)*, 2020, Association for Computing Machinery, New York, 134, 1–3.

[258] M. Kaptein, D. Eckles, « Selecting Effective Means to Any End: Futures and Ethics of Persuasion Profiling », *Persuasive Technology*, 2010, 82–93 ; CNIL, *Comment permettre à l'homme de garder la main ? Les enjeux éthiques des algorithmes et de l'intelligence artificielle*, Synthèse du débat public animé par la CNIL dans le cadre de la mission de réflexion éthique confiée par la loi pour une République Numérique, 15 décembre 2017.

[259] CNIL, *ibid*.

[260] Pour une étude plus détaillée des risques liés au traitement de données personnelles par l'IA, et notamment au profilage, V. A. Debet, « Intelligence artificielle et données à caractère personnel », in *Droit de l'intelligence artificielle*, dir. A. Bensamoun, G. Loiseau, LGDJ, coll. Les intégrales, 2019, p. 269.

souligne l'importance de leur encadrement, allant au-delà des obligations contraignantes déjà applicables à tous les traitements de données personnelles. À ce titre, l'information des personnes concernées occupe une place de choix et semble être une véritable priorité en matière de profilage puisqu'elle est mentionnée à six reprises dans les considérants[261].

109. **Un encadrement strict pour limiter les risques du profilage.** En plus de l'application des règles contraignantes régissant le traitement de données personnelles telles que la licéité du traitement ou les autres principes de protection des données[262], le profilage fait l'objet d'une attention particulière dans les textes[263]. En particulier, le RGPD précise dans ses considérants que « *le responsable de traitement devrait utiliser des procédures mathématiques ou statistiques adéquates aux fins du profilage* » et « *appliquer les mesures techniques et organisationnelles* » pour limiter au maximum le risque d'erreur dans les données utilisées, et pour prévenir tout effet discriminatoire, direct ou indirect[264]. Une telle rédaction pourrait laisser penser qu'il pèserait, sur les responsables de traitement, une obligation de moyens à mettre en place toute mesure réalisable pour limiter le risque d'effets discriminatoires liés au profilage. Pourtant, le corps du RGPD ne reprend pas cette idée et se contente de reconnaître le droit d'opposition au profilage fondé sur l'intérêt légitime du responsable de traitement ou la réalisation d'une mission d'intérêt public, et au profilage utilisé aux fins de prospection[265]. Pour le reste, le régime du profilage est adossé à celui de la prise de décision automatisée. La loi Informatique et Libertés, quant à elle, prohibe tout profilage entraînant une discrimination sur la base des critères mentionnés en son article 6, interdisant lui-même tous les traitements des données considérées comme sensibles[266]. Il convient tout de même de constater que le profilage conduisant à une prise de décision, automatisée ou non, figure au rang des traitements « à

[261] RGPD, Considérants 24 ; 60 ; 63 ; 70 ; 71 ; 72.
[262] RGPD, Considérant 72.
[263] Notons que le terme « profilage » n'apparaît que 3 fois dans le corps du texte de la loi Informatique et Libertés du 6 janvier 1978 contre 22 fois dans le RGPD, puisque le premier texte effectue de nombreux renvois vers le second. Le RGPD présente donc plus de matière pour analyser l'esprit de la régulation sur le sujet.
[264] RGPD, considérant 71.
[265] RGPD, article 21 (1) et (2).
[266] Loi Informatique et Libertés du 6 janvier 1978, articles 6 et 95.

risque » et devant systématiquement faire l'objet d'une analyse d'impact relative à la protection des données[267].

110. **Conclusion et transition.** L'IA peut être utilisée pour profiler des personnes à partir de nombreuses données les concernant. Dans le secteur de l'énergie, il peut s'agir des habitudes de consommation, du type de contrat souscrit ou encore de l'historique des réclamations. Ce cas d'usage n'est pas sans risque pour les droits des personnes et il convient de noter qu'il est déjà bien couvert par le corpus législatif. En effet, dès lors qu'un système d'IA sera utilisé pour évaluer des aspects personnels ou prédire des informations sur des personnes physiques, il devra respecter non seulement les règles et principes relatifs à la protection des données, mais aussi les obligations spécifiques au profilage. Ces dernières sont contraignantes, notamment du fait du risque d'effets discriminatoires à prévenir et de l'analyse d'impact à réaliser si des décisions sont prises sur son fondement. De plus, cet encadrement se trouvera encore plus contraignant si le profilage réalisé au moyen d'un système d'IA produit « *des effets juridiques* » pour les personnes concernées (accès à une offre de fourniture d'électricité plutôt qu'une autre, octroi d'une réduction de prix sur un contrat en cours…) ou « *qu'il l'affecte de manière significative* »[268]. En effet, dans ce cas précis, le profilage basculera dans le champ de la prise de décision automatisée, répondant d'un régime encore plus strict.

[267] RGPD, article 35.
[268] RGPD, considérant 71.

§2 : Des gardes fous pour garantir un contrôle humain sur les prises de décision automatisées

111. **Décision automatisée et IA.** Les résultats obtenus à partir du traitement de données par un système d'IA peuvent être utilisés par l'homme pour prendre des décisions. Ces décisions peuvent concerner tant des objets, par exemple dans le cas de la maintenance prédictive, que des personnes[269], notamment dans le cas du profilage. Dans le secteur de l'électricité, on pourrait imaginer des systèmes établissant un profil détaillé de consommation pour les particuliers ou professionnels qui conditionneraient l'accès à certaines offres de fourniture. L'utilisation de tels systèmes aurait pour conséquence de discriminer les clients lors de la conclusion du contrat, en excluant certains profils de l'accès à certaines offres (puisqu'elles ne leur seraient pas proposées). Cette application, si elle est dépourvue d'intervention humaine, entrerait dans la qualification de la décision automatisée au sens du RGPD et de la loi Informatique et Libertés. Le premier texte reconnaît en son article 22 le droit des personnes à « *ne pas faire l'objet d'une décision fondée exclusivement sur un traitement automatisé, y compris le profilage, produisant des effets juridiques la concernant ou l'affectant de manière significative de façon similaire* »[270]. La formulation amène une question fondamentale : s'agit-il d'un droit que les personnes doivent exercer ou la décision automatisée est-elle par défaut interdite ? Le second texte contient une nuance non négligeable en son article 95 puisque, plutôt que de reconnaître un droit à ne pas faire l'objet d'une décision automatisée, il interdit *stricto sensu* les décisions « *produisant des effets juridiques à l'égard d'une personne ou l'affectant de manière significative* » prises sur « *le seul fondement d'un traitement automatisé de données* »[271], levant ainsi tous les doutes issus de la formulation du RGPD. Dès lors, les systèmes d'IA pour la prise de décision, dont des exemples ont été cités précédemment, seraient prohibés en France s'ils produisaient des effets juridiques ou significatifs pour les personnes et étaient totalement automatisés. Dans l'Union européenne, le responsable du traitement en

[269] A titre d'exemple en dehors du secteur de l'électricité, des banques ou organismes d'assurance utilisent des systèmes de « scoring » (une forme de profilage) basés sur de l'IA pour conditionner l'accès des personnes à certains clients dont le profil est le moins risqué : V. notamm. J. Morel-Maroger, « La protection des données personnelles des clients des banques : Bilan et perspectives », *Revue de Droit bancaire et financier*, mars 2011, p. 7 et s.
[270] RGPD, article 22.
[271] Loi Informatique et Libertés du 6 janvier 1978, article 95.

question devra respecter les obligations renforcées d'information pour les prises de décision automatisée et garantir le droit des personnes à obtenir une intervention humaine.

112. **Plan.** Ces contraintes au déploiement de systèmes de prise de décision automatisée sont justifiées par les risques liés à l'automatisation, notamment en termes de discrimination et de reproduction de biais présents dans la société. Leur mise en œuvre requiert néanmoins des précisions au regard de ce qui constitue une « intervention humaine », excluant *de facto* le traitement du champ de l'interdiction contenue dans la LIL (**A**), et de ce que doit contenir l'information de la personne concernée dans le cas d'une décision algorithmique (**B**). En effet, sur ce dernier point, l'opacité de certains systèmes d'IA peut constituer une contrainte supplémentaire à la mise en œuvre des règles du RGPD.

A/ Un indispensable contrôle humain sur la prise de décision algorithmique

113. **La nécessité d'une intervention humaine et non artificielle.** Le groupe de l'article 29 a eu l'occasion de préciser les contours du régime des décisions automatisées dans ses lignes directrices du 3 octobre 2017[272]. Ainsi, le principe de l'interdiction figurant à l'article 22 du RGPD concerne les « *décisions prises par des moyens techniques sans aucune intervention humaine* »[273]. L'implication de l'humain permet donc au responsable de traitement de ne pas avoir à se plier au régime strict des décisions automatisées. C'est une des raisons pour laquelle peu d'utilisateurs de systèmes d'IA délèguent intégralement le processus de décision à la machine. En revanche, cette intervention humaine doit être significative : elle ne doit pas être « insignifiante ou artificielle »[274]. La CNIL souhaite véritablement qu'un « contrôle humain garantisse la maîtrise de l'algorithme »[275]. Pourtant, la question de savoir à partir de quel

[272] GROUPE DE TRAVAIL « ARTICLE 29 » SUR LA PROTECTION DES DONNÉES, *Lignes directrices relatives à la prise de décision individuelle automatisée et au profilage aux fins du règlement (UE) 2016/679*, 3 octobre 2017, n°WP251.
[273] *Ibid.*, p. 8.
[274] N. Martial-Braz, « Le profilage », *Communication Commerce Electronique*, avril 2018, p. 70 et s., dossier spécial « Entrée en vigueur du Règlement général sur la protection des données ».
[275] CNIL, *Délibération n°2017-299 du 30 novembre 2017 portant avis sur un projet de loi d'adaptation au droit de l'Union européenne de la loi n°78-17 du 6 janvier 1978*, p. 28.

moment un système d'aide à la décision bascule dans le champ de la décision automatisée se pose toujours fréquemment dans la pratique[276]. On notera sur ce point que le Conseil d'État s'est déjà prononcé sur la notion d'intervention humaine à l'occasion d'un arrêt datant de 1983[277]. Le juge y considère que la prise en compte par une personne humaine d'un ensemble d'informations (et non seulement du résultat d'un traitement informatique) ainsi que des explications fournies par la personne concernée permettrait de démontrer la réalité de l'intervention humaine. Bien que l'arrêt soit ancien, le raisonnement de la juridiction administrative pourrait être toujours d'actualité.

114. **Une solution à travers les systèmes d'aide à la décision.** Cet encadrement vise à éviter que des aspects essentiels de la vie des individus se voient régis automatiquement par des algorithmes, avec tout ce que cela implique en termes de risques de reproduction de biais présents dans la société, souvent au détriment des minorités. Dans le secteur de l'électricité, il n'existe à notre connaissance aucun système d'IA utilisé aux fins de prise de décision entièrement automatisée. La supervision humaine est pour le moment toujours garantie avec des systèmes « d'aide » à la prise de décision. Il convient tout de même de noter l'importance du terme « aide » puisque, pour ne pas rentrer dans le champ des décisions automatisées, il faut bien que le traitement algorithmique ne soit qu'un des éléments fondant la décision. L'analyse humaine doit être réelle et ne saurait consister en une validation systématique des résultats de l'algorithme[278].

115. **Transition.** De plus, le principe d'interdiction des décisions automatisées vient s'ajouter aux droits des personnes concernées, qui se voient renforcés dans ce cas précis.

---

[276] Pour une analyse des difficultés à saisir les éléments constitutifs de la décision automatisée dans le RGPD, V. notamm. A. Danis-Fatôme, « Décisions automatisées et profilage », *in Le règlement général sur la protection des données : Aspects institutionnels et matériels*, dir. A. Bensamoun, B. Bertrand, Mare & Martin, 2020, p. 193.

[277] CONSEIL D'ÉTAT, 29 juillet 1983, *Docteur Cloarec*, n°32172 ; H. Maisl note ss. CE, 29 juillet 1983, *Recueil Dalloz*, 1985, 49.

[278] R. Brauneis, E.P. Goodman, « Algorithmic Transparency for the Smart City », *Yale Journal of Law & Technology*, vol. 20, n°103, 2018, p. 126-127 : « *Avec le temps, s'en remettre aveuglément aux algorithmes pourrait affaiblir la capacité de prise de décision des agents de l'État de même que leur sens de l'engagement et de l'action* » (traduction libre).

B/ Une obligation d'information renforcée pour la prise de décision automatisée

116. **L'information de la personne concernée.** L'article 13 du RGPD prévoit une obligation d'information spéciale dans le cas des décisions automatisées[279]. Pour être complète, l'information de la personne concernée doit mentionner l'existence d'une prise de décision automatisée, y compris un profilage, les données traitées, la logique sous-jacente de l'algorithme utilisé, l'objectif du traitement et ses conséquences pour la personne[280]. En pratique, l'information sur la logique sous-jacente du traitement pose des difficultés, notamment quant à la détermination de son contenu[281] et la protection du secret des affaires, pouvant justifier la volonté de certaines entreprises de ne pas communiquer sur leurs algorithmes. Des enseignements peuvent être tirés de différentes affaires concernant des algorithmes utilisés par l'administration en France. En effet, la CNIL a pu préciser à l'occasion de l'affaire « Parcoursup » que le responsable de traitement mettant en œuvre une décision automatisée doit communiquer aux personnes concernées « *la méthode ayant permis de développer l'algorithme, le score obtenu par [la personne concernée], les seuils de scoring et leur signification* »[282]. Ces éléments donnent des indications, y compris pour le secteur privé, sur ce que doit contenir une information complète de la personne concernée lorsque ses données sont traitées à des fins de prise de décision automatisée.

117. **L'information de la personne concernée à travers le droit d'accès.** En complément du droit à l'information en amont de la prise de décision automatisée, les individus peuvent également exercer leur droit d'accès afin de recueillir de précieuses informations sur la façon dont sont traitées leurs données[283]. Ce droit leur permet de demander au responsable de traitement la confirmation que leurs données sont traitées, de connaître la finalité du traitement,

---

[279] RGPD, article 13 (2) f).
[280] A. Danis-Fatôme, *op. cit.*, p. 207 ; V. aussi J. Rochfeld, « L'encadrement des décisions prises par algorithme », *Dalloz IP/IT*, 2018, 474.
[281] Selon le Groupe de l'article 29, l'information doit être intelligible pour la personne concernée, ce qui exclut de fait la communication de principes mathématiques ou techniques.
[282] CNIL, *Délibération du 30 août 2017 prise à l'encontre du Ministère de l'Enseignement supérieur, de la Recherche et de l'Innovation, n°2017-053* ; sur le même sujet, voir aussi : QPC n°2020-834, 3 avril 2020, aff. « Parcoursup ».
[283] RGPD, article 15 (1) h) ; V. aussi : J. Rochfeld, *op. cit.*

d'en connaître les destinataires et de se faire communiquer les données en question[284]. Une partie de la doctrine considère que ce droit d'accès permettrait également d'obtenir des explications sur les motifs d'une décision prise sur le fondement d'un algorithme[285]. L'ensemble de ces obligations doivent être prises en compte par les responsables de traitement utilisant des systèmes d'IA pour prendre des décisions à propos d'individus. Il convient alors de mettre en place des mesures techniques et organisationnelles pour permettre à la fois de garantir la bonne information des personnes (mise en place de bannières d'informations, recueil du consentement explicite…) et l'effectivité du droit d'accès (identification de la personne à contacter, mise à disposition d'une boîte mail générique, traçabilité des prises de décision pour pouvoir générer des explications…). L'ensemble de ces mesures représentent un coût non négligeable, notamment lorsque les systèmes d'IA peuvent concerner des centaines de milliers d'individus comme c'est souvent le cas dans le secteur de l'énergie, un marché concentré autour de grands acteurs historiques. Ainsi, malgré leur légitimité, elles peuvent constituer des contraintes non négligeables à l'adoption des systèmes d'IA.

118. **Transition.** Ainsi, les règles et principes contenus dans le RGPD, tant généraux que spécifiques aux traitements automatisés, sont pleinement applicables aux systèmes d'IA. S'il n'y a pas de problème d'applicabilité, leur mise en œuvre peut en revanche être rendue bien plus difficile lorsque les traitements concernés sont réalisés au moyen de systèmes d'IA, du fait de leurs spécificités. Dès lors, il apparaît que l'application du régime actuel de protection des données à caractère personnel, découlant du RGPD et de la loi Informatique et Libertés, peut générer des contraintes non négligeables au développement de l'IA. Dans notre recherche continue de proportionnalité entre prévention des risques et promotion de l'innovation, il apparaît que certaines de ces contraintes sont disproportionnées au regard des bénéfices potentiels que l'IA peut apporter.

[284] A. Danis-Fatôme, *op. cit.*, p. 209.
[285] *Ibid*, p. 210.

Section 2 : Un encadrement freinant le développement des systèmes d'IA

119. **Plan.** La confrontation des règles existantes en matière de protection des données avec les spécificités de l'IA génère des contraintes parfois disproportionnées à son développement. En effet, les principes généraux de la protection des données contenus dans le RGPD semblent à de nombreux égards incompatibles dans leur esprit avec la nature même de l'IA (**§1**). Naissent alors des tensions entre les différents principes et les éléments constitutifs des systèmes d'IA tels que l'autonomie ou la dépendance à la donnée. Ces tensions peuvent créer des réticences chez les entreprises souhaitant développer des systèmes d'IA traitant de données personnelles. De plus, le défaut d'explicabilité de ces systèmes fait que les responsables de traitement seront bien souvent dans l'impossibilité matérielle de remplir les obligations de transparence et d'information contenues dans les textes (**§2**). Enfin, la complexité structurelle des systèmes d'IA nécessite souvent le recours au *Cloud Computing* et donc au transfert transfrontalier de données. Il ressort de l'analyse que le régime des transferts de données en dehors de l'Union européenne, notamment au vu de la jurisprudence récente de la Cour de Justice de l'Union européenne (CJUE), constitue un blocage majeur pour le recours à l'IA, notamment dans le secteur de l'électricité qui débute sa transition numérique (**§3**). En effet, les éditeurs des solutions d'IA les plus performantes et sécurisées ne se trouvent pas toujours dans un État jugé adéquat au sens du RGPD.

§1 : Des tensions entre les principes du RGPD et les spécificités des systèmes d'IA

120. **Une neutralité technologique de façade.** L'esprit du RGPD et les objectifs qu'il poursuit se retrouvent principalement dans les principes relatifs au traitement de données à caractère personnel évoqués en son article 5 : licéité, loyauté et transparence ; limitation des finalités ; minimisation des données ; exactitude ; limitation de la conservation ; intégrité et confidentialité ; responsabilité. Bien que le RGPD soit présenté comme neutre technologiquement, l'articulation entre ces différents principes peut s'avérer très difficile lorsqu'ils sont confrontés à des technologies de rupture telles que la blockchain, l'internet des

objets ou l'IA[286]. Ces difficultés peuvent avoir pour conséquence d'imposer aux entreprises de renoncer à certains cas d'usage[287]. L'applicabilité du contenu du RGPD telle que démontrée précédemment n'est pas remise en cause. Il s'agit ici de démontrer que son application conduit à des incohérences et amène les responsables de traitement à devoir réaliser des arbitrages entre les différents principes qu'il contient, voire à renoncer au traitement uniquement parce qu'il a recours à une technologie particulière. Les paragraphes suivants présentent plusieurs exemples de tensions entre les principes du RGPD et les spécificités des systèmes d'IA[288].

121. **La limitation des finalités crée un frein à l'apprentissage sur des données.** Le fonctionnement des systèmes d'IA fondés sur des techniques d'apprentissage automatique repose sur l'utilisation de grandes quantités de données. Elles peuvent être issues de différentes sources et avoir été collectées à différentes occasions. Le processus de conception d'un système informatique de résolution de problème consiste en l'identification du problème, des données disponibles et enfin au développement d'un algorithme de résolution du problème. Pour être conforme au RGPD, l'utilisation des données aux fins d'apprentissage du système d'IA et de son fonctionnement doit être compatible avec la finalité initiale pour laquelle elles ont été collectées[289]. On se rend compte de la complexité du processus puisque la conception du système peut intervenir bien après cette collecte. En pratique, les développeurs se retrouvent donc dans l'impossibilité d'utiliser les données qu'ils ont à leur disposition mais qui ont été collectées pour d'autres finalités considérées incompatibles. Cette situation est justifiée lorsque les données révèlent de nombreux éléments sur la vie privée des individus. En revanche, elle

[286] B. Bertrand, « Perfectibilité de la protection des données personnelles », *Revue Trimestrielle de Droit européen*, 2021, Chronique Droit européen du numérique, 143 ; V. aussi sur les difficultés similaires se présentant lors de l'application des principes du RGPD à des technologies de blockchain : O. Lasmoles, « Difficulties faced by the legal system in coming to terms with blockchains », *Revue internationale de droit économique*, 2018, vol. 4, 453-469.

[287] K. Brousmiche, G. Deleuze, P. Tellier, « Les enjeux de protection des données dans les usages des Blockchains pour la transition énergétique », in *Vie privée et numérisation. Des enjeux pour le monde de l'énergie*, Lavoisier, 2023, à paraître.

[288] Certaines de ces tensions ont déjà pu être mises en évidence dans différents rapports, notamm. THE NORVEGIAN DATA PROTECTION AUTHORITY, *Artificial intelligence and privacy,* Rapport, Janvier 2018, pp. 15-19, disponible en ligne : <https://www.datatilsynet.no/globalassets/global/english/ai-and-privacy.pdf> ; CNIL, *Comment permettre à l'homme de garder la main ? Les enjeux éthiques des algorithmes et de l'intelligence artificielle*, Synthèse du débat public animé par la CNIL dans le cadre de la mission de réflexion éthique confiée par la loi pour une République Numérique, 15 décembre 2017.

[289] V. notamm. A. Bensamoun, G. Loiseau, « L'intelligence artificielle : faut-il légiférer ? », *Recueil Dalloz*, 2017, p. 581.

pourrait être assouplie dans le cas de données pseudonymisées, ou si la conception du système respecte suffisamment le principe du *privacy by design*, par exemple en anonymisant les données avant leur utilisation[290] ou en utilisant des jeux de données fictives.

122. **La minimisation nuit à la précision du modèle d'IA.** Les systèmes basés sur des techniques d'apprentissage automatique nécessitent de grandes quantités de données lors de leur phase d'entraînement. Généralement, la performance du modèle pour simuler un phénomène donné (simulation de consommation électrique…) ou réaliser une tâche (reconnaissance et classification d'images…) sera proportionnelle à la quantité de données mobilisables pour l'apprentissage. Le principe de minimisation contenu dans le RGPD pourrait donc constituer un obstacle à la performance de tels systèmes, en ce qu'il contraint les responsables de traitement à ne traiter que les données strictement nécessaires au regard de la finalité poursuivie[291]. En pratique, ce critère de la stricte nécessité conduit les développeurs à ne pas explorer et expérimenter des applications de l'IA qui nécessiteraient de grandes quantités de données. Une nuance doit toutefois être apportée puisque la performance du modèle d'IA est également fonction de la qualité des données utilisées. Ces données, en plus d'être en quantité suffisante, doivent en effet être exemptes de biais et pertinentes au regard de la tâche à réaliser[292]. De plus, il existe un risque de surentraînement si le modèle apprend sur une quantité trop importante de données. Il convient donc de rechercher l'équilibre entre le recours à une quantité de données minimisée, en application du principe du RGPD, et une masse trop importante, aboutissant dans les deux cas à des performances sous-optimales. En pratique, pour ne pas bloquer le développement de systèmes d'IA traitant des données personnelles, la justification du respect du principe de minimisation se réalise *a posteriori* : les données sont sélectionnées, le modèle est expérimenté sur différentes quantités de données, puis l'on cherchera à justifier que la quantité de données a été minimisée au regard de la finalité

[290] En l'état actuel de la législation, le processus d'anonymisation constitue en lui-même un traitement de données à caractère personnel, tout comme le transfert des données en vue de leur anonymisation.
[291] RGPD, article 5 (1) c) ; C. Féral-Schuhl, *Cyberdroit*, Dalloz, coll. Praxis, 8ème éd., 2020, 1852 p., spec. 113.151 ; pour un exemple de l'application du principe de minimisation par la CNIL, voir CNIL, *Décision n°2016-058 du 30 juin 2016 mettant en demeure la société Microsoft Corporation*.
[292] A. Jain, H. Patel, L. Nagalapatti, *et al.*, « Overview and Importance of Data Quality for Machine Learning Tasks », *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, 2020, 3561–3562.

poursuivie. L'effet n'est donc pas celui escompté puisque le RGPD a pour vocation d'instaurer une logique de *privacy by design*, ou prise en compte de ses principes dès la conception.

123. **La confidentialité conduit à altérer les données utilisées et la performance du modèle.** L'utilisation de techniques permettant de garantir la confidentialité des données affectent la précision du résultat. Tout d'abord, l'anonymisation de données à caractère personnel présente de nombreuses limites[293]. En effet, l'état actuel de la technique ne permet que difficilement une désidentification totale des données et les modèles d'IA peuvent toujours être victimes d'attaques adverses visant à la réidentification des données utilisées pour la phase d'apprentissage. De plus, des données aujourd'hui anonymisées pourraient devenir identifiantes dans le futur en raison de l'apparition de nouvelles techniques de réidentification. L'anonymisation des données peut également conduire à des effets pervers pour l'innovation puisque retirer le rattachement de certaines données à des personnes physiques rend plus difficile le développement d'un certain nombre d'applications potentielles, telles que des systèmes d'optimisation des consommations électriques à large échelle[294]. Enfin, l'utilisation de procédés cryptographiques pour garantir la sécurité des données, souvent considérées comme de la simple pseudonymisation[295], peuvent conduire à altérer la performance globale du système d'IA. À titre d'exemple, la confidentialité différentielle consiste à introduire de fausses données non identifiantes dans un jeu de données afin d'empêcher toute personne mal intentionnée de pouvoir déterminer avec certitude si les données auxquelles il a réussi à accéder sont réelles ou non[296]. Par nature donc, cette technique atteint à la véracité, l'exactitude et la

[293] GROUPE DE TRAVAIL « ARTICLE 29 » SUR LA PROTECTION DES DONNÉES, *Avis 05/2014 sur les Techniques d'anonymisation*, 0829/14/FR, WP216, spec. p. 11 et s.

[294] A titre d'exemple, il est impossible de concevoir un système d'optimisation des courbes de consommation électrique à l'échelle d'un quartier sans collecter et traiter des données énergétiques attachées à chaque foyer. Ces données, puisqu'elles sont liées à un « Point De Livraison » fixe (ou « PDL »), seront indirectement identifiantes. La mise en conformité avec le RGPD sera alors extrêmement contraignante et rendra impossible la mise en place du système qui, pourtant, pourrait garantir la sécurité des données par des procédés cryptographiques et une supervision par l'opérateur du réseau électrique en tant que tiers de confiance responsable (car soumis aux obligations du service public de la distribution d'énergie).

[295] GROUPE DE TRAVAIL « ARTICLE29 » SUR LA PROTECTION DES DONNÉES, *Avis 05/2014 sur les Techniques d'anonymisation*, *op. cit.*

[296] *Ibid.*, p. 16 ; C. Dwork, « Differential privacy », in *Automata, languages and programming*, Springer Berlin Heidelberg, 2006, pp. 1-12 ; V. aussi, spec. dans le secteur de l'électricité : F. Leukam Lako, *Protection des données à caractère personnel pour les services énergétiques*, thèse pour le doctorat en informatique, réseaux et télécommunications, Institut polytechnique de Paris, 2021, spec. p. 40 et s.

pertinence des données sur lesquelles le modèle d'IA va apprendre ou inférer. L'objectif des développeurs consistera à introduire suffisamment de données fictives pour empêcher la réidentification, tout en préservant au maximum la précision du modèle. Un autre exemple peut être donné avec les techniques de cryptographie homomorphe, consistant à faire opérer les algorithmes sur les données chiffrées directement sans avoir besoin de les déchiffrer avant leur traitement[297]. Toutefois, cette technique complexifie grandement le traitement des données, qui doit prendre en compte le fait qu'elles soient chiffrées. Elle peut conduire également à allonger les temps de calcul[298], ce qui n'est pas sans effet en termes de coût financier et d'empreinte environnementale. La protection des données à caractère personnel est une nécessité absolue et justifie en grande partie les contraintes évoquées ci-dessus. Cela dit, le respect de ces obligations ne devrait pas avoir pour effet de limiter la performance des systèmes d'IA. Des souplesses pourraient être accordées aux responsables de traitement qui acceptent d'employer ces techniques, hautement protectrices mais très contraignantes puisqu'elles ne dispensent que rarement des autres obligations du RGPD, l'anonymisation étant elle-même un traitement de données personnelles…

124. **Le principe de licéité du traitement est un obstacle à la prévention de l'apparition de biais discriminatoires.** Le RGPD et la Loi Informatique et Libertés proscrivent le traitement de données considérées comme sensibles[299], tout en encourageant les responsables de traitement à lutter contre les effets discriminatoires des algorithmes[300]. Or, la communauté scientifique et la doctrine juridique s'accordent pour dire que la lutte contre les biais discriminatoires dans les données nécessite le traitement de données jugées comme sensibles[301]. En effet, les audits algorithmiques requièrent de disposer de bases de tests, biaisées et non biaisées, pour mettre en

[297] *Ibid.*, pp. 31-33.
[298] *Ibid.*, p. 43 : « *Cependant […] effectuer des opérations complexes (multiplications et additions) sur des chiffrés prend encore beaucoup de temps.* ».
[299] RGPD, article 9 et Loi Informatique et Libertés du 6 janvier 1978, article 6.
[300] RGPD, considérant 71 ; CONSEIL DE L'EUROPE, *Lignes directrices sur l'intelligence artificielle et la protection des données*, Rapport, novembre 2019, p. 9, disponible en ligne : <https://rm.coe.int/2018-lignes-directrices-sur-l-intelligence-artificielle-et-la-protecti/168098e1b8>, consulté le 3 août 2021.
[301] M. Gérot, W. Maxwell, « Le RGPD pourrait freiner les ambitions de l'Europe en matière d'intelligence artificielle », *ecommercemag (blog)*, 30 mars 2020, disponible en ligne : < https://www.ecommercemag.fr/Thematique/management-1225/Breves/Tribune-RGPD-pourrait-freiner-ambitions-Europe-matiere-intelligence-artificielle-348359.htm> consulté le 15 avril 2020 ; INSTITUT MONTAIGNE, *Algorithmes : contrôle des biais S.V.P.*, Rapport, mars 2020, p. 63 et s. ; J. Photopoulos, « Fighting algorithm bias », *Phys. World*, 2021, vol. 34, n°5, 42.

avant de potentiels effets discriminatoires[302]. À l'heure actuelle, les bases utilisées doivent être fictives, et ne seront donc pas parfaitement représentatives de la réalité. De la même manière, la conception de systèmes d'IA non discriminants peut nécessiter le traitement de données sensibles. À titre d'exemple, un système de reconnaissance vocale devrait, pour ne discriminer aucune minorité, être en mesure de comprendre sans distinction tous les accents d'une même langue. Pour remplir cet objectif, il faudrait que sa base d'apprentissage soit suffisamment diverse pour représenter tous les accents existants, ce qui ne correspond pas à la réalité pratique. La construction de bases de données représentatives nécessiterait une classification qui pourrait être assimilée à un traitement révélant l'origine ethnique des individus[303], interdit en l'état actuel de la législation.

125. **La transparence ne peut être que partielle avec des systèmes fondés sur de l'apprentissage profond (*Deep learning*).** L'opacité de l'IA peut avoir plusieurs sources mais, quelle qu'elle soit, cette spécificité rend très difficile la transparence prônée par le RGPD. Le responsable de traitement peut se retrouver dans l'impossibilité de fournir une information complète[304] à la personne concernée ou de donner une explication de la logique sous-jacente de l'algorithme dans le cadre de l'exercice du droit d'accès[305]. *A contrario*, une trop grande transparence sur la façon dont a été conçu le modèle d'IA peut également avoir des effets négatifs. En effet, elle pourrait conduire à révéler des secrets d'affaires[306], permettre

[302] S. Brown, J. Davidovic, A. Hasan, « The algorithm audit: Scoring the algorithms that score us », *Big Data & Society*, 2021, 1-8 ; S. Barocas, S. Hood, M. Ziewitz, « Governing Algorithms: A Provocation Piece », *Proceedings of the « Governing Algorithms » conference*, 16 mai 2013, New York University ; ADA LOVELACE INSTITUTE, *Examining the black box : Tools for assessing algorithmic systems*, Rapport, 2020, disponible en ligne : <https://www.adalovelaceinstitute.org/wp-content/uploads/2020/04/Ada-Lovelace-Institute-DataKind-UK-Examining-the-Black-Box-Report-2020.pdf>, consulté le 3 août 2021.

[303] Un projet de Chatbot observé dans une grande entreprise de télécommunications au cours de la thèse a mis en évidence le fait que le système de reconnaissance vocale utilisé ne reconnaissait pas les accents « africains ». Le seul palliatif possible était alors de réentraîner le modèle sur une nouvelle base d'apprentissage, enrichie de la contribution d'individus ayant un accent très prononcé. Cela revenait à reconstruire la base d'apprentissage (diffusée en open-source) entièrement avec des individus identifiés expressément comme ayant un fort accent « africain ». Le projet a finalement été abandonné face à l'ampleur de la tâche.

[304] GROUPE DE TRAVAIL « ARTICLE 29 » SUR LA PROTECTION DES DONNÉES, *Lignes directrices relatives à la prise de décision individuelle automatisée et au profilage aux fins du règlement (UE) 2016/679*, 3 octobre 2017, n°WP251.

[305] A. Danis-Fatôme, *op. cit.*, p. 209.

[306] La valeur de l'algorithme de Google vient en grande partie du secret donc il fait l'objet : P. Nirwan, « Le secret d'affaires : le droit de propriété intellectuelle caché sous le boisseau », *OMPI Magazine (blog)*, décembre 2017, disponible en ligne : <https://www.wipo.int/wipo_magazine/fr/2017/06/article_0006.html>, consulté le 3 août 2021 ; F. Pasquale, « Restoring Transparency to Automated Authority », *Journal on Telecommunications and*

l'exploitation de certaines failles par les utilisateurs[307], ou divulguer des informations sur les personnes dont les données ont été traitées. La transparence, lorsqu'elle est appliquée à des systèmes d'IA, présente donc des risques nouveaux et doit en conséquence être raisonnée.

Si le défaut d'explicabilité de ces systèmes peut constituer un frein majeur à son développement, de nombreux travaux de recherche visent à développer de nouvelles approches pour adapter le principe de la transparence à l'IA.

§2 : Le frein du défaut d'explicabilité des systèmes d'IA

126. **Des conséquences du défaut d'explicabilité de l'IA.** Comme cela a été démontré précédemment, les responsables de traitement ne sont pas toujours en capacité d'expliquer les justifications derrière les résultats d'un modèle d'IA, en particulier lorsque celui-ci repose sur des techniques d'apprentissage profond. Au mieux, ils auront la possibilité de communiquer sur les choix réalisés lors de la conception du modèle, les données utilisées pour son apprentissage le cas-échéant, les paramètres mathématiques que le modèle prend en compte et éventuellement le poids attribué à chaque paramètre[308]. Ces informations, éminemment techniques, sont difficilement compréhensibles pour les personnes concernées par les traitements. L'explicabilité doit toutefois être différenciée de l'interprétabilité. La première consiste en la « *capacité à donner à l'ensemble des utilisateurs, quel que soit leur bagage éducatif, une vision claire des procédures employées et des fonctionnalités remplies par l'algorithme, afin de permettre un usage informé* »[309]. La seconde correspond à la « *capacité des concepteurs de comprendre [le] fonctionnement [du modèle] et de vérifier s'il satisfait bien*

High Technology Law, 2011, vol. 9, n° 235, pp. 235-256 : « *Keeping [Google's]search algorithm private is the key to defeating gamers who might propagate link farms or other disfavoured methods to gain salience in search results* » ; M. Maggiolino, « EU Trade Secrets Law and Algorithmic Transparency », *Bocconi Legal Studies Research Paper*, 31 mars 2019, n° 3363178, disponible en ligne : <https://ssrn.com/abstract=3363178>, consulté le 7 août 2021.

[307] Des utilisateurs qui auraient connaissance de la façon dont a été conçu l'algorithme et des paramètres qu'il prend en compte pourraient en détourner le fonctionnement à leur avantage.

[308] Il est possible de déterminer le poids des paramètres traités par le modèle par tests successifs, afin d'identifier lesquels ont le plus d'influence sur le résultat final.

[309] M. Pégny, I. Ibnouhsein, « Quelle transparence pour les algorithmes d'apprentissage machine ? », *Revue d'intelligence artificielle,* 28 août 2018, 32, n°4, 447, 78.

les propriétés désirées »[310]). Nous nous intéresserons ici presque exclusivement à l'explicabilité, puisque la réponse à l'enjeu d'interprétabilité pour les concepteurs est principalement technique et donc hors champ de l'analyse juridique. Comme cela a été démontré précédemment, le défaut d'explicabilité peut être un obstacle aux droits à l'information préalable et d'accès pour les personnes concernées par un traitement de données personnelles basé sur de l'IA. En effet, dans ce cas, le responsable de traitement sera souvent dans l'incapacité de donner une explication complète et intelligible du fonctionnement de l'algorithme ou de sa logique sous-jacente. Cette incapacité peut être liée aux propriétés opaques du modèle utilisé, à une incompétence technique du responsable de traitement qui n'est pas forcément l'éditeur de la solution d'IA, ou encore à la difficulté d'interprétation du RGPD[311], rendant difficile de savoir précisément quel degré d'explication est attendu du responsable de traitement.

127. **Un obstacle au contrôle externe par les autorités de régulation.** Le défaut d'explicabilité des systèmes d'IA justifie la réticence des autorités de régulation à leur égard. En effet, l'opacité complique grandement les contrôles et éventuels processus de certification externes. Le Comité européen à la protection des données (CEPD), dans son avis du 29 juin 2020 sur le livre blanc de la Commission européenne sur l'IA soulève à cet égard « *le manque de moyens des autorités de contrôle et l'impossible explicabilité pour certaines IA* » et rappelle que c'est au responsable de traitement d'apporter la preuve de sa conformité avec le RGPD[312]. Cette preuve peut passer par la réalisation d'audits. Cette interprétation du RGPD rend la mise en conformité extrêmement lourde pour les responsables de traitement qui n'ont pas forcément les moyens ou la compétence technique pour la réalisation de tels audits, auquel cas ils n'auront pas d'autres choix que de renoncer au traitement par l'IA, même si ce dernier promettait de nombreux bénéfices.

---

[310] *Ibid.*

[311] Sur les difficultés d'interprétation naissant de la généralité des termes du RGPD, en particulier lorsque des traitements automatisés ou d'IA sont concernés, voir S. Wachter, B. Mittelstadt, L. Floridi, « Why a right to explanation of automated decision-making does not exist in the General Data Protection Regulation », *International Data Privacy Law*, 2017, 7(2), 76–99.

[312] CEPD, *Avis 4/2020 du 29 juin 2020 sur le livre blanc de la Commission sur l'intelligence artificielle*, spec. pts 46-52.

128. **Explicabilité et confiance dans les algorithmes d'IA.** Il est admis que la confiance dans les algorithmes, et en particulier dans les nouvelles techniques d'IA, est une condition à leur adoption. Cette confiance doit être créée dès la phase de conception des modèles, par une transparence sur les choix réalisés, ainsi que sur les capacités et limites des systèmes conçus. L'opacité de certaines techniques d'IA telles que l'apprentissage profond nécessite de trouver d'autres moyens que l'explication du fonctionnement du système pour créer la confiance. En effet, pour ne pas se priver de leurs incroyables performances, il convient de trouver de nouvelles formes d'explicabilité adaptées à ces techniques.

129. **Vers de nouvelles formes d'explicabilité pour l'IA.** Jusqu'à présent, plusieurs méthodes permettaient d'analyser le fonctionnement d'un logiciel informatique afin de le comprendre ou de l'expliquer[313]. Aujourd'hui, de nombreuses réflexions sont en cours afin de déterminer si ces méthodes pourraient être adaptées pour permettre la compréhension des systèmes d'IA dans une logique « d'interprétabilité »[314]. D'autres réflexions ont également été lancées sur la question de « l'explicabilité » de l'IA. À cette fin, il est nécessaire de déterminer les situations dans lesquelles une explication du fonctionnement de l'IA devrait être donnée. Dans un rapport de 2020, l'organe de supervision français de la banque et de l'assurance, l'ACPR, distingue par exemple l'explication à donner suivant l'audience destinataire[315].

130. **Premier niveau d'explication : l'observation.** Le premier niveau, appelé « observation », est destiné à un public non initié comme des consommateurs finaux. Il ne consiste pas en une explication exhaustive du fonctionnement de l'algorithme comme dans le cas de l'interprétabilité mais vise à présenter ce que fait l'algorithme et son utilité. Ainsi, il

[313] P. J. Roache, *Verification and validation in Computational Science and Engineering*, Hermosa publishers, 1998, 446 p., spec. p. 3-14.
[314] Pour un exemple, voir : EDF, « AI for Humanity : EDF, Thales et Total ouvrent le premier laboratoire industriel commun en Intelligence Artificielle », *Communiqué de presse*, 6 février 2020, disponible en ligne : <https://www.edf.fr/groupe-edf/espaces-dedies/journalistes/tous-les-communiques-de-presse/ai-for-humanity-edf-thales-et-total-ouvrent-le-premier-laboratoire-industriel-commun-en-intelligence-artificielle>, consulté le 7 août 2021.
[315] ACPR, *Gouvernance des algorithmes d'intelligence artificielle dans le secteur financier*, Document de réflexion, spec. p. 12 et s., disponible en ligne : <https://acpr.banque-france.fr/sites/default/files/medias/documents/20200612_gouvernance_evaluation_ia.pdf>, consulté le 24 mai 2021.

correspond plutôt au niveau d'information qu'il serait nécessaire de donner à des consommateurs pour créer la confiance dans l'IA utilisée.

131. **Deuxième niveau d'explication : la justification.** Le deuxième niveau, la « justification », consiste à expliquer les raisons pour lesquelles l'algorithme a donné un tel résultat. Il peut être accessible à un public non-expert, ou être utile à des fins de contrôle interne des algorithmes. On peut également trouver ici une référence au contenu du droit d'accès du RGPD qui inclut, dans le cas d'une décision automatisée, le droit d'obtenir une explication du résultat obtenu.

132. **Troisième niveau d'explication : l'approximation.** Le troisième niveau correspond à « l'approximation » et vise à présenter le fonctionnement du modèle d'IA. Sa compréhension peut nécessiter un bagage technique. Il est donc destiné à un public initié.

133. **Quatrième niveau d'explication : la réplication.** Le dernier niveau d'explication proposée par l'ACPR est la « réplication » et vise à prouver que l'algorithme fonctionne correctement. Il répond à un besoin d'analyse détaillée du modèle d'IA et des données nécessaires à l'explication. La réponse à la question du « bon » fonctionnement d'un algorithme étant éminemment technique, ce niveau d'explication est destiné à une audience experte.

134. **La nécessaire contextualisation de l'explicabilité de l'IA.** Il nous semble particulièrement pertinent de différencier le niveau d'explication d'un système d'IA suivant le risque que son opération génère ainsi que le public visé[316]. La forme de l'explication devrait également être suffisamment souple pour s'adapter à toutes les techniques que l'on peut faire entrer dans la vaste catégorie de l'IA[317]. Il apparaît opportun d'utiliser ces réflexions pour pallier les lacunes du droit existant lorsqu'il est confronté à l'opacité de l'IA. Penser de nouvelles formes d'explicabilité adaptées à l'IA permettrait une plus grande effectivité des droits à l'information et d'accès pour les individus, ainsi qu'un meilleur contrôle par les

[316] V. Beaudouin, I. Bloch, D. Bounie, *et al.*, « Flexible and Context-Specific AI Explainability: A Multidisciplinary Approach », *SSRN Electronic Journal*, 2020, disponible en ligne : < https://www.ssrn.com/abstract=3559477>, consulté le 29 avril 2020.

[317] Voir sur ce point les travaux de l'autorité de protection des données du Royaum-Uni : ICO et ALAN TURING INSTITUTE, *Explaining decisions made with AI*, Rapport d'étude dans le cadre du projet Explain, 20 mai 2020, 1.0.41, spec. p. 45 et s., disponible en ligne : <https://ico.org.uk/media/for-organisations/guide-to-data-protection/key-data-protection-themes/explaining-decisions-made-with-artificial-intelligence-1-0.pdf>.

autorités de régulation. Prenant acte de l'opacité des systèmes fondés sur des techniques d'IA opaques, il semble inadéquat de chercher à expliquer le fonctionnement de l'algorithme à un public non initié et pour lequel l'explication donnée ne serait pas forcément intelligible. Pour ce dernier, pouvant être composé de personnes concernées par une prise de décision automatisée au sens du RGPD, il apparaît plus pertinent de proposer une information visant à permettre aux individus d'agir sur la décision plutôt que de comprendre précisément le fonctionnement de l'algorithme. Cette information peut consister à expliquer les raisons justifiant le résultat (correspondant à la « justification » dans la classification proposée par l'ACPR), à donner à la personne concernée les moyens et fondements pour contester la décision prise, ou encore à proposer des pistes pour atteindre un résultat plus favorable[318]. En revanche, lorsque le destinataire visé par l'explication est un public expert (autorité de régulation, contrôle de conformité ou autres), le contenu de l'explication devrait être beaucoup plus technique.

135. **Synthèse sur la contrainte juridique liée au défaut d'explicabilité des systèmes d'IA.** Le défaut d'explicabilité de l'IA peut être un frein majeur à son adoption et à son acceptabilité. En effet, il risque de rendre inopérants les droits des individus, notamment au regard de leurs données personnelles et des décisions algorithmiques prises à leur égard. Il rend également plus difficile le contrôle des systèmes d'IA, tant par les entreprises qui les développent que par les autorités de régulation. Toutefois, il est possible de penser de nouvelles formes d'explication adaptées à l'IA, différenciées suivant les situations et les publics visés. Ces explications devraient se concentrer sur les informations nécessaires à l'exercice des droits des individus plutôt que sur l'explication précise du fonctionnement des algorithmes ou sa « logique sous-jacente ». Les travaux de recherche en cours sur le sujet sont prometteurs et pourraient donner lieu à des standards intersectoriels, sous la forme de lignes directrices ou autres supports de droit souple.

[318] S. Wachter, B. Mittelstadt, C. Russell, « Counterfactual explanations without opening the black box: Automated decisions and the GDPR », *Harvard Journal of Law & Technology*, 2017, vol. 31, n°2, 841–888.

§3 : Un impossible recours à certaines solutions d'IA développées par des entreprises établies en dehors de l'Union européenne

136. **Plan.** Le régime juridique applicable aux transferts de données en dehors de l'Union européenne génère des difficultés majeures pour les entreprises européennes souhaitant avoir recours à des solutions d'IA étrangers (**A**). Toutefois, des solutions peuvent être apportées à cette problématique (**B**).

A/ Des contraintes liées au régime des transferts de données en dehors de l'Union européenne

137. **Les transferts de données à caractère personnel en dehors de l'Union européenne.** Aux termes du RGPD, lorsque des données personnelles sont transférées vers des territoires dont la réglementation locale en matière de données personnelles n'est pas considérée comme équivalente au RGPD par la Commission Européenne, c'est-à-dire vers des pays « non adéquats », des mesures de protection dites « garanties appropriées » doivent être mises en place (ex : clauses contractuelles types, règles contraignantes d'entreprises…) par les responsables de traitement[319]. En particulier, dans le cadre de ces « garanties appropriées » l'importateur des données doit s'engager à traiter les données reçues dans le respect des principes contenus dans le RGPD. Le Comité européen à la protection des données (CEPD) a eu l'occasion de rappeler dans deux recommandations du 10 novembre 2020[320] à la suite de l'arrêt Schrems II de la CJUE[321] que tout transfert de données hors UE (hors pays considérés comme « adéquats ») doit faire l'objet d'une analyse préalable du droit local de l'importateur pour déterminer si l'une des « garanties appropriées » habituelles peut être valablement mise en place ou si le droit local (ou la façon dont le droit local est appliqué) priverait ces garanties d'effet. Lorsqu'au regard de cette analyse les « garanties appropriées » ne permettent pas de garantir le niveau de protection

[319] RGPD, article 46. Les garanties appropriées peuvent être de différentes natures et consister, par exemple, en l'adoption de clauses contractuelles types ou des règles d'entreprise contraignantes (une politique de protection des données intra-groupe en matière de transferts de données personnelles hors de l'Union européenne).
[320] CEPD, *Recommandations n°01/2020 sur les mesures qui complètent les instruments de transfert destinés à garantir le respect du niveau de protection des données à caractère personnel de l'UE*, 10 novembre 2020.
[321] CJUE, 16 juillet 2020, aff. C-311/18, *Data Protection Commissioner c/ Facebook Ireland*, CCE 2020. Comm. 4, obs. N. Metallinos.

requis, le responsable de traitement doit les cumuler avec des « mesures supplémentaires », qui sont, la plupart du temps, particulièrement difficiles à mettre en place[322], ou renoncer au transfert. Ainsi, le développement d'un système d'IA ayant recours, pour sa conception ou son fonctionnement, à un transfert de données à caractère personnel vers un pays « non-adéquat » se heurte à des nombreuses contraintes. Certaines de ces contraintes, telles que l'analyse du droit local pour en vérifier la compatibilité avec le droit de l'UE ou la vérification du respect des principes du droit de l'UE par l'importateur, sont surmontables. Il convient néanmoins de souligner qu'en pratique elles font peser une charge très importante sur les responsables de traitement, qui peuvent se retrouver face à des importateurs qui ne coopèrent pas ou face à une trop grande quantité d'analyses à mener suivant le nombre de sous-traitants, de premier rang ou de degrés ultérieurs. D'autant plus que la notion de « transfert de données » est particulièrement large et recouvre tout accès potentiel à des données personnelles, incluant l'hébergement de données, les accès pour maintenance ou encore les prestations infonuagiques (dites « *Cloud* »). Reconnaître en bonne pratique le partage systématique, par l'importateur, d'éléments utiles à la démonstration de la compatibilité de son droit local avec le RGPD et de son respect de ses principes, permettrait d'alléger la charge administrative qui pèse sur les seuls responsables de traitement. Enfin, l'impossibilité technique, dans la majorité des cas, de mettre en place des « mesures supplémentaires » lorsqu'elles sont nécessaires conduit les développeurs de solutions d'IA d'abandonner de nombreux projets.

138. **L'exemple du recours aux solutions des GAFAM et du transfert de données vers les États-Unis.** Les grandes entreprises américaines du numérique telles qu'Amazon avec sa filiale AWS, Google avec Google Cloud, ou Microsoft avec son service Azure, occupent une place omnipotente sur le marché du *Cloud computing*, ou « informatique en nuage ». Le recours à leurs solutions, ainsi qu'à leurs capacités de calcul décentralisées, est bien souvent

[322] CEPD, *Recommandations n°01/2020 sur les mesures qui complètent les instruments de transfert destinés à garantir le respect du niveau de protection des données à caractère personnel de l'UE,* 10 novembre 2020, spec. 84 et s. « Scénarios dans lesquels aucune mesure efficace n'a été trouvée » : Les « mesures supplémentaires » doivent aboutir à ce que l'importateur des données soit incapable d'accéder aux données en clair (non cryptées) ou que des mesures de pseudonymisation très robustes rendent impossible l'identification des personnes concernées par les données transmises (même par les services de renseignements locaux à l'aide des moyens dont sont susceptibles de disposer les services en question…). En pratique, ces mesures sont techniquement très compliquées à mettre en place.

incontournable pour le développement de solutions d'IA performantes. Il est possible de s'en passer, mais le temps requis pour développer les modèles seraient grandement allongés, leur performance amoindrie, et les solutions d'autres éditeurs, dont des européens, ne présentent pas le même niveau de sécurité que celles des GAFAM. Pour faciliter le commerce international des solutions de *Cloud computing*, les États-Unis avaient mis en place le *Privacy Shield*, un mécanisme d'auto-certification « validé » par la Commission Européenne et considéré comme une « garantie appropriée » au sens du RGPD. Selon ce dispositif, les entreprises désireuses de bénéficier du mécanisme devaient certifier leur conformité à un corpus de règles prédéfinies et s'inscrire sur la liste des organismes adhérents au *Privacy Shield*. Or, la CJUE a invalidé le *Privacy Shield* le 16 juillet 2020 par son arrêt « Schrems II »[323]. La Cour a en effet considéré que les lois américaines sur le renseignement portent des atteintes disproportionnées aux droits des personnes et mettent en échec les protections offertes par le *Privacy Shield* en raison de la primauté du droit américain sur les engagements des entreprises y adhérant. Or, le droit américain prévaut également sur les autres types de « garanties appropriées » qui ne sont que de « simples » accords contractuels. Les transferts de données personnelles vers les États-Unis sur le fondement de ces seules garanties deviennent donc illicites. Ainsi, tous les transferts de données personnelles vers les États-Unis, et donc chaque recours à des solutions éditées par des GAFAM ou leurs filiales[324], devront faire l'objet de « mesures supplémentaires » qui sont impossibles à mettre en œuvre en pratique. Des précisions sur les nouvelles versions des clauses contractuelles types, post-Schrems II, ont été apportées par la Commission européenne[325] mais il conviendrait de proposer aux responsables de traitement un cadre clair et réaliste en ce qui concerne les « mesures supplémentaires » à mettre en place.

---

[323] CJUE, 16 juillet 2020, aff. C-311/18, *Data Protection Commissioner c/ Facebook Ireland*, CCE 2020. Comm. 4, obs. N. Metallinos.

[324] En pratique, même les solutions proposées par les filiales européennes des entreprises américaines contiennent, à un degré plus ou moins proche, un transfert de données soit vers les États-Unis pour leur traitement par les ingénieurs de la maison mère, soit pour leur transfert ultérieur aux sous-traitants de la maison mère, basés partout dans le monde. Tracer le parcours des données transférées et s'assurer de la conformité de chaque transfert (si tant est qu'il en ait connaissance) devient rapidement très difficile pour les responsables de traitement.

[325] *Décision d'exécution (UE)2021/915 de la Commission du 4 juin 2021 relative aux clauses contractuelles types entre les responsables du traitement et les sous-traitants au titre de l'article 28, paragraphe 7, du règlement (UE) 2016/679 du Parlement européen et du Conseil et de l'article 29, paragraphe 7, du règlement (UE) 2018/1725 du Parlement européen et du Conseil*, publiée au JOUE n° L199/18 du 7 juin 2021.

139. **Les conséquences sur le développement de systèmes d'IA traitant de données à caractère personnel.** Les règles sur les transferts de données à caractère personnel contenues dans le RGPD, mises en œuvre par les autorités de protection des données et telles qu'interprétées par la CJUE, constituent un frein majeur au développement de l'IA, tous secteurs confondus. Ce constat est particulièrement dommageable dans le secteur de l'électricité où nombre d'applications de l'IA pourraient avoir un intérêt en matière de réduction de la consommation électrique ou de lutte contre les émissions de gaz à effet de serre. D'abord, l'application de ces règles génère de trop nombreuses contraintes règlementaires pour les responsables de traitement qui n'ont pas forcément les moyens humains, techniques et financiers pour conduire des analyses détaillées sous-traitants par sous-traitants. En effet, ces analyses représentent un coût non-négligeable, ce qui mène à l'abandon de nombreux projets prometteurs. Ensuite, il est souvent impossible de mettre en place des « mesures supplémentaires » telles que définies par le CEPD sans qu'elles ne privent le système d'IA de toute utilité. Enfin, privilégier des solutions d'éditeurs de pays « adéquats » ne nous apparaît pas comme une solution idéale à court terme en ce que ces solutions présentent, à l'heure actuelle, des niveaux de sécurité et de performance moindres que ceux des solutions des grands éditeurs. La situation des entreprises est donc très délicate puisqu'un niveau adéquat de conformité ne semble pas atteignable en l'état actuel de la règlementation et de la jurisprudence de la CJUE. Toutefois, plusieurs solutions semblent envisageables.

B/ Des solutions aux contraintes liées au régime des transferts de données en dehors de l'Union européenne

140. **Plan.** Plusieurs solutions sont envisageables. À court terme, il faudrait préciser les mesures supplémentaires à mettre en œuvre par les exportateurs de données pour sécuriser les transferts de données (**1**). À moyen terme, l'adoption d'une décision d'adéquation par la Commission européenne qui permettrait le transfert licite de données vers les États-Unis (**2**) et les nouvelles offres de Cloud de confiance (**3**) sont également très attendues par les entreprises. En effet, ces deux solutions leur permettraient d'accéder aux meilleures technologies tout en assurant leur conformité à la règlementation en matière de protection des données personnelles.

1. La précision des mesures supplémentaires requises pour encadrer les transferts de données

141. **La nécessaire complétion de la doctrine du CEPD en matière de transferts de données.** L'absence de mesures supplémentaires identifiées par le CEPD pour sécuriser les transferts de données vers des organismes étrangers pour des prestations nécessitant un accès en clair aux données, y compris si ce dernier est ponctuel, empêche les entreprises européennes d'avoir recours à ces solutions de façon licite. Il apparaît donc urgent de préciser les mesures supplémentaires complétant les garanties appropriées pour les transferts de données en dehors de l'Union européenne. Ces mesures devraient être suffisamment réalistes pour pouvoir être mises en œuvre par les entreprises. En particulier, les lignes directrices du CEPD[326] sur les mesures supplémentaires devraient être complétées pour les situations dans lesquelles les données sont transférées à des entreprises devant avoir accès à la donnée en clair. Cette situation couvre dans la pratique la majorité des services infonuagiques. La précision recommandée dans le présent paragraphe pourrait par exemple consister en l'exigence d'un chiffrement des données par des clés non accessibles à l'importateur de données[327] ainsi qu'en la mise en œuvre de mesures visant à garantir que ce dernier ne dispose pas d'un accès continu aux données. Ces mesures nous semblent, d'une part, pertinentes car elles minimisent le risque d'accès illicite aux données par des autorités étrangères. En effet, l'importateur ne sera vraisemblablement pas en capacité de fournir les données déchiffrées sauf à copier illicitement les données lors des accès ponctuels aux données. Les mesures sont, d'autre part, réalistes puisqu'elles correspondent à des solutions techniques disponibles sur le marché[328] ou déjà pratiquées par

[326] CEPD, *Recommandations n°01/2020 sur les mesures qui complètent les instruments de transfert destinés à garantir le respect du niveau de protection des données à caractère personnel de l'UE,* 10 novembre 2020, spec. 84.

[327] E. Maury, « Actualité Informatique et Libertés – Transfert de données à caractère personnel vers des pays tiers », *AJDA*, 2022, p. 1433. Le Conseil d'État a déjà jugé que le recours à des mécanismes de chiffrement par clés externes, complétant des engagements contractuels forts, permettait de fournir un niveau de protection des données ne pouvant être regardé comme manifestement insuffisant : CE, 12 mars 2021, *Doctolib*, ord. réf., n° 450163.

[328] Voir par exemple : THALES, « Les innovations de Thales en matière de gestion des clés de chiffrement aident les organisations à atteindre la souveraineté numérique dans les environnements hybrides et multicloud », *Site de Thalès Group (blog)*, 12 octobre 2022, disponible en ligne : <https://www.thalesgroup.com/fr/monde/groupe/press_release/innovations-thales-matiere-gestion-des-cles-chiffrement-aident>, consulté le 16 janvier 2023.

certains éditeurs[329]. Toutefois, une évolution du contexte règlementaire pourrait assouplir les contraintes pesant sur les exportateurs de données.

2. La perspective d'une décision d'adéquation concernant les États-Unis

142. **La perspective d'une décision d'adéquation concernant les États-Unis.** Il est probable que les transferts de données vers des importateurs établis sur le territoire américain soient à nouveau rendus licites par un nouveau texte européen. En effet, la Commission européenne a publié, le 13 décembre 2022, son projet de décision d'adéquation concernant la circulation sécurisée de données avec les États-Unis[330]. Ce projet intervient après un accord de principe conclu entre la Commission et le gouvernement américain en mars 2022[331] et l'adoption par l'administration Biden, le 7 octobre 2022, d'un décret exécutif relatif aux activités de renseignement[332]. Ce dernier prévoit notamment l'ajout de critères de proportionnalité et de nécessité pour les demandes d'accès aux données par des agences de renseignement ainsi que la création d'une entité devant laquelle des recours pourront être formés par les personnes placées sous surveillance. Si le projet est adopté, il permettra de fonder licitement les transferts de données vers les entreprises américaines qui respecteront les obligations contenues dans le « *EU-US Data Privacy Framework* » (par exemple l'obligation de supprimer les données à caractère personnel lorsqu'elles ne sont plus nécessaires à la finalité pour laquelle elles ont été collectées, et d'assurer la continuité de la protection lorsque des données à caractère personnel sont partagées avec des tiers)[333]. Les prochaines étapes sont la consultation (pour avis, non

[329] AMAZON WEB SERVICES, « FAQ sur la confidentialité des données », *Site d'AWS (blog)*, disponible en ligne : <https://aws.amazon.com/fr/compliance/data-privacy-faq/>, consulté le 16 janvier 2023.
[330] COMMISSION EUROPÉENNE, « Data protection: Commission starts process to adopt adequacy decision for safe data flows with the US », *Communiqué de presse*, 13 décembre 2022.
[331] *Ibid.*
[332] WHITE HOUSE, *Executive Order On Enhancing Safeguards For United States Signals Intelligence Activities*, 7 octobre 2022, disponible en ligne : <https://www.whitehouse.gov/briefing-room/presidential-actions/2022/10/07/executive-order-on-enhancing-safeguards-for-united-states-signals-intelligence-activities/>, consulté le 16 janvier 2023.
[333] COMMISSION EUROPÉENNE, « Questions & Answers: EU-U.S. Data Privacy Framework, draft adequacy decision », *Site de la Commission européenne*, 13 décembre 2022, disponible en ligne : <https://ec.europa.eu/commission/presscorner/detail/en/qanda_22_7632>, consulté le 16 janvier 2023.

contraignant) du Comité européen de la protection des données (CEPD) et des États membres de l'UE, puis la probable adoption de la décision d'adéquation au printemps 2023. Comme pour les précédents instruments juridiques permettant les transferts vers les États-Unis (le *Safe Harbor* puis le *Privacy Shield*), il est probable que la décision fasse l'objet d'un recours devant un tribunal national, qui pourra ensuite saisir la CJUE d'une question préjudicielle[334]. Cette dernière devra alors valider ou invalider la décision. Ce processus de recours a pris 2 ans dans le cadre du *Privacy Shield* (procédure devant la Haute Cour irlandaise en mai 2018, décision d'invalidation de la CJUE en juillet 2020). Dans l'intervalle, les transferts vers les États-Unis étaient licites. Ainsi, la probable adoption d'une décision d'adéquation par la Commission européenne au printemps 2023 pourrait donc ouvrir une période où les transferts de données vers les États-Unis seront autorisés. Un tel assouplissement de la législation permettrait aux entreprises, y compris dans le secteur de l'électricité, d'avoir recours de façon licite aux solutions technologies américaines. Toutefois, de nombreux arguments laissent penser que cette décision ne survivrait pas à un recours devant la CJUE[335]. De plus, il serait préférable que des solutions performantes et offrant un niveau élevé de protection des données à caractère personnel se développent au sein de l'Union européenne.

[334] L'avocat Maximilian Schrems, à l'origine de l'invalidation du *Privacy Shield*, a déjà annoncé son intention de remettre en cause la validité de la décision d'adéquation : NOYB, « Le nouveau décret américain a peu de chances de satisfaire à la législation européenne », *NOYB (blog)*, 7 octobre 2022, disponible en ligne : <https://noyb.eu/fr/le-nouveau-decret-americain-peu-de-chances-de-satisfaire-la-legislation-europeenne>, consulté le 16 janvier 2023 ; NOYB, « Déclaration sur la décision d'adéquation de la Commission européenne concernant les États-Unis », *NOYB (blog)*, 13 décembre 2022, disponible en ligne : <https://noyb.eu/fr/declaration-sur-la-decision-dadequation-de-la-commission-europeenne-concernant-les-etats-unis>, consulté le 16 janvier 2023.

[335] Ces arguments, qui ne seront pas étudiés dans la thèse car décorrélés du sujet, sont notamment présentés dans NOYB, « Déclaration sur la décision d'adéquation de la Commission européenne concernant les États-Unis », *op. cit.*

3. Le développement des offres de Cloud de confiance

143. **Le développement des offres de Cloud de confiance.** Une dernière solution pour permettre aux entreprises du secteur de l'électricité de disposer des meilleures solutions technologiques pour développer des applications d'IA vertueuses consisterait dans le recours à des solutions européennes offrant un niveau de protection élevé des données. Ces solutions sont souvent dénommées « Cloud de confiance », sans que ce terme ne soit défini dans un texte. Le développement de ces offres est devenu une priorité européenne[336] et encore plus en France[337]. À ce titre, l'ANSSI a développé un référentiel d'exigences offrant de nombreuses garanties de sécurité, notamment au regard des potentiels accès aux données par des autorités étrangères[338]. La CNIL considère par ailleurs que ce référentiel, appelé SecNumCloud 3.2[339], *« fournit une réponse qui est conforme by design aux exigences de la Cour en matière de protection des données dans le cloud »*[340]. De nombreuses offres commerciales répondant à ces exigences sont en cours de développement[341]. Il est donc très probable que les entreprises européennes disposent de solutions de Cloud conformes à la législation européenne dans les prochaines années.

[336] A. Cherki, « Gaia-X, ou les illusions perdues d'un cloud européen », *Contexte*, 30 mai 2022, disponible en ligne : <https://www.contexte.com/article/numerique/gaia-x-souverainete-cloud_150712.html>, consulté le 16 janvier 2023.
[337] J. Cheminat, « Le gouvernement réoriente sa stratégie sur le cloud de confiance », *Le Monde de l'Informatique*, 17 mai 2021, disponible en ligne : <https://www.lemondeinformatique.fr/actualites/lire-le-gouvernement-reoriente-sa-strategie-sur-le-cloud-de-confiance-82937.html>, consulté le 16 janvier 2023.
[338] ANSSI, « L'ANSSI actualise le référentiel SecNumCloud », *Site de l'ANSSI*, disponible en ligne : <https://www.ssi.gouv.fr/actualite/lanssi-actualise-le-referentiel-secnumcloud/>, consulté le 16 janvier 2023.
[339] ANSSI, *Référentiel d'exigences « Prestataires de services d'informatique en nuage » (SecNumCloud)*, version 3.2 du 21 septembre 2021.
[340] ANSSI, « L'ANSSI actualise le référentiel SecNumCloud », *op. cit.*
[341] DGE, « NUMSPOT : un partenariat pour un nouveau service de cloud de confiance », Site de la DGE (blog), 26 octobre 2022, disponible en ligne : <https://www.entreprises.gouv.fr/fr/actualites/numerique/numspot-partenariat-pour-nouveau-service-de-cloud-de-confiance>, consulté le 15 janvier 2023 ; D. Filippone, « Cloud de confiance : Thales et Google Cloud répliquent à Bleu avec S3NS », *Le Monde de l'Informatique*, 30 juin 2022, disponible en ligne : <https://www.lemondeinformatique.fr/actualites/lire-cloud-de-confiance-thales-et-google-cloud-repliquent-a-bleu-avec-s3ns-87262.html>, consulté le 16 janvier 2023.

144. **Conclusion du Chapitre 2 sur l'excès de la protection accordée par le régime juridique applicable au traitement de données à caractère personnel.** Deux constats principaux se dégagent de l'analyse du régime de protection des données lorsqu'il est appliqué à des systèmes d'IA. Premièrement, les règles et obligations contraignantes qu'il contient (licéité, transparence, limitation des finalités…) s'y appliquent sans difficulté du fait de la neutralité technologique des textes : il n'existe pas de flou juridique. Deuxièmement, en revanche, cette application indifférenciée contraint le développement de systèmes d'IA, conduisant des responsables de traitement à abandonner de nombreuses potentielles applications vertueuses. Certaines de ces contraintes sont justifiées, mais d'autres apparaissent comme disproportionnées. L'articulation entre les principes du RGPD et ceux sous-tendant le fonctionnement de l'IA mériterait d'être clarifiée afin de ne pas entraver sa performance, source de bénéfices potentiels[342]. De plus, il est urgent d'utiliser les récentes avancées dans le domaine de l'explicabilité des systèmes d'IA qui peuvent pallier leur opacité en proposant de nouvelles formes d'explications, distinctes suivant le public visé et le but poursuivi[343]. Enfin, l'absence de consensus sur les « mesures supplémentaires » requises pour les transferts de données en dehors de l'Union européenne, en particulier vers les États-Unis, ainsi que la réticence des autorités de régulations quant à leur efficacité constituent un handicap majeur pour les entreprises européennes dans le développement de l'IA[344]. En effet, en pratique, le régime des transferts de données en dehors de l'Union européenne post-Schrems II les prive de la possibilité d'avoir recours aux solutions d'IA les plus performantes et sécurisées, développées par les grands éditeurs. Ces constats s'inscrivent dans la doctrine juridique majoritaire considérant que le droit européen de la protection des données présente à la fois des forces (sa neutralité technologique) et des faiblesses (l'absence de prise en compte des bénéfices attendus de la technologie) lorsqu'il est appliqué à des systèmes d'IA[345]. Les propositions présentées

[342] Voir Supra, 120-125.
[343] Voir Supra, 126-135.
[344] Voir Supra, 137-139.
[345] C. Copain-Héritier, « Le cadre européen de la protection des données entre forces et faiblesses intrinsèques », *Revue de l'Union européenne*, 2021, p. 163 ; M. Gérot et W. Maxwell, « Le RGPD pourrait freiner les ambitions de l'Europe en matière d'intelligence artificielle », *ecommercemag (blog)*, 30 mars 2020, disponible en ligne : < https://www.ecommercemag.fr/Thematique/management-1225/Breves/Tribune-RGPD-pourrait-freiner-ambitions-Europe-matiere-intelligence-artificielle-348359.htm> consulté le 15 avril 2020 ; C. Castets-Renard, « L'intelligence artificielle, les droits fondamentaux et la protection des données personnelles dans l'Union européenne et les États-Unis », *Revue de Droit International d'Assas*, 2019, 2, 158-174.

précédemment nous semblent à même de minimiser ses faiblesses et de faire du droit des données personnelles un véritable rempart contre les risques de l'IA, tout en préservant suffisamment de souplesse pour permettre son développement dans l'Union européenne.

145. **Conclusion du Titre 1 sur les incertitudes juridiques résultant de l'application du Droit actuel aux systèmes d'IA.** Les spécificités de l'IA, à savoir la complexité, l'opacité, l'autonomie et la dépendance à la donnée, complexifient parfois l'application des règles de droit existantes. Le présent Titre a permis de mettre en évidence l'applicabilité et les limites de deux régimes juridiques non-sectoriels, perçus comme les plus contraignants par les acteurs de l'IA dans l'énergie : les régimes de responsabilité et de la protection des données. Les premiers apparaissent comme suffisamment adaptables pour couvrir les risques de dommages générés par les systèmes d'IA. Les fonctions du droit de la responsabilité ont été éprouvées par les âges et méritent d'être préservées. Un encadrement de la conception des systèmes d'IA, concentré sur ses usages qui présenteraient des risques particuliers ou mettraient à mal l'applicabilité des règles existantes (autonomie complète ou auto-apprentissage au cours du cycle de vie notamment), permettrait de pallier ses limites résiduelles. En second lieu, les droits français et européen de la protection des données sont pleinement applicables aux systèmes d'IA malgré leurs spécificités.

Toutefois, nous avons pu constater que cette application indifférenciée, si elle permet de couvrir la majorité des risques de l'IA pour la vie privée des individus, génère un frein considérable à son développement. En pratique, l'application des principes du RGPD, l'impossible information sur le fonctionnement d'un algorithme qu'on ne peut expliquer ou l'obstacle au recours aux technologies d'IA les plus performantes et sécurisées, situées en dehors de l'Union européenne, constituent autant de contraintes disproportionnées à l'adoption de l'IA. Pourtant, il existe des moyens pour adapter le droit existant de façon raisonnée et prendre en considération les spécificités de l'IA, en préservant la protection des individus, tout en donnant suffisamment de souplesse pour ne pas entraver l'innovation.

A défaut d'une telle adaptation, les entreprises seraient placées dans une situation d'incertitudes quant aux règles applicables, à leur interprétation au regard des spécificités des

systèmes d'IA, ou encore au risque de superposition avec d'autres législations[346]. Ces incertitudes conduisent en pratique à une hausse du coût de la mise en conformité, dissuadent l'investissement dans la technologie et, selon le niveau d'aversion au risque de l'entreprise concernée, peuvent causer l'abandon de nombreux projets prometteurs. Comme cela a été démontré, ces incertitudes peuvent toutefois être gommées par une adaptation raisonnée du corpus juridique actuel ou sa précision par la voie du droit dit « souple ».

146. **Transition.** Afin de poursuivre notre étude visant à identifier les adaptations nécessaires du Droit pour favoriser le développement éthique et durable de l'IA dans le secteur de l'électricité, notre réflexion doit également être étendue aux règlementations sectorielles qui sont particulièrement nombreuses dans ce domaine.

[346] CEPD, *Avis 4/2020, 29 juin 2020 sur le livre blanc de la Commission sur l'intelligence artificielle*, spec. pts 55-61, mettant en évidence les risque de chevauchement entre les exigences d'un régime propre à l'IA et les règles existantes, notamment en matière de protection des données.

Titre 2 : Une contrainte disproportionnée résultant de l'application de réglementations sectorielles

147. **Plan.** Des règles spécifiques au secteur de l'électricité viennent régir la plupart des activités le composant : de la production à la commercialisation, en passant par le transport et la distribution. Toutes n'ont pas vocation à s'appliquer à l'utilisation de systèmes d'IA mais une partie d'entre elles peuvent contraindre leur développement. Les observations réalisées sur le terrain et les entretiens réalisés dans le cadre de notre étude ont mis en évidence deux cas d'usage de l'IA dans le secteur de l'électricité qui se trouvent grandement contraints en raison de l'application de règlementations sectorielles. Le premier correspond à l'utilisation de l'IA dans les infrastructures critiques du système électrique, à savoir les réseaux et les centrales de production, y compris nucléaires. La conception et l'exploitation de ces infrastructures doivent respecter un important corpus de règles relatives à la sécurité, notamment informatique mais pas seulement. Avec le temps, de nombreux standards se sont développés et sont suivis à la lettre par les opérateurs, laissant peu de place à l'innovation. En effet, l'ensemble de ces règles encadrant la sécurité des infrastructures critiques de l'énergie ont été pensées dans un contexte technologique différent, où les logiciels informatiques étaient majoritairement déterministes et fondés sur l'application de règles. L'apparition de l'apprentissage automatique remet en cause ces paradigmes, ce qui génère une incompatibilité entre les systèmes d'IA fondés sur ces techniques et la règlementation relative à la sécurité des infrastructures critiques qui n'y est pas adaptée (**Chapitre 1**). Le second cas d'usage contraint par des règlementations sectorielles relève du traitement de données énergétiques par des systèmes d'IA. Les données énergétiques sont trop peu disponibles ou alors leur traitement fait l'objet d'un encadrement très strict, empêchant leur mobilisation pour la recherche de solutions innovantes (**Chapitre 2**).

Chapitre 1 : Une inadaptation manifeste des règles de sécurité dans les systèmes critiques

148. **Introduction aux contraintes de sécurité applicables aux systèmes d'information utilisés dans des infrastructures critiques.** Au début des années 2000 et à la suite des attentats de New-York, Londres et Madrid, la France a engagé une réflexion sur la protection de ses infrastructures les plus critiques. Elle a abouti à la création d'une législation relative à la sécurité des « activités d'importance vitale » (AIV), codifiée dans le Code de la défense en 2005[347]. Ces activités ont été définies pour la première fois dans un décret du 23 février 2006[348]. Elles regroupent, d'une part, la production et la distribution de biens ou de services indispensables à la satisfaction des besoins essentiels de la population, à l'exercice de l'autorité de l'État, au fonctionnement de l'économie, à la défense nationale ou à la sécurité de la nation, dès lors que ces activités sont difficilement substituables ou remplaçables, et toute activité présentant un danger grave pour la population, d'autre part. Les secteurs d'activités considérés comme « d'importance vitale » sont listés dans un arrêté du 2 juin 2006 et incluent notamment l'industrie, les transports, la santé, les télécommunications, la finance et, pour ce qui concerne la présente thèse, l'énergie[349]. Chaque secteur d'activité est placé sous la tutelle d'un ministère. Pour le secteur objet de notre étude, l'énergie, il s'agit aujourd'hui du Ministère de la transition écologique. Dans chacun des secteurs d'activités visés, les opérateurs gérant ou utilisant *« une ou des installations dont le dommage ou l'indisponibilité ou la destruction par suite d'un acte de malveillance, de sabotage ou de terrorisme risquerait, directement ou indirectement : a) D'obérer gravement le potentiel de guerre ou économique, la sécurité ou la capacité de survie de la Nation ; b) Ou de mettre gravement en cause la santé ou la vie de la population »*[350] sont qualifiés « d'opérateurs d'importance vitale » (OIV). En raison de leur criticité, les activités composant le secteur de l'électricité sont concernées par ces dispositions. En particulier, la

[347] *Loi n° 2005-1550 du 12 décembre 2005 modifiant diverses dispositions relatives à la défense créant les articles L1332-1 et suivants du Code de la défense*, publiée au JORF n°289 du 13 décembre 2005, texte n° 2.
[348] *Décret n°2006-212 du 23 février 2006 relatif à la sécurité des activités d'importance vitale*, publié au JORF n°47 du 24 février 2006, texte n° 1, article 2, aujourd'hui codifié aux articles R1332-1 et suivants du Code de la Défense.
[349] *Arrêté du 2 juin 2006 fixant la liste des secteurs d'activités d'importance vitale et désignant les ministres coordonnateurs desdits secteurs*, publié au JORF n°129 du 4 juin 2006, texte n° 1, annexe.
[350] Code de la Défense, article R1332-1.

gestion des réseaux de transport et des ouvrages de production d'électricité concourent à la sécurité d'approvisionnement du pays et constituent donc un intérêt vital pour la Nation. Les OIV doivent identifier et recenser au sein de leurs installations les *« points d'importance vitale »*[351], traduction locale de l'anglais *« critical infrastructures »* ou *« critical systems »*[352]. Les systèmes d'information constituant des points d'importance vitale sont qualifiés de *« systèmes d'information d'importance vitale »* (SIIV) et doivent faire l'objet de mesures renforcées de protection depuis la loi de programmation militaire de 2013[353].

149. **Déclinaison dans le secteur étudié.** Dans le secteur de l'électricité, un grand nombre de logiciels utilisés dans les réseaux de transport et de distribution ainsi que dans la gestion, l'exploitation et la protection des centrales nucléaires de production d'électricité peuvent être qualifiés de SIIV. Les systèmes d'IA utilisés dans le cadre de ces activités pourraient aussi relever de cette qualification et donc se conformer à un important corpus de règles qui seront étudiées dans ce Chapitre. De plus, d'autres règles peuvent venir s'additionner à ces contraintes en fonction de la nature de l'activité concernée. Ainsi, la production électronucléaire relève d'une règlementation spécifique relative à la « sûreté nucléaire » dont certaines dispositions peuvent également entraver le recours à l'IA. Notre objectif ici sera de questionner la règlementation en vigueur pour déterminer si la contrainte qu'elle impose sur l'utilisation de l'IA est justifiée et équilibrée. Peut-on réconcilier innovation et protection de la sécurité dans les infrastructures critiques du secteur de l'électricité ?

150. **Plan.** Après avoir présenté, à titre liminaire, les raisons pour lesquelles l'utilisation de l'IA dans des systèmes critiques de l'énergie pourrait être souhaitable (**Section liminaire**), le contenu de la règlementation sectorielle en matière de sécurité des infrastructures critiques devra être étudié afin d'identifier les limites qu'elle impose à l'utilisation de l'IA dans ce contexte (**Section 1**). Cette analyse permettra, en second lieu, de réfléchir aux potentielles actions ou adaptations qui pourraient être entreprises pour rééquilibrer le corpus existant et

---

[351] SECRETARIAT GENERAL DE LA DEFENSE ET DE LA SECURITE NATIONALE, *Instruction générale interministérielle relative à la sécurité des activités d'importance vitale*, n°6600/SGDSN/PSE/PSN, 7 janvier 2014, spec. p. 21.
[352] J.-P. Galland, « Critique de la notion d'infrastructure critique », *Flux*, 2010, vol. 81, n°3, pp. 6-18, note 18.
[353] *Loi n° 2013-1168 du 18 décembre 2013 relative à la programmation militaire pour les années 2014 à 2019 et portant diverses dispositions concernant la défense et la sécurité nationale*, publiée au JORF n°0294 du 19 décembre 2013, article 22.

permettre le recours à des systèmes d'IA sûrs tout en conservant des garde-fous suffisants pour garantir la sécurité des infrastructures critiques (**Section 2**).

Section liminaire : L'utilité de l'IA dans les infrastructures critiques du secteur de l'électricité

151. **Plan.** Des systèmes d'IA peuvent être employés dans toutes les activités du secteur de l'électricité. Leurs potentielles applications dans certaines fonctions critiques des réseaux de transport et de distribution (**§1**) et des ouvrages de production d'électricité, notamment nucléaire (**§2**), pourraient apporter de nombreux bénéfices, accompagnés d'autant de risques nouveaux.

§1 : L'utilisation souhaitable de l'IA dans la gestion des réseaux de transport et de distribution d'électricité

152. **L'enjeu de la numérisation des réseaux de transport et de distribution.** De nombreux cas d'usage de l'IA dans la gestion et la maintenance des réseaux peuvent concerner des fonctions critiques. En effet, ces activités nécessitent un pilotage en temps quasi-réel pour équilibrer la quantité d'électricité injectée sur le réseau avec celle effectivement consommée par les consommateurs finals. Il s'agit d'un maillon essentiel de la sécurité d'approvisionnement en électricité. Tout défaut de fonctionnement peut conduire à des coupures de courant pouvant entraîner des conséquences désastreuses pour la Nation : arrêt des usines de production ou des hôpitaux, effondrement du système informatique bancaire… Pour assurer son pilotage, des opérateurs humains doivent en surveiller continuellement le fonctionnement et déclencher les actions nécessaires en cas d'alertes (défaillance d'un matériel, anomalie dans les données remontées, dégradation intentionnelle, panne de courant localisée). Cette mission est aujourd'hui assurée en France par RTE pour le réseau de transport d'électricité et Enedis pour le réseau de distribution. Elle requiert le traitement de grandes quantités d'information en temps réel et des prises de décision rapides. À cet égard, le recours aux technologies numériques peut être d'un grand secours, si bien que la gestion des réseaux s'est progressivement digitalisée même si elle reste encore aujourd'hui majoritairement manuelle. Au cours des deux dernières décennies, s'est développé un nouvel intérêt pour les techniques d'apprentissage automatique qui promettent des performances et des capacités de traitement de données encore accrues. Le présent Chapitre se concentrera sur ces dernières en excluant les

systèmes d'IA « déterministes » correspondant aux systèmes experts déjà utilisés depuis longtemps et pour lesquels la règlementation en matière de sécurité a été pensée à l'origine[354].

153. **Les bénéfices de l'utilisation de l'IA dans les réseaux de transport et de distribution.** L'utilisation potentielle de systèmes d'IA dans le pilotage des réseaux de transport et de distribution d'électricité a déjà fait l'objet de nombreux articles académiques[355]. L'IA pourrait ainsi être utilisée pour automatiser la résolution d'une panne de courant ou pour les prévenir[356], pour garantir la stabilité de la fréquence des réseaux électriques[357], ainsi que pour aider en temps réel les opérateurs humains à prendre les bonnes décisions[358]. Ce dernier exemple est notamment étudié en France par RTE[359]. Ces applications pourraient apporter de nombreux bénéfices. Tout d'abord, elles permettraient d'améliorer la qualité d'approvisionnement en électricité puisqu'elles faciliteraient la maintenance des réseaux et la résolution des incidents techniques en les détectant plus rapidement. Ensuite, ces cas d'usage permettraient de rendre la gestion du réseau plus flexible puisque des actions pourraient être déclenchées automatiquement ou du moins beaucoup plus rapidement. Il s'agit là d'un enjeu clé dans un contexte de développement des énergies renouvelables qui multiplie les sources de production d'électricité raccordée aux réseaux et complexifie donc sa gestion. Enfin, l'utilisation de systèmes d'IA dans les réseaux permettrait aux opérateurs historiques, Enedis et RTE, d'optimiser leur gestion, notamment en facilitant la maintenance. Cette optimisation contribuera à la réduction des coûts pour la collectivité, qui constituent une partie du prix de l'électricité. Toutefois, dans le cas où de tels systèmes interagissaient avec des fonctions critiques des réseaux concernés, ils pourraient être qualifiés de SIIV ou à tout le moins de

[354] Pour une taxonomie plus précise des techniques auxquelles nos développements s'intéressent, voir F. Mamalet, E. Jenn, G. Flandin, *et al.*, *Machine learning in certified systems*, livre blanc, ANITI – IRT Saint-Exupéry, 2021, p. 14, table 1.
[355] Pour une étude systémique de la littérature sur le sujet, voir notamm. A. Sozontov, M. Ivanona, A. Gibadullin, « Implementation of artificial intelligence in the electric power industry », *E3S Web of Conferences*, 2019, vol. 114, n°01009, spec. p. 2 ; A. Mogilenko, « Application of artificial intelligence algorithms in the global energy industry », *Energy and Industry of Russia*, 2018, n°7.
[356] S. Ivanov, « Artificial Intelligence will be able to deal with blackouts in power grids », *Hi-Tech*, 2019, n°1, pp. 28-31.
[357] J. Kruse, B. Schäfer, D. Witthaut, « Revealing drivers and risks for power grid frequency stability with explainable AI », *Patterns*, vol. 2, n°11, 2021.
[358] A. Marot, A. Rozier, M. Dussartre, *et al.*, « Towards an AI assistant for human grid operators », *Hybrid Human Artificial Intelligence (HHAI) Conference*, juin 2022, Amsterdam.
[359] *Ibid.*

composant d'un SIIV en raison des risques qu'ils pourraient générer. Cette qualification conduit à l'application d'un important corpus de règles contraignantes. Le constat est similaire lorsque l'on s'intéresse aux cas d'usage potentiels de l'IA dans la production d'électricité, notamment électronucléaire.

§2 : L'utilisation souhaitable de l'IA dans la production électronucléaire

154. **Une intégration dans des systèmes d'information déjà complexes.** Des systèmes d'IA peuvent être utilisés dans le système d'information des centrales nucléaires de production d'électricité afin d'optimiser leur fonctionnement, automatiser des actions récurrentes ou encore pour renforcer la sécurité. Le système d'information des centrales est particulièrement riche et complexe mais repose aujourd'hui exclusivement sur des logiciels déterministes classiques, aisément auditables. Toutefois, le développement de nouveaux types de logiciels, fondés non pas sur l'application de règles mais sur des techniques d'apprentissage automatique, fait aujourd'hui l'objet de nombreuses recherches très actives, dont quelques exemples sont donnés ci-après.

155. **Des applications potentielles de l'IA dans les centrales nucléaires.** Des chercheurs russes ont réalisé un inventaire des principales applications potentielles de l'IA dans le secteur de l'énergie atomique en 2019, dans lequel ils ont recensé neuf grandes catégories de cas d'usage[360]. Cinq d'entre elles pourraient donner lieu à la conception de systèmes d'IA intégrés au système d'information de la centrale et pouvant interagir avec certaines fonctions critiques pour la sécurité : l'identification de défauts dans le réacteur ; l'analyse en temps réel des données de fonctionnement ; la réponse aux incidents nucléaires ; l'automatisation de fonctions de sûreté ; la cybersécurité, la détection et la gestion automatiques des attaques informatiques[361].

[360] A. Sozontov, M. Ivanona, A. Gibadullin, *op. cit.*, spec. p. 2.
[361] *Ibid.*

156. **Bénéfices et risques de l'utilisation de l'IA dans les centrales nucléaires.** Les capacités de l'IA en matière de quantité de données traitées, de précision et de rapidité laissent certains experts penser que son utilisation dans les centrales nucléaires permettrait d'en optimiser le fonctionnement et la sûreté[362]. En effet, les catégories de cas d'usage mentionnées précédemment illustrent bien les potentiels bénéfices que l'IA pourrait apporter à l'industrie nucléaire : optimisation de la maintenance, identification des défauts pouvant conduire à des incidents de sûreté, diminution du risque d'erreurs humaines ou encore accroissement du niveau de sécurité informatique des installations. Pourtant, son introduction dans les SIIV d'une centrale pourrait faire naître de nouveaux risques du fait de ses différences (autonomie, opacité, dépendance à la donnée, complexité) avec les systèmes déterministes utilisés aujourd'hui. Certains auteurs ont ainsi mis en évidence le paradoxe existant entre la volonté des hommes à automatiser les infrastructures qu'ils considèrent vulnérables (pour optimiser le fonctionnement et réduire le risque d'erreur humaine) et le fait que cette automatisation génèrerait en réalité de nouvelles vulnérabilités (informatiques)[363]. En effet, le recours à l'IA dans un environnement si complexe et critique pour la sécurité qu'une centrale nucléaire peut, s'il n'est pas correctement encadré, conduire à la création de nouvelles cibles pour les pirates informatiques. Fort heureusement, le régime juridique applicable aux SIIV constitue un garde-fou efficace, au risque de brider l'innovation.

[362] S. Suman, « Artificial Intelligence in Nuclear Industry: Chimera or Solution? », *Journal of Cleaner Production*, 1er janvier 2021, vol. 278, 124022.

[363] M.C. Horowitz, « Artificial intelligence and nuclear stability: Euro-Atlantic perspectives », in *The Impact of Artificial Intelligence on Strategic Stability and Nuclear Risk*, dir. V. Boulanin, SIPRI Press, 2019, vol. 1, p. 79 ; E. Swaton, V. Neboyan, L. Lederman, « Human factors in the operation of nuclear power plants : Improving the way man and machines work together », *IAEA Bulletin*, 4/1987, disponible en ligne <https://www.iaea.org/sites/default/files/publications/magazines/bulletin/bull29-4/29405042730.pdf>, consulté le 9 septembre 2022.

Section 1 : La sécurité des infrastructures critiques : une limite légitime à l'utilisation de l'IA dans le secteur de l'électricité

157. **Plan.** Les limites à l'utilisation de l'IA dans les systèmes critiques du secteur de l'électricité résultent de l'application du régime juridique encadrant les SIIV, d'une part (**§1**), et des règles spécifiques à la sécurité nucléaire, d'autre part (**§2**). Ces limites sont justifiées dans la mesure où la défaillance ou le manque de contrôle sur les systèmes d'information utilisés dans ces environnements sensibles pourraient entraîner des conséquences désastreuses pour la sécurité des biens et des personnes, que ce soit en raison des pannes de courant causées par un dysfonctionnement des infrastructures de réseaux ou des potentiels rejets de matières radioactives dans l'environnement causés par un incident dans une centrale nucléaire. Certaines caractéristiques spécifiques des systèmes d'IA telles que le défaut d'explicabilité ou la potentielle autonomie les rendent difficilement compatibles avec les contraintes de sécurité dans les infrastructures critiques du secteur de l'électricité.

§1 : Des limites liées au régime juridique de la sécurité des systèmes d'information d'importance vitale (SIIV)

158. **Plan.** Le régime juridique des SIIV est particulièrement contraignant (**A**). Son contenu, pensé dans un contexte technologique antérieur où régnaient les systèmes logiciels déterministes, ne semble pas adapté pour appréhender les spécificités des systèmes d'IA (**B**).

A/ Le régime juridique des SIIV appliqué aux systèmes d'IA

159. **Les textes de référence en matière de sécurité des SIIV.** L'approvisionnement en énergie électrique fait partie des secteurs d'activité considérés comme « d'importance vitale » pour les intérêts de la Nation. Bien que la liste des OIV soit couverte par le secret de la défense nationale, les gestionnaires des réseaux de transport et de distribution, ainsi que les entreprises exploitant des ouvrages importants pour l'approvisionnement en électricité du pays, tels que les centrales nucléaires ou les barrages hydroélectriques, sont naturellement visés par les règles

issues de la loi du 12 décembre 2005[364], aujourd'hui codifiées aux articles L1332-1 et suivants du Code de la Défense. Les systèmes d'information que ces OIV du secteur de l'électricité utilisent dans les installations critiques dont ils ont la charge peuvent être qualifiés de SIIV et sont encadrés par un corpus de règles spécifiques depuis la loi de programmation militaire (LPM) de 2013[365]. Ces règles ont été précisées par le pouvoir règlementaire et déclinées dans des arrêtés sectoriels en 2016, dont un dédié à l'activité « d'approvisionnement en énergie électrique »[366] qui constituera la source principale étudiée dans la suite du présent Paragraphe. Il existe également de nombreux textes d'application, de référentiels techniques ou guides pratiques publiés par l'ANSSI dont l'étude permet une meilleure compréhension du contexte règlementaire à l'usage de l'IA dans des systèmes critiques. Ce corpus juridique, très fourni et aujourd'hui éprouvé par le temps, contribue à un ensemble plus large que le Professeur Thibault Douville nomme le « droit commun de la cybersécurité »[367]. En effet, il existe d'autres règlementations relatives à la sécurité informatique qui ne sont pas spécifiques aux SIIV. C'est le cas, par exemple, du régime encadrant les activités des « opérateurs de services essentiels » (OSE) issu du droit européen.

160. **Le régime redondant des systèmes d'information essentiels (SIE) issu de la directive européenne « Network Information Security » (NIS).** La directive NIS est le premier texte législatif à l'échelle de l'UE en matière de cybersécurité[368]. Il prévoit des mesures juridiques visant à instaurer un socle commun de sécurité informatique pour tous les systèmes

[364] *Loi n° 2005-1550 du 12 décembre 2005 modifiant diverses dispositions relatives à la défense créant les articles L1332-1 et suivants du Code de la défense*, publiée au JORF n°289 du 13 décembre 2005, texte n° 2.
[365] *Loi n° 2013-1168 du 18 décembre 2013 relative à la programmation militaire pour les années 2014 à 2019 et portant diverses dispositions concernant la défense et la sécurité nationale*, publiée au JORF n°0294 du 19 décembre 2013, article 22.
[366] *Arrêté du 11 août 2016 fixant les règles de sécurité et les modalités de déclaration des systèmes d'information d'importance vitale et des incidents de sécurité relatives au sous-secteur d'activités d'importance vitale « Approvisionnement en énergie électrique » et pris en application des articles R. 1332-41-1, R. 1332-41-2 et R. 1332-41-10 du code de la défense*, publié au JORF n°0197 du 25 août 2016, texte n° 4.
[367] T. Douville, « L'émergence d'un droit commun de la cybersécurité », *Recueil Dalloz*, 2017, p. 2255, dressant notamment un panorama de l'ensemble des réponses juridiques, législatives ou règlementaires, apportées face à l'accroissement du risque lié à la cybersécurité et commençant à composer un ensemble cohérent, véritable droit commun de la sécurité informatique.
[368] E. Maupin, « Adoption de la loi sur la cyber-sécurité, les armes à feu et Galiléo », *AJDA*, 2018, p. 368 ; *Directive (UE) 2016/1148 du Parlement européen et du Conseil du 6 juillet 2016 concernant des mesures destinées à assurer un niveau élevé commun de sécurité des réseaux et des systèmes d'information dans l'Union*, publiée au JOUE n°L194/1 du 19 juillet 2016, dite « directive NIS » ou « directive SRI ».

d'information utilisés par des OSE[369], défini en droit français dans une loi de transposition de 2018 comme les *« opérateurs, publics ou privés, offrant des services essentiels au fonctionnement de la société ou de l'économie et dont la continuité pourrait être gravement affectée par des incidents touchant les réseaux et systèmes d'information nécessaires à la fourniture desdits services »*[370]. On peut à première vue s'étonner de la similarité de cette définition avec celle des OIV figurant à l'article L1332-1 du Code de la Défense : *« opérateurs publics ou privés exploitant des établissements ou utilisant des installations et ouvrages, dont l'indisponibilité risquerait de diminuer d'une façon importante le potentiel de guerre ou économique, la sécurité ou la capacité de survie de la nation »*[371]. L'annexe II de la directive NIS vise expressément les opérateurs impliqués dans la fourniture, le transport et la distribution d'électricité[372] alors même que l'approvisionnement en énergie électrique figure déjà au rang des secteurs d'activités d'importance vitale au titre du Code de la Défense[373]. Au-delà des définitions, le contenu des règles à respecter dans chacun des deux régimes est très proche. Dans une recommandation du 18 décembre 2020 relative à la protection des systèmes d'information essentiels (SIE), l'ANSSI détaille les règles applicables aux SIE en faisant référence à celles encadrant les SIIV et va même jusqu'à proposer un tableau de correspondance complet[374]. Les référentiels techniques auxquels l'autorité nationale renvoie dans sa bibliographie sont également les mêmes que ceux utilisés pour la mise en conformité des OIV[375]. Dès lors, pourquoi ces deux régimes cohabitent-ils et quelle est la justification de cette apparente redondance ? La lecture des travaux parlementaires préalables à l'adoption de la loi

[369] Pour une analyse juridique des objectifs et du contenu de la directive NIS, voir notamm. B. Bertrand, « Chronique Droit européen du numérique - La nouvelle approche de la cybersécurité européenne », *RTD Eur.*, 2021, p. 155.
[370] *Loi n° 2018-133 du 26 février 2018 portant diverses dispositions d'adaptation au droit de l'Union européenne dans le domaine de la sécurité*, publiée au JORF n°0048 du 27 février 2018, texte n° 2, article 5.
[371] Code de la Défense, article L1332-1.
[372] Directive NIS, Annexe II, a).
[373] *Arrêté du 2 juin 2006 fixant la liste des secteurs d'activités d'importance vitale et désignant les ministres coordonnateurs desdits secteurs*, publié au JORF n°129 du 4 juin 2006, texte n° 1, annexe.
[374] ANSSI, *Recommandations pour la protection des systèmes d'information essentiels*, guide ANSSI, 18 décembre 2020, ANSSI-PA-085, Annexe A.
[375] *Ibid.*, Bibliographie, p. 101 et s. concernant les SIE ; ANSSI, « Les règles de sécurité », *Site de l'ANSSI (blog)*, disponible en ligne : <https://www.ssi.gouv.fr/administration/protection-des-oiv/les-regles-de-securite/>, consulté le 14 septembre 2022, concernant les SIIV. Voir notamm. les référentiels PDIS, PRIS, PASSI, EBIOS pour l'analyse des risques…

de transposition de la directive NIS[376], et notamment l'étude d'impact[377], apporte un début de réponse. En effet, les parlementaires y expliquent que les deux dispositifs s'appuient sur des fondements juridiques distincts et des finalités différentes. Selon eux, la directive NIS vise « *à assurer le fonctionnement des activités économiques et sociétales dans le cadre du marché intérieur* »[378], tandis que le dispositif applicable aux OIV *« s'inscrit dans une stratégie de sécurité nationale »*[379]. Ces derniers sont définis à partir de critères plus discriminants et matériels, tels que l'exploitation ou l'utilisation « d'installations et d'ouvrages », là où les OSE sont uniquement identifiés à partir des services fournis et de leur dépendance à des systèmes d'information. Le régime des OSE a donc une portée plus large et ne se limite pas aux opérateurs exploitant des infrastructures dont la défaillance ferait courir un danger grave à la défense et à la sécurité de la Nation. De plus, le régime des OSE se concentre uniquement sur les systèmes d'information, alors que celui des OIV impose également la protection physique des « points d'importance vitale ». On regrettera cependant que les parlementaires, dans leur étude d'impact, n'aient pas justifié l'articulation entre les règles des deux régimes qui s'avèrent en réalité similaires. Heureusement, l'article 5 de la loi de transposition[380] prévoit que les règles applicables aux OSE ne sont pas applicables aux systèmes d'information des OIV dans le cas où le secteur d'activité concerné serait couvert par les deux régimes, ce qui est le cas de l'approvisionnement en énergie électrique. Le présent Chapitre entend étudier les règles applicables aux systèmes d'IA utilisés dans les infrastructures critiques du secteur de l'électricité. Au vu de la sensibilité de ces installations (centrales de production, réseaux de transport et de distribution), leurs exploitants peuvent être qualifiés d'OIV au sens du Code de la Défense pour une grande partie de leur activité. Les systèmes d'IA utilisés dans leur système

[376] C. Euzet, *Rapport fait au nom de la commission des lois constitutionnelles, de la législation et de l'administration générale de la république, sur le projet de loi, adopté par le sénat après engagement de la procédure accélérée, portant diverses dispositions d'adaptation au droit de l'Union européenne dans le domaine de la sécurité (n° 530)*, texte n°554.
[377] SENAT, *Etude d'impact sur le projet de loi portant diverses dispositions d'adaptation au droit de l'union européenne dans le domaine de la sécurité*, 17 novembre 2017, NOR : INTX1728622L/Bleue-1.
[378] *Ibid.*, p. 17.
[379] *Ibid.*
[380] *Loi n° 2018-133 du 26 février 2018 portant diverses dispositions d'adaptation au droit de l'Union européenne dans le domaine de la sécurité*, publiée au JORF n°0048 du 27 février 2018, texte n° 2, article 5.

d'information devront donc respecter les règles applicables aux SIIV, c'est pourquoi la suite de notre propos se concentrera sur ce régime et non sur celui de la directive NIS.

161. **Une innovation bridée ?** L'ANSSI est chargée de la mise en œuvre de la règlementation applicable aux SIIV et a adopté une doctrine elle-même très restrictive. Ces règles, particulièrement lourdes, se justifient au regard des risques que peuvent faire peser ces systèmes sur la vie de la Nation ou sur la population (accident nucléaire, coupures de courant…). C'est d'autant plus vrai dans le secteur de l'électricité puisque nous avons déjà pu assister à des attaques informatiques d'ampleur sur des infrastructures critiques telles qu'une centrale nucléaire en Iran[381] ou encore à des erreurs logicielles ayant conduit à des pannes de courant à grande échelle au Japon dans les années 2000[382]. Les risques informatiques sont donc réels dans les infrastructures critiques de l'énergie et pourraient être aggravés si des systèmes d'IA opaques et autonomes y étaient intégrés. Il n'est ainsi pas surprenant que nos rencontres avec des experts[383] aient révélé que l'utilisation de l'IA était bridée par la règlementation sectorielle en matière de sécurité.

B/ L'incompatibilité de l'IA avec les règles de sécurité applicables aux SIIV

162. **Plan.** L'étude des règles de sécurité applicables aux SIIV du secteur de l'approvisionnement en énergie électrique, contenues dans l'arrêté sectoriel du 11 août 2016[384] ou dans les différents référentiels publiés par l'ANSSI, permet de mettre en lumière plusieurs limites à l'utilisation de systèmes d'IA dans ce contexte. Ces règles imposent notamment d'identifier et de notifier à l'ANSSI la liste de tous les SIIV opérés par l'OIV, de rédiger une

[381] T.Y. Ebrahim, « National Cybersecurity Innovation », *West Virginia Law Review*, 2020, vol. 123, n° 2, pp. 483-546, spec. 485-486.
[382] M. Williams, « Computer problems hit three nuclear plants in Japan », *CNN (blog)*, 3 janvier 2000, disponible en ligne : <https://edition.cnn.com/2000/TECH/computing/01/03/japan.nukes.y2k.idg/index.html>, consulté le 9 septembre 2022.
[383] Voir Supra, 63.
[384] *Arrêté du 11 août 2016 fixant les règles de sécurité et les modalités de déclaration des systèmes d'information d'importance vitale et des incidents de sécurité relatives au sous-secteur d'activités d'importance vitale « Approvisionnement en énergie électrique » et pris en application des articles R. 1332-41-1, R. 1332-41-2 et R. 1332-41-10 du code de la défense*, publié au JORF n°0197 du 25 août 2016, texte n° 4.

politique de sécurité des systèmes d'information, d'auditer les SIIV et de mettre en place de nombreuses mesures pour garantir le maintien en conditions de sécurité, de détecter et traiter tous les incidents ou encore de garantir le cloisonnement des systèmes ou sous-systèmes[385]. Dès lors, le fait que l'IA soit difficilement explicable et auditable peut poser des difficultés (**1**). L'absence de standards institutionnels relatifs à la sécurité des systèmes d'IA rendrait difficile tout contrôle par l'ANSSI et place les OIV dans une situation d'insécurité juridique (**2**). Les règles relatives à l'homologation *ex-ante* et au maintien en conditions de sécurité excluent, de fait, le recours à des systèmes d'IA auto-apprenants (**3**), bien qu'ils ne soient pas encore fréquemment utilisés dans la pratique. Enfin, le nécessaire cloisonnement des systèmes empêche l'utilisation d'un système d'IA qui fonctionnerait grâce à des données collectées par des capteurs ou générées par d'autres systèmes opérés par l'OIV (**4**).

1. L'auditabilité imparfaite des systèmes d'IA

163. **Une double obligation d'audit dans la règlementation des SIIV.** Deux dispositions du régime applicable aux SIIV imposent qu'il soit possible de les auditer.

164. **La capacité d'audit requise au titre du Code de la défense.** La première est prévue à l'article L1332-6-3 du Code de la Défense qui dispose que les SIIV doivent pouvoir être audités à tout moment par l'ANSSI afin de contrôler la sécurité et la conformité aux règles fixées par les arrêtés sectoriels. Pour rappel, ces règles s'appliquent aux *« systèmes pour lesquels l'atteinte à la sécurité ou au fonctionnement risquerait de diminuer d'une façon importante le potentiel de guerre ou économique, la sécurité ou la capacité de survie de la Nation ou pourrait présenter un danger grave pour la population »*[386], ce qui couvre une partie des systèmes logiciels utilisés dans les fonctions critiques des centrales de production d'électricité ou des réseaux de transport et de distribution. Des systèmes d'IA utilisés dans de tels systèmes devraient donc pouvoir être audités par l'ANSSI à tout moment.

[385] *Ibid.*, annexe I.
[386] Code de la Défense, article L1332-6-1.

165. **La capacité d'audit requise par la règlementation spécifique au secteur de l'électricité.** La deuxième disposition imposant que les SIIV soient auditables est issue des règles de sécurité sectorielles prévues par l'arrêté du 11 août 2016 pour ce qui concerne le secteur d'activité « approvisionnement en énergie électrique »[387]. Sa deuxième règle impose en effet à l'OIV de procéder à l'homologation de sécurité de chaque SIIV en amont de leur utilisation[388]. L'homologation d'un système est une décision formelle prise par l'opérateur qui atteste que les risques pesant sur la sécurité de ce système ont été identifiés et que les mesures nécessaires pour le protéger sont mises en œuvre[389]. L'arrêté précise également que, dans le cadre de l'homologation, un audit de la sécurité du SIIV doit être réalisé afin d'évaluer le niveau de sécurité du SIIV au regard des menaces et des vulnérabilités connues. La procédure consiste en réalité à un triple audit puisque la règle impose la réalisation *« d'un audit d'architecture, d'un audit de configuration et d'un audit organisationnel et physique »*[390]. Sans avoir besoin d'entrer dans le détail du contenu de chaque audit, on comprend aisément que si un système d'IA venait à être utilisé dans ce contexte, il devrait être parfaitement auditable. Pourtant, certains modèles d'IA fonctionnent encore en « boîte noire » et il reste difficile de comprendre précisément leur fonctionnement intrinsèque. C'est notamment le cas de certains algorithmes d'apprentissage automatique fondés sur des « réseaux de neurones profonds ». Dès lors, il est probable que le défaut d'explicabilité de certains systèmes d'IA vienne complexifier leur homologation. Si les OIV exploitant les SIIV sont *a priori* les seuls à définir les méthodes employées pour homologuer leurs systèmes et à réaliser les audits mentionnés, il n'en reste pas moins que l'ANSSI a le pouvoir, d'une part, de déclencher des audits à son initiative afin d'évaluer la sécurité des systèmes et, d'autre part, de faire part à l'OIV de ses observations sur les décisions d'homologation, pouvant aller jusqu'à les invalider. Le non-respect des règles de sécurité constaté par l'ANSSI peut conduire à une amende pénale allant jusqu'à 150 000€ pour les dirigeants des OIV concernés. *In fine*, c'est bien l'autorité nationale qui décide si les

[387] *Arrêté du 11 août 2016 fixant les règles de sécurité et les modalités de déclaration des systèmes d'information d'importance vitale et des incidents de sécurité relatives au sous-secteur d'activités d'importance vitale « Approvisionnement en énergie électrique » et pris en application des articles R. 1332-41-1, R. 1332-41-2 et R. 1332-41-10 du code de la défense*, publié au JORF n°0197 du 25 août 2016, texte n° 4.
[388] *Ibid.*, Annexe I, 2.
[389] *Ibid.*
[390] *Ibid.*

techniques employées et les mesures de protection mises en œuvre permettent d'atteindre un niveau suffisant de sécurité[391]. Or, il est peu probable qu'elle parvienne à cette conclusion si l'OIV ou elle-même n'est pas en mesure de comprendre le fonctionnement du système audité. C'est pourquoi l'utilisation de l'IA dans les systèmes critiques du secteur de l'électricité est encore si peu répandue. C'est aussi la raison pour laquelle de nombreux programmes de recherche ont été lancés afin de développer des systèmes d'IA suffisamment explicables pour pouvoir être audités et ainsi utilisés dans des environnements critiques[392].

2. L'absence de standards permettant de contrôler la sécurité des systèmes d'IA

166. **L'utilité de standards adaptés aux spécificités de l'IA.** L'établissement de standards harmonisés contenant des règles de sécurité spécifiques aux systèmes d'IA permettrait de les homologuer et de se mettre en conformité plus facilement avec les règles générales applicables à tous les SIIV. En effet, certaines dispositions de l'arrêté de 2016 sont difficilement compatibles avec les caractéristiques des systèmes d'apprentissage automatique. Par exemple, la règle n°8 impose la mise en place d'une procédure de traitement des incidents de sécurité et la conservation des informations relatives aux incidents telles que les relevés d'informations et les analyses des incidents[393]. *Quid* si l'opacité du système d'IA utilisé rend impossible l'identification de l'origine de l'incident ?

167. **L'incompatibilité des méthodologies d'analyse des risques prescrites par la règlementation.** Une autre incompatibilité vient du fait que les règles de sécurité pour les SIIV se concentrent sur l'analyse des risques de vulnérabilités, la détection des incidents et la mise en place de procédures pour y répondre en temps voulu. Les menaces prises en référence sont

[391] Voir notamm. le guide exhaustif de l'ANSSI encadrant la procédure d'homologation (pourtant censée être à la main des OIV) : ANSSI, « L'homologation de sécurité », *Guide pratique*, Version 1.0, août 2014, n°20140821-1128.
[392] Voir par exemple le laboratoire « SINCLAIR » établit en 2020 par EDF, TotalEnergies et Thalès : https://sinclair-lab.com/.
[393] Arrêté du 11 août 2016, *op. cit.*, Annexe I, 8.

issues de la méthodologie d'analyse des risques « EBIOS »[394], pensée pour des logiciels déterministes classiques. Or, les systèmes d'IA peuvent présenter des vulnérabilités spécifiques. Notamment, ils peuvent avoir été conçus à partir de données non représentatives, incomplètes ou de mauvaise qualité. Dans ce cas, les éventuels défauts ne viennent pas d'une manipulation par un tiers ou d'un traitement inapproprié des données mais simplement des choix réalisés lors de la conception. Ces vulnérabilités spécifiques devraient être prises en compte dans les règles de sécurité applicables aux systèmes d'IA et pourraient utilement faire l'objet d'un référentiel technique dédié.

168. **Transition.** En plus des difficultés évoquées, l'application de certaines règles de sécurité conduit à exclure l'utilisation de plusieurs techniques d'IA dans les systèmes critiques.

3. L'exclusion de fait des systèmes d'IA auto-apprenants

169. **Le nécessaire maintien en conditions de sécurité dans les infrastructures critiques.** Les systèmes d'IA capables d'évoluer par apprentissage au cours de leur utilisation sont aujourd'hui encore rares, notamment parce qu'il est difficile d'anticiper leur évolution. Si la problématique est commune à tous les secteurs d'activité, elle est encore plus importante dans les environnements critiques pour la sécurité. Ces derniers imposent une parfaite maîtrise dans le temps des technologies employées. Il n'est donc pas surprenant que le corpus de règles de sécurité applicables aux SIIV du secteur de l'électricité contienne une disposition imposant la mise en œuvre d'une procédure de *« maintien en conditions de sécurité »*[395]. Cette procédure doit prévoir des mesures pour prendre en compte l'évolution des menaces et vulnérabilités, en insistant sur la définition d'une politique d'installation de toute nouvelle version ou mesure correctrice[396]. Par ailleurs, l'homologation des SIIV préalable à leur utilisation impose également la définition de mesures visant à garantir le maintien en conditions de sécurité. Le

[394] ANSSI, « L'homologation de sécurité », *Guide pratique*, Version 1.0, août 2014, n°20140821-1128, annexe IV contenant la liste des menaces de référence, issues de la méthode d'analyse des risques EBIOS.
[395] Arrêté du 11 août 2016, *op. cit.*, Annexe I, 4.
[396] *Ibid.*

guide dédié de l'ANSSI précise à cet égard, d'une part, que l'homologation doit être revue annuellement pour contrôler que les règles de sécurité sont toujours respectées et, d'autre part, qu'une veille technologique doit être mise en place pour identifier les vulnérabilités qui apparaîtraient au cours de l'utilisation du système[397].

170. **L'impossible maintien en conditions de sécurité de systèmes apprenants.** Il est regrettable que les règles de l'arrêté se focalisent exclusivement sur la politique de mise à jour et le suivi des vulnérabilités. En effet, si des systèmes d'IA capables d'apprendre au cours de leur utilisation venaient à être utilisés, il conviendrait d'assurer un suivi continu des évolutions qu'ils subissent. Un système dont l'apprentissage se ferait tout au long de son cycle de vie pourrait fonctionner parfaitement pendant une longue période puis dysfonctionner bien plus tard, lorsque ses conditions d'exploitation évolueraient ou lorsqu'il sera victime du phénomène de *« surentraînement »* dégradant ses performances[398]. Dans ces cas, le dysfonctionnement n'est identifié qu'*a posteriori*, ce qui n'est pas acceptable dans des environnements critiques comme la production ou le transport d'électricité. Il est donc peu probable que des systèmes d'IA apprenants puissent satisfaire les conditions de maintien en conditions de sécurité prévues par la règlementation sectorielle. Cela se traduit en pratique par une priorité donnée aux modèles dont l'apprentissage se cantonne à la phase de conception et qui n'évoluent plus une fois mis en service, sauf mise à jour ou réentraînement.

171. **Transition.** Outre les systèmes auto-apprenants, ceux dont le fonctionnement requiert leur interconnexion avec d'autres systèmes d'information apparaissent également difficilement compatibles avec la règlementation.

[397] ANSSI, « L'homologation de sécurité », *op. cit.*, pp. 46–47.
[398] I. Bilbao, J. Bilbao, « Overfitting problem and the over-training in the era of data: Particularly for Artificial Neural Networks », *2017 eighth international conference on intelligent computing and information systems (ICICIS)*, IEEE, 2017.

4. Le difficile cloisonnement des systèmes d'IA

172. **Le cloisonnement obligatoire des SIIV.** La seizième règle de sécurité contenue dans l'arrêté sectoriel du 11 août 2016 impose le cloisonnement des SIIV afin de limiter la propagation des attaques informatiques au sein des autres systèmes de l'OIV qui les exploitent[399]. Concrètement, l'OIV doit cloisonner chaque SIIV *« physiquement ou logiquement vis-à-vis des autres systèmes de l'opérateur ou des systèmes tiers »* ainsi que chaque sous-système composant le SIIV s'il en existe[400]. Cela signifie que la règle s'appliquerait tant à un système d'IA utilisé comme SIIV qu'à un système d'IA utilisé comme composant d'un SIIV. Seules les interconnexions strictement nécessaires à la sécurité d'un SIIV sont tolérées. Le cloisonnement, non défini dans les textes, est présenté dans le guide relatif au cloisonnement système de l'ANSSI en référence au principe du moindre privilège : le composant d'un système ne doit avoir la possibilité de mener à bien que les actions dont l'utilité fonctionnelle est avérée[401]. Autrement dit, le cloisonnement consiste à la mise en place de mesures techniques pour que les systèmes ne puissent exécuter que la tâche pour laquelle ils ont été conçus et n'interagissent qu'avec les composants strictement nécessaires à leur fonctionnement. De cette manière, si un dysfonctionnement intervient, seule une tâche est affectée et non l'ensemble du système.

173. **L'interconnexion nécessaire au fonctionnement des systèmes d'IA.** L'une des caractéristiques des systèmes d'IA est leur dépendance à la donnée. En effet, ils consistent pour la majorité en des systèmes de traitement automatisé de données, eux-mêmes conçus à partir d'un apprentissage sur des données d'exemple. Dès lors, leur interconnexion avec d'autres composants fait partie de leur nature. Evidemment, il est possible d'appliquer le cloisonnement à l'IA, en limitant le nombre d'interconnexion au strict nécessaire et en mettant en place des mesures minimisant les effets de bord en cas de dysfonctionnement. Toutefois, la contrainte du cloisonnement risque de limiter le nombre d'applications de l'IA dans les systèmes critiques, qui seront toujours plus interconnectés que des logiciels classiques. À défaut de les

[399] Arrêté du 11 août 2016, *op. cit.*, Annexe I, 16.
[400] *Ibid.*, alinéa 2.
[401] ANSSI, « Recommandations pour la mise en place de cloisonnement système », *Guide pratique*, 14 décembre 2017, ANSSI-PG-040.

interconnecter avec d'autres sources de données, les entrées – ainsi que les vérifications nécessaires – devront se faire manuellement, ce qui génère un aléa humain et un délai supplémentaire. L'application du cloisonnement à l'IA pourrait finalement avoir un effet contre-productif et priver les systèmes d'apprentissage automatique de leurs avantages initiaux (automatisme, rapidité de traitement, capacité à traiter de grandes quantités de données en temps quasi-réel).

174. **Conclusion du §1 relatif aux contraintes liées au régime de la sécurité informatique des infrastructures d'importance vitale.** Le régime juridique applicable aux SIIV du secteur de l'électricité génère ainsi plusieurs contraintes non négligeables aux cas d'usage les plus critiques envisagés dans la présente thèse. Au vu de la règlementation et de l'état actuel de la recherche scientifique sur la compréhension des systèmes d'IA, il n'est donc pas surprenant que l'apprentissage automatique soit très peu – voire pas du tout – employé dans les systèmes critiques de l'énergie tels que les réseaux de transport et de distribution, ainsi que les centrales de production.

175. **Transition.** Concernant les centrales de production, il convient de noter cependant qu'elles présentent des niveaux de risques différents suivant l'origine de l'énergie produite. La production électronucléaire, par exemple, est reconnue plus dangereuse que les productions photovoltaïque ou éolienne. C'est la raison pour laquelle les centrales nucléaires répondent d'un régime juridique encore plus strict, générant des contraintes additionnelles à l'utilisation de l'IA dans ce contexte.

§2 : Des limites liées au régime juridique de la sécurité nucléaire

176. **Plan.** L'IA n'est aujourd'hui utilisée qu'en dehors du système informatique des centrales nucléaires de production d'électricité, bien qu'elle puisse être plus performante que certains logiciels déterministes. Plusieurs raisons ont conduit à cette situation. Notamment, la relative nouveauté des techniques d'apprentissage automatique par rapport aux systèmes experts utilisés depuis des décennies conduit à des appréhensions dans l'industrie nucléaire, qui préfère conserver une technologie connue et éprouvée par le temps, fusse-t-elle moins performante. Ce constat conduit à un certain immobilisme technologique dans les centrales nucléaires, accentué par une règlementation exhaustive qui ne laisse que peu de place à l'innovation. Dès lors, si

l'on se réfère aux textes règlementaires applicables à l'utilisation de l'IA dans les installations nucléaires (**A**), il semble qu'un tel usage soit difficilement compatible avec plusieurs règles fondamentales en matière de sûreté nucléaire (**B**).

A/ L'encadrement juridique de l'utilisation de l'IA dans des installations nucléaires

177. **Plan.** Le recours à des systèmes d'IA dans un environnement critique comme celui d'une centrale nucléaire se fait dans un contexte règlementaire très strict. Un tel cas d'usage devrait se conformer au régime juridique de la sécurité nucléaire (**1**) comprenant un ensemble plus spécifique appelé la sûreté nucléaire (**2**). Cette dernière relève d'une règlementation *ad hoc* générant des contraintes additionnelles à l'utilisation d'IA en centrale. Avant d'étudier les limites de leur application à l'IA, ces deux notions doivent être définies et leur régime précisé.

1. L'applicabilité des règles de sécurité nucléaire aux systèmes d'IA

178. **Les sources de la sécurité nucléaire.** Les articles L591-1 à L597-46 du Code de l'environnement prévoient des dispositions générales encadrant la sécurité nucléaire et les Installations Nucléaires de Base (INB)[402] mais la majeure partie des règles relèvent du pouvoir règlementaire[403]. De plus, plusieurs textes publiés par des organisations internationales telles que l'Agence Internationale de l'Energie Atomique (AIEA) font office de standards de référence et ont une grande influence sur les règlementations nationales[404]. Les normes étudiées

[402] L'article L593-2 du Code de l'environnement liste les installations qualifiées d'INB : les réacteurs nucléaires ; les installations de préparation, d'enrichissement, de fabrication, de traitement ou d'entreposage de combustibles nucléaires ou de traitement, d'entreposage ou de stockage de déchets radioactifs ; les installations contenant des substances radioactives ou fissiles ; les accélérateurs de particules ; ainsi que les centres de stockage en couche géologique de déchets radioactifs. Le présent paragraphe se concentrera sur l'utilisation de systèmes d'IA dans les centrales nucléaires de production d'électricité, entrant dans la première des catégories précitées.
[403] Code de l'environnement, article L591-2.
[404] Voir par exemple l'ensemble des standards techniques publiés par l'AIEA en matière de sûreté nucléaire : AIEA, « Specific Safety Guides », *Site de l'AIEA (blog)*, disponible en ligne : <https://www.iaea.org/publications/search/type/safety-standards-series>, consulté le 16 septembre 2022.

dans le cadre du présent Paragraphe seront donc diverses tant par leur origine (législative, règlementaire, privée) que leur contenu (grands principes, règles juridiques, standards techniques non contraignants).

179. **La définition de la sécurité nucléaire.** En droit français, la sécurité nucléaire est définie depuis la loi relative à la transparence et à la sécurité en matière nucléaire du 13 juin 2006 comme comprenant *« la sûreté nucléaire, la radioprotection, la prévention et la lutte contre les actes de malveillance, ainsi que les actions de sécurité civile en cas d'accident »*[405]. Sa définition, inchangée, est aujourd'hui codifiée à l'article L591-1 du Code de l'environnement.

180. **L'encadrement de l'utilisation de systèmes d'information au titre de la sécurité nucléaire.** Dans le cadre de notre étude, nous nous intéressons exclusivement aux règles de sécurité pouvant encadrer l'usage de systèmes informatiques. Dès lors, les règles relatives à la radioprotection (protection des personnes contre les rayonnements ionisants) et aux actions de sécurité civile en cas d'accident ne retiendront pas notre attention. La prévention et la lutte contre les actes de malveillance correspondent à la mise en œuvre de mesures visant à assurer la protection physique et informatique des INB. Or, la sécurité informatique des INB est régie par les règles incombant aux OIV puisque le nucléaire fait partie des secteurs d'activités considérés comme d'importance vitale pour les intérêts de la Nation[406]. Le contenu des règles en matière de sécurité applicables aux systèmes d'IA au titre de la sécurité nucléaire sont donc les mêmes que celles applicables aux SIIV, décrites dans le précédent Paragraphe. Par conséquent, elles ne seront pas étudiées dans la suite du propos. Toutefois, la dernière composante de la sécurité nucléaire, la sûreté nucléaire, présente des spécificités et peut également générer des contraintes à l'utilisation de l'IA.

---

[405] *Loi n° 2006-686 du 13 juin 2006 relative à la transparence et à la sécurité en matière nucléaire*, publiée au JORF n°136 du 14 juin 2006, texte n° 2, article 1er.

[406] Comme pour le secteur de l'approvisionnement électrique, un arrêté sectoriel vient encadrer les SIIV utilisés par les OIV du domaine nucléaire : *Arrêté du 10 mars 2017 fixant les règles de sécurité et les modalités de déclaration des systèmes d'information d'importance vitale et des incidents de sécurité relatives au sous-secteur d'activités d'importance vitale « Nucléaire » et pris en application des articles R. 1332-41-1, R. 1332-41-2 et R. 1332-41-10 du code de la défense*, publié au JORF n°0065 du 17 mars 2017, texte n° 2

2. L'applicabilité des règles de sûreté nucléaire aux systèmes d'IA

181. **Définition de la sûreté nucléaire.** Le deuxième alinéa de l'article L591-1 du Code de l'environnement définit la sûreté nucléaire comme *« l'ensemble des dispositions techniques et des mesures d'organisation relatives à la conception, à la construction, au fonctionnement, à l'arrêt et au démantèlement des installations nucléaires de base, ainsi qu'au transport des substances radioactives, prises en vue de prévenir les accidents ou d'en limiter les effets »*[407]. En d'autres termes, la sûreté nucléaire vise à garantir le bon fonctionnement d'une INB dans des conditions d'exploitation normales. Les conditions « normales » sont ici entendues largement : des catastrophes naturelles telles que des séismes ou des tsunamis sont par exemple des situations auxquelles une installation nucléaire sûre doit pouvoir résister. À l'inverse, la sécurité nucléaire englobe les mesures de protection contre des attaques extérieures (informatiques ou physiques comme des attentats ou des détournements d'aéronefs). La sûreté vise donc à prévenir les accidents d'exploitation, là où la sécurité consiste en la protection contre des actes malveillants[408]. Evidemment, les deux notions sont intimement liées puisque la sûreté est une des composantes de la sécurité nucléaire et que certaines mesures techniques (telles que le cloisonnement des systèmes informatiques) peuvent avoir un intérêt commun. Il n'en reste pas moins qu'elles répondent de régimes distincts.

182. **Sources du régime de la sûreté nucléaire.** En France[409], la sûreté nucléaire est encadrée par le Code de l'environnement, plusieurs décisions règlementaires d'application[410] et est contrôlée par une autorité administrative indépendante : l'Autorité de Sûreté Nucléaire (ASN)[411]. Cette dernière a le pouvoir de prendre des décisions règlementaires à caractère

[407] Code de l'environnement, article L591-1, alinéa 2.
[408] Pour une étude des distinctions entre sûreté et sécurité nucléaires, voir : IRSN, *Approche comparative entre sûreté et sécurité nucléaires*, Rapport IRSN 2009/117, 21 avril 2009, 26 p.
[409] Les régimes nationaux relatifs à la sûreté nucléaire sont en grande partie inspirés de normes ou conventions internationales telles que la Convention sur la sûreté nucléaire adoptée en 1994 (*Convention sur la sûreté nucléaire*, adoptée le 17 juin 1994, INFCIRC/449) ou le Traité Euratom (*Version consolidée du Traité instituant la communauté européenne de l'énergie atomique (Euratom)*, 2010/C 84/01, publiée au JOUE n°C84, 30 mars 2010) en droit européen. La présente thèse se concentrera sur le droit français.
[410] Notamment l'arrêté dit « INB » : *Arrêté du 7 février 2012 fixant les règles générales relatives aux installations nucléaires de base*, publié au JORF n°0033 du 8 février 2012, texte n°12.
[411] Code de l'environnement, article L592-1.

technique pour compléter les modalités d'application des décrets et arrêtés pris notamment dans le domaine des INB. La doctrine de l'ASN et les actes règlementaires qu'elle prend relèvent donc d'une grande importance dans la règlementation de la sûreté nucléaire.

183. **Contenu du régime de la sûreté nucléaire.** Pour gérer les risques liés à l'exploitation des INB, le Code de l'environnement prévoit que leur création, leur mise en service et leur exploitation sont soumises à l'autorisation de l'ASN[412]. Cette autorisation ne peut être délivrée que si l'exploitant démontre que les dispositions techniques ou d'organisation prises ou envisagées aux stades de la conception, de la construction et de l'exploitation ainsi que les principes généraux proposés pour le démantèlement sont de nature à prévenir ou à limiter de manière suffisante les risques ou inconvénients que l'installation présente[413]. Pour espérer obtenir l'autorisation nécessaire à la création d'une INB, l'exploitation doit donc démontrer la sûreté de son installation : c'est la « démonstration de sûreté »[414]. Cette démonstration se fait essentiellement sur la base de rapports ou de calculs scientifiques (réalisés manuellement ou en utilisant des logiciels qualifiés selon une procédure prévue par l'ASN). Elle vise notamment à prouver à l'ASN le respect du principe de base de la sûreté nucléaire : la défense en profondeur. Ce principe consiste, pour l'exploitant, en la mise en œuvre de plusieurs niveaux de protection comprenant des barrières successives et indépendantes visant à empêcher la dispersion de substances radioactives dans l'environnement[415]. Les différentes barrières peuvent être physiques (confinement matériel des équipements) ou logiciels (systèmes de sûreté automatiques).

184. **Un régime contraignant pour le recours à l'IA dans les installations nucléaires.** Dans un tel contexte règlementaire, une attention particulière doit être portée aux systèmes d'information utilisés dans les INB. En effet, l'exploitant doit veiller à ce que ses systèmes logiciels n'affaiblissent pas la sûreté de fonctionnement globale de l'installation. Les logiciels doivent eux-mêmes présenter le plus haut niveau de sûreté et l'exploitant doit être en capacité de le démontrer à l'ASN, sans quoi il ne pourrait pas les utiliser. De tels systèmes sont rares,

[412] Code de l'environnement, article L593-7.
[413] Code de l'environnement, article L593-7, I.
[414] M. Moliner-Dubost, « Nucléaire », *JurisClasseur Administratif*, version mise à jour le 14 avril 2021, Fasc. 378, spec. 191–206.
[415] *Ibid.*, 186–188.

c'est pourquoi l'industrie nucléaire, où la culture de la sûreté est forte, n'utilise finalement que peu de systèmes entièrement informatisés. En effet, l'utilisation de l'IA dans cet environnement s'avère fortement contrainte par les règles de la sûreté nucléaire qui ne semblent pas adaptée aux spécificités de cette technologie.

B/ L'incompatibilité de l'IA avec les règles de sûreté nucléaire

185. **Plan.** L'application du principe de défense en profondeur, au cœur de la démarche de sûreté nucléaire, semble condamner les systèmes d'IA à être exclus du système d'information des centrales nucléaires (**1**). De plus, les modalités de la démonstration de sûreté, et notamment les conditions de qualification des outils logiciels utilisés dans ce cadre, ne semblent pas plus adaptés (**2**).

1. L'exclusion du recours à l'IA en application du principe de défense en profondeur

186. **L'origine du principe de la défense en profondeur.** La défense en profondeur est un des principes fondateurs de la sûreté nucléaire et est inscrit dans la règlementation française depuis 2012[416]. Il constitue une norme internationale depuis que l'AIEA en a fait le pilier de sa doctrine dans les années 90[417]. Depuis, plusieurs standards ont été établis afin de préciser les modalités de la mise en œuvre la défense en profondeur dans la conception, l'exploitation et le démantèlement des centrales nucléaires[418]. Elle s'applique à tous les éléments composant une INB et dont le fonctionnement (ou dysfonctionnement) peut avoir des impacts sur la sûreté de l'installation, y compris les systèmes d'information. D'autres normes sont venues préciser comment ce principe pouvait être adapté à la conception et à l'utilisation de systèmes

[416] *Arrêté du 7 février 2012 fixant les règles générales relatives aux installations nucléaires de base*, publié au JORF n°0033 du 8 février 2012, texte n° 12, article 3.1., I : *« L'exploitant applique le principe de défense en profondeur, consistant en la mise en œuvre de niveaux de défense successifs et suffisamment indépendants […] »*.
[417] AIEA, *La sûreté des installations nucléaires*, 1993, collection Sécurité n°110.
[418] GROUPE CONSULTATIF INTERNATIONAL POUR LA SÛRETÉ NUCLÉAIRE, *La défense en profondeur en sûreté nucléaire*, 1997, INSAG-10 ; GROUPE CONSULTATIF INTERNATIONAL POUR LA SÛRETÉ NUCLÉAIRE, *Basic Safety Principles for Nuclear Power Plants*, 1999, 75-INSAG-3 Rev. 1, INSAG-12.

d'information, notamment en France avec un guide dédié dès 2004, avant même la création de l'ANSSI[419].

187. **La défense en profondeur appliquée aux systèmes critiques.** Appliquer le principe de la défense en profondeur aux systèmes d'information utilisés dans des environnements critiques nécessite la mise en place de différents niveaux de protection, y compris au niveau logiciel, afin d'éviter la propagation d'éventuels dysfonctionnements. La défense en profondeur s'articule autour d'un triptyque constitué de *« barrières »*, de *« lignes de défense »* et de différents *« niveaux de protection »*[420]. Les deux premières correspondent essentiellement à des moyens physiques de protection des installations (enceinte de confinement…) ainsi qu'à des mesures structurelles (systèmes de détection, ergonomie des interfaces homme/machine…) et organisationnelles (consignes de sécurité, procédures…). Les *« niveaux de protection »* en revanche consistent en cinq niveaux : la prévention des incidents (premier niveau), la détection des incidents et le rétablissement d'une situation de fonctionnement normal s'ils conduisent à un accident (deuxième niveau), la maîtrise des accidents n'ayant pu être évités ou la limitation de leur gravité (troisième niveau), la gestion des accidents graves consécutifs à l'échec des trois premiers niveaux de défense pour en limiter les conséquences sur les personnes ou l'environnement (quatrième niveau), la gestion de crise par les pouvoirs publics en vue d'atténuer les conséquences radiologiques des rejets accidentels n'ayant pu être évités (cinquième niveau)[421]. L'ensemble des *« barrières »* et *« lignes de défense »* mises en place dans chacun de ces *« niveaux de protection »* sont ce qui constitue la sûreté d'une installation nucléaire. Des normes techniques précisent les mesures techniques et organisationnelles pouvant constituer des barrières et lignes de défense pour chaque niveau de protection[422].

---

[419] SECRETARIAT GENERAL DE LA DEFENSE NATIONALE, *La défense en profondeur appliquée aux systèmes d'information*, Mémento, 19 juillet 2004, version 1.1, disponible en ligne : <https://www.ssi.gouv.fr/uploads/IMG/pdf/mementodep-v1-1.pdf>, consulté le 16 septembre 2022.

[420] M. Moliner-Dubost, « Nucléaire », *JurisClasseur Administratif*, version mise à jour le 14 avril 2021, Fasc. 378, spec. 189.

[421] IRSN, « Démarche générale de prévention des accidents – La défense en profondeur », *Site de l'IRSN (blog)*, disponible en ligne : <https://www.irsn.fr/FR/connaissances/Installations_nucleaires/La_surete_Nucleaire/risque-nucleaire/demarche-prevention/Pages/1-defense-profondeur.aspx#.YySu7N8682w>, consulté le 16 septembre 2022.

[422] ASN, *Conception des réacteurs à eau sous pression*, Guide ASN n°22, 19 juillet 2017.

188. **Les conséquences de l'application du principe de la défense en profondeur sur l'utilisation de logiciels dans les centrales nucléaires.** En raison des principes évoqués précédemment, le recours à des systèmes logiciels dans les fonctions de sûreté des centrales nucléaires est rare. Il se limite à des logiciels déterministes très simples qui n'exécutent qu'une tâche bien précise et dont le fonctionnement ne risque pas d'évoluer dans le temps, conformément aux guides publiés par l'AIEA pour l'emploi de systèmes informatisés dans les fonctions de contrôle-commande[423] ou autres fonctions critiques[424]. En d'autres termes, dans un tel contexte, seuls peuvent être utilisés des systèmes dont la compréhension par l'homme est parfaite et le fonctionnement uniforme dans le temps. À défaut, l'ASN ne validera pas la démonstration de sûreté. L'utilisation de systèmes d'IA, par nature complexes et parfois opaques, semble ainsi incompatible avec la culture de la sûreté telle qu'elle existe dans l'industrie nucléaire.

189. **Transition.** L'inadaptation de la règlementation en matière de sûreté nucléaire semble provenir plus précisément des modalités de la démonstration de sûreté que doit présenter l'exploitant d'une INB à l'ASN.

2. L'exclusion du recours à l'IA à défaut de méthode de qualification adaptée

190. **L'absence de consensus sur la méthode pour prouver la fiabilité d'un système d'IA.** À l'heure actuelle, il n'existe aucun standard technique, d'origine privée ou institutionnelle, sur la sécurité des systèmes d'IA. En d'autres termes, il n'existe pas encore de consensus sur la façon de prouver qu'un tel système est sûr. En l'absence d'une norme harmonisée, à quoi une autorité de régulation comme l'ASN pourrait-elle se référer pour évaluer le niveau de sécurité d'un système fondé sur l'apprentissage automatique ? Dans ce contexte, les systèmes d'IA semblent condamnés à n'être jamais utilisés pour des fonctions dont la criticité nécessite

[423] AIEA, *Systèmes d'instrumentation et de contrôle-commande importants pour la sûreté des centrales nucléaires*, 2005, collection Normes de sûreté nº NS-G-1.3.
[424] AIEA, *Logiciels destinés aux systèmes programmés important s pour la sûreté des centrales nucléaires*, 2004, collection Normes de sûreté n° NS-G-1.1.

l'homologation, la certification ou la démonstration de sûreté. Heureusement, de nombreux travaux de normalisation sont en cours, à l'initiative d'acteurs privés[425], d'organisations de normalisation[426] ou d'autorités publiques[427], laissant penser que la situation pourrait évoluer dans les années à venir. En attendant, il est très difficile de démontrer la sûreté des systèmes d'IA en utilisant les méthodes de certification conçues pour les logiciels déterministes traditionnels. L'exemple de la méthode de qualification des logiciels utilisés dans le cadre de la démonstration de sûreté nucléaire auprès de l'ASN illustre ce constat.

191. **L'inadaptation de la méthode de qualification des logiciels utilisés dans le cadre de la démonstration de sûreté nucléaire.** L'acceptabilité de l'utilisation de logiciels dans le système de sûreté d'une installation critique comme une centrale nucléaire dépend du degré de confiance que l'on peut avoir dans leur fonctionnement et leur résilience face à une possible défaillance. Pour garantir un haut niveau de confiance, les logiciels de sûreté nucléaire étaient initialement évalués par l'Institut de Protection et de Sûreté Nucléaire (IPSN)[428], appui technique de l'ASN, selon deux critères. D'une part, la phase de conception du logiciel devait avoir respecté les standards pratiqués dans l'industrie. Pour ce critère, l'IPSN contrôlait la documentation relative à la conception du système et fournie par le développeur pour vérifier sa conformité avec l'état de l'art. D'autre part, le respect des bonnes pratiques en matière de conception était contrôlé à travers un examen détaillé de l'architecture du système, la relecture exhaustive des lignes de code source et de tests sur le comportement du logiciel[429]. Cet examen en deux étapes a finalement été remplacé par la méthode Vérification, Validation et

[425] LABORATOIRE NATIONAL DE METROLOGIE ET D'ESSAIS (LNE), « Certification de processus pour l'IA », *Site du LNE (blog)*, disponible en ligne : <https://www.lne.fr/fr/service/certification/certification-processus-ia>, consulté le 18 septembre 2022.

[426] Voir l'avancement de l'ISO sur la normalisation de l'IA : ISO, ISO/IEC JTC 1/SC 42 Artificial intelligence, disponible en ligne : <https://www.iso.org/committee/6794475.html>, consulté le 18 septembre 2022 ; ETSI, *Securing Artificial Intelligence (SAI) : Problem Statement*, rapport, décembre 2020, GR SAI 004, V1.1.1.

[427] ANSSI, *Rapport d'activité 2018*, p. 50, disponible en ligne : <https://www.ssi.gouv.fr/uploads/2019/04/anssi_rapport_annuel_2018.pdf>, consulté le 18 septembre 2022 ; voir aussi les outils référencés par la CNIL sur le sujet : CNIL, « Conformité des systèmes d'IA : les autres guides, outils et bonnes pratiques », *Site de la CNIL (blog)*, 5 avril 2022, disponible en ligne : <https://www.cnil.fr/fr/intelligence-artificielle/guide/conformite-des-systemes-dia-les-autres-guides-outils-et-bonnes-pratiques>, consulté le 18 septembre 2022.

[428] Aujourd'hui devenue l'*Institut de Radioprotection et de Sûreté Nucléaire* (IRSN).

[429] J.-Y. Henry, « La sûreté des logiciels dans les installations nucléaires », *Revue Contrôle de l'Autorité de sûreté nucléaire*, octobre 1999, n°131, Octobre 1999, Dossier : Les systèmes informatiques dans l'industrie nucléaire.

Quantification des Incertitudes (VVQI), jugée plus fiable et devenue l'état de l'art en matière de certification de logiciels.

192. **La méthode de qualification des « outils de calcul scientifique ».** La suite du propos se concentrera sur une catégorie spécifique de logiciels pouvant être utilisés dans le contexte de la sûreté nucléaire, dont la procédure de qualification fait l'objet de textes précis : les *« outils de calcul scientifique »*[430]. Ces derniers sont des programmes logiciels capables de réaliser des simulations numériques de phénomènes physiques. Ils sont composés d'un ou plusieurs résolveurs capables de calculer et résoudre un problème mathématique (des équations) ainsi que de préprocesseur et post-processeur. Les résolveurs sont conçus en plusieurs étapes : modélisation mathématique d'un phénomène physique conduisant à construire un système d'équations, développement d'algorithmes capables de résoudre les équations ainsi obtenues et traduction de ces algorithmes en langage informatique. Les préprocesseurs permettent de traduire les données brutes d'entrées (mesures de phénomènes physiques, pression…) en données mobilisables par un algorithme, tandis que les post-processeurs traitent les résultats des calculs effectués par les résolveurs pour qu'ils soient interprétables par l'être humain, par exemple sous forme de graphiques. Le recours à des techniques d'apprentissage automatique permettrait de simplifier ce processus en construisant des modèles de simulation numérique à partir de données d'exemple, plutôt qu'à partir d'équations mathématiques. Toutefois, de tels systèmes, qu'ils relèvent de l'IA ou non, devraient être qualifiés selon la procédure prévue dans un guide de 2017 établi par l'ASN[431] avant de pouvoir être utilisés dans le cadre la démonstration de sûreté nucléaire.

193. **Plan.** La méthode de qualification des logiciels de sûreté nucléaire vise précisément à prendre en considération l'ensemble de ces facteurs en créant un cadre qui, s'il est respecté, permet de garantir que le système est digne de confiance. Ce processus doit être mis en œuvre par l'exploitant de l'INB et comprend trois étapes : la vérification, la validation et la quantification de l'incertitude (**a**). Cette méthode, appelée « VVQI » n'est pas nouvelle et est

[430] ASN, *Qualification des outils de calcul scientifique utilisés dans la démonstration de sûreté nucléaire – 1re barrière*, Guide pratique, 2017, n° 28.
[431] *Ibid.*

aussi applicable à la fois dans d'autres pays[432] mais aussi dans d'autres industries[433]. L'analyse du contenu de ces étapes montre qu'elles ne sont pas suffisamment adaptées aux spécificités des systèmes d'IA, rendant impossible leur qualification (**b**).

a) La qualification par la méthode « VVQI »

194. **Les trois étapes de la méthode « VVQI ».** La procédure prévue par le guide de l'ASN est fondée sur la méthode « VVQI », très répandue dans l'industrie des logiciels critiques. Elle comprend trois étapes : la vérification, la validation et la quantification des incertitudes.

195. **L'étape de vérification.** La première étape peut être définie comme le processus visant à établir que le système de calcul représente de façon adéquate le modèle mathématique sous-jacent et sa solution[434]. En d'autres termes, elle vise à répondre à la question « Le logiciel a-t-il été développé correctement ? » et plus précisément « Le logiciel réalise-t-il correctement la tâche pour laquelle il a été conçu ? ». Ainsi, la vérification consiste à vérifier l'absence d'anomalies dans le code, dans la transposition des équations mathématiques en algorithmes et dans les résultats obtenus lorsqu'on les compare à des résultats théoriques obtenus en résolvant les équations physiques manuellement[435]. À ce stade, on ne questionne pas le bienfondé des équations retenues pour la modélisation mais seulement leur bonne traduction sous un format « solvable » par un programme informatique.

196. **L'étape de validation.** La deuxième étape va plus loin que la vérification. Elle consiste en un processus visant à déterminer si le modèle retenu et les données utilisées représentent

---

[432] Voir notamm. NATIONAL RESEARCH COUNCIL, *Assessing the Reliability of Complex Models: Mathematical and Statistical Foundations of Verification, Validation, and Uncertainty Quantification*, The National Academies Press, 2012, 144 p.

[433] SOCIETE FRANÇAISE DES MECANICIENS, *Guide de validation des progiciels de calcul des structures*, guide AFNOR technique, 1990, 372 p. ; AMERICAN SOCIETY OF MECHANICAL ENGINEERS (ASME), *Standard for Verification and Validation in Computational Fluid Dynamics and Heat Transfer*, VV 20–2009(R2021), 100 p., disponible en ligne : < https://www.asme.org/codes-standards/find-codes-standards/v-v-20-standard-verification-validation-computational-fluid-dynamics-heat-transfer>, consulté le 18 septembre 2022.

[434] AMERICAN SOCIETY OF MECHANICAL ENGINEERS (ASME), *Standard for Verification and Validation in Computational Fluid Dynamics and Heat Transfer*, Standard technique, 2009, V&V 20-2009.

[435] W. Oberkampf, C. Roy, « Verification and validation in scientific computing », *Cambridge University Press*, 2010, Cambridge, UK, p. 249 et s.

précisément le monde réel[436]. Son objectif est donc de répondre aux questions « Le logiciel représente-t-il précisément le phénomène réel qu'il est censé simuler ? » ou « A-t-on correctement choisi le logiciel à concevoir ? ». Là où la vérification se contente de contrôler si le logiciel résout bien les équations modélisées numériquement, la validation va jusqu'à questionner le choix de l'équation et des paramètres à modéliser. L'ASN recommande également dans son guide de comparer les résultats du système logiciel contrôlé avec des « données de validation », c'est-à-dire des données expérimentales, ou des données issues de modèles de référence[437]. L'étude de cette comparaison permet de passer à la dernière étape : la quantification des incertitudes.

197. **L'étape de quantification des incertitudes.** La troisième et dernière étape consiste en l'analyse des divergences entre les résultats du système logiciel testé et les données de validation ou mesures expérimentales obtenues à partir des mêmes données d'entrée. Cette analyse doit permettre de mettre en lumière une marge d'incertitudes (exemple : 95% de probabilité de réussite) qui devrait se situer dans un intervalle dit « de confiance », qui diffère suivant les pratiques des industries. Si une telle analyse ne peut être réalisée, il est possible de recourir à une étude de sensibilité, consistant à tester le modèle avec des paramètres d'entrée différents afin de déterminer l'impact de chaque paramètre sur les résultats finaux.

198. **Transition.** Cette procédure en trois étapes, fondée sur la méthode VVQI, est utilisée depuis de nombreuses années puisque les techniques de simulation numérique n'ont que peu évolué. Toutefois, les récents progrès des techniques d'IA, notamment dans le domaine de l'apprentissage automatique, laissent penser qu'elles peuvent être plus performantes dans la simulation numérique que les systèmes déterministes classiques. Or, la méthode de qualification des logiciels de sûreté nucléaire apparaît peu adaptée à ces nouvelles techniques.

---

[436] AMERICAN SOCIETY OF MECHANICAL ENGINEERS (ASME), *Standard for Verification and Validation in Computational Fluid Dynamics and Heat Transfer*, *op. cit.*

[437] ASN, *Qualification des outils de calcul scientifique utilisés dans la démonstration de sûreté nucléaire – 1re barrière*, Guide pratique, 2017, n° 28, p. 8.

b) La difficile application de la méthode VVQI aux techniques d'apprentissage automatique

199. **Des techniques radicalement différentes.** Contrairement aux logiciels utilisés jusqu'à présent qui simulaient numériquement un phénomène en faisant résoudre à la machine une série d'équations issues des sciences physique et nucléaire, les systèmes fondés sur des techniques d'apprentissage automatique construisent eux-mêmes les équations représentant le phénomène à partir de l'analyse de jeux de données d'observation dites « d'entraînement ». La construction d'un modèle mathématique n'est possible que s'il existe un consensus sur les causes « X » d'un phénomène « Y » ainsi que sur les équations décrivant la relation entre X et Y selon les lois de la physique. Ce sont ces équations qui sont traditionnellement représentées de façon numérique dans un logiciel de calcul. En revanche, un modèle d'apprentissage automatique va reproduire le phénomène Y à partir de corrélations qu'il identifiera dans un jeu de données qu'on lui aura fourni : le jeu de données d'entraînement. Ainsi, la conception d'un logiciel de simulation numérique est donc complètement différente selon la technique utilisée. D'un côté, avec une modélisation mathématique classique, il sera essentiel d'identifier dans la littérature scientifique les équations décrivant le phénomène à simuler et de les transposer fidèlement dans le code informatique. De l'autre, avec l'apprentissage automatique, il sera plus important de sélectionner avec attention les données pertinentes et suffisamment représentatives ainsi que de choisir le mode d'apprentissage le plus adapté (supervisé, non supervisé ou par renforcement).

200. **L'inadéquation de la méthode de qualification aux nouvelles approches de conception logicielle.** Cette différence fondamentale entre les approches laisse penser que la méthode de qualification décrite précédemment ne pourra s'appliquer pleinement aux systèmes d'IA, privant donc les opérateurs de centrales nucléaires de la possibilité de les utiliser dans le cadre de la démonstration de sûreté.

D'abord, concernant la phase de vérification et étant donné qu'un logiciel fondé sur des techniques d'apprentissage automatique peut être opaque contrairement à un logiciel traditionnel exécutant des règles précises, il serait très laborieux de vérifier la bonne transposition des équations mathématiques dans le système et l'absence d'anomalies dans le code. Dans le domaine de l'IA, c'est la donnée et non le code informatique qui est au cœur du fonctionnement du système.

Ensuite, concernant la phase de validation et en l'absence de règles précises ancrées dans le marbre du code informatique, il ne serait pas pertinent de chercher si les équations ont été correctement sélectionnées pour simuler le phénomène physique souhaité.

De plus, comme un grand nombre de modèles d'apprentissage automatique tels que les réseaux de neurones profonds sont difficilement explicables et fonctionnent comme des boîtes noires, il serait très difficile de déterminer avec précision les données d'entrée et paramètres ayant un impact sur les résultats obtenus par le système.

Enfin, les modèles d'apprentissage automatique visent à reproduire des phénomènes à partir de corrélations identifiées entre les données utilisées dans la phase d'entraînement du système. Toutefois, une forte corrélation entre des variables ne garantit pas un lien de causalité et peut rester le fruit d'une coïncidence. Il peut également exister une variable explicative non connue ou non présente dans le jeu de données d'entraînement. Il serait donc dangereux de se fier intégralement à un système conçu à partir de simples corrélations au vu des risques qu'une erreur pourrait causer.

201. **La nécessaire adaptation de la méthode de qualification.** Au regard des différences fondamentales existant entre les systèmes déterministes classiques et ceux fondés sur des techniques d'apprentissage automatique, la méthode prescrite par la règlementation pour qualifier des logiciels utilisés dans le cadre de la démonstration de sûreté nucléaire mériterait d'être repensée. Ses principes sont pertinents et doivent être conservés mais leur application en pratique doit être adaptée pour prendre en considération les nouvelles approches fondées sur la donnée et non sur la rédaction de règles logiques.

202. **Transition.** Toutefois, la seule adaptation de la procédure de qualification des logiciels utilisés dans le cadre de la sûreté nucléaire ne saurait garantir à elle seule une utilisation sûre de l'IA dans les infrastructures critiques du secteur de l'électricité. D'autres pistes d'évolution de la règlementation pourraient faire l'objet de discussions entre les parties prenantes

Section 2 : Des adaptations nécessaires pour permettre une utilisation sûre de l'IA dans les infrastructures critiques

203. **Plan.** S'il apparaît aujourd'hui justifié que l'usage de l'IA dans les infrastructures critiques du secteur de l'électricité soit limité pour des raisons de sécurité, ce constat pourrait être amené à évoluer. En effet, de nombreux chercheurs estiment que certaines applications de cette technologie pourraient au contraire être employées pour renforcer la sécurité, notamment contre de potentielles attaques informatiques. De plus, des systèmes d'IA peuvent aussi être utilisés dans les réseaux de transport d'électricité ou dans les centrales nucléaires sans accroître le risque déjà existant. En effet, certaines applications permettent seulement de réaliser des calculs plus rapidement, d'analyser des flux de données si importants qu'ils n'étaient simplement pas contrôlés auparavant ou encore d'aider les opérateurs dans leurs décisions, sans jamais se substituer aux procédures décisionnelles existantes. Pourtant, ces applications vertueuses souffrent également de l'incompatibilité manifeste des règles de sécurité en vigueur, ce qui entrave leur développement. Dès lors, il apparaît nécessaire d'engager une réflexion sur les leviers à notre disposition pour permettre le recours à l'IA dans les infrastructures critiques du secteur de l'électricité tout en garantissant la sûreté. Cette réflexion doit faire l'objet d'une discussion pluridisciplinaire et impliquer à la fois les régulateurs sectoriels (ANSSI, ASN), les OIV concernés (EDF S.A., RTE et autres), des experts indépendants par exemple issus du milieu universitaire, ainsi que le grand public pour mesurer l'acceptabilité des souplesses règlementaires envisagées. L'étude de la littérature scientifique et la rencontre avec des professionnels[438] nous ont permis de mettre en lumière plusieurs pistes de solution. Les premières consistent en l'éventuel assouplissement des règles en vigueur et de la doctrine des autorités sectorielles pour s'adapter aux spécificités des systèmes d'IA (**§1**). Ces assouplissements devront être pensés rigoureusement pour n'autoriser que les applications non risquées, ce qui n'est pas chose aisée. Les autres pistes de solution relèvent plutôt de la technique. Elles consistent en l'adaptation des systèmes d'IA aux règles existantes (**§2**), ce qui peut nécessiter un effort de recherche supplémentaire, par exemple pour parvenir à concevoir

[438] Voir Supra, 63.

des systèmes explicables, ou de nouvelles approches innovantes combinant performance de l'IA et auditabilité des logiciels classiques.

§1 : Une adaptation de l'ordre juridique aux spécificités des systèmes d'IA

204. **Plan.** Les cadres juridiques relatifs à la sécurité des SIIV et à la sûreté nucléaire sont particulièrement exhaustifs. Ils le sont tellement qu'ils ont formaté toute une pratique comme un référentiel, qu'il suffit de suivre à la lettre pour garantir le résultat escompté (en l'espèce la sécurité des systèmes critiques). Ce constat est paradoxal. D'une part, il est positif car il facilite la mise en conformité des opérateurs d'infrastructures critiques et garantit un niveau de sécurité élevé et uniforme sur le territoire d'application de la norme. D'autre part, il présente naturellement un effet négatif sur l'innovation. Pourquoi un opérateur irait-il investir pour rechercher de nouvelles solutions innovantes s'il se trouve bridé en fin de processus par un carcan règlementaire ? De plus, ces modèles de règlementation exhaustive souffrent d'immobilisme. Le corpus de règles est si étoffé et précis qu'il est impossible de le faire évoluer au fil des avancées technologiques. Les logiciels les plus sûrs et performants aujourd'hui ne sont pas forcément ceux d'hier : *quid* si la règlementation ne s'est pas adaptée suffisamment rapidement ? Nous pensons qu'il est essentiel que les autorités règlementaires et les agences publiques chargées de la mise en œuvre de la règlementation lancent une réflexion d'ampleur pour réformer le cadre en vigueur afin de créer un environnement juridique favorable à l'innovation dans les systèmes critiques (**A**). Ces derniers doivent présenter le plus haut niveau de sécurité, il est donc pertinent qu'ils disposent des systèmes de protection les plus avancés et innovants. Pour ce qui concerne les systèmes logiciels, dont les systèmes d'IA font partie, la fiabilité et la sécurité sont souvent évaluées à travers une procédure de certification. Si le logiciel passe toutes les étapes de la procédure alors il sera considéré comme sûr et apte à l'utilisation. Les procédures de certification utilisées dans les règlementations sectorielles du secteur de l'électricité ont été pensées pour des logiciels déterministes. Une règlementation équilibrée ne peut être fondée que sur des méthodes de certification adaptées aux systèmes les plus actuels (**B**).

A/ Un possible assouplissement du cadre juridique en faveur de l'innovation dans les systèmes critiques

205. **La nécessaire discussion collégiale sur l'opportunité d'adapter le corpus règlementaire.** L'objet du présent Paragraphe n'est pas de dire qu'il faut à tout prix assouplir la règlementation sectorielle relative à la sécurité des systèmes d'information utilisés dans les infrastructures critiques de l'énergie. Cette conclusion ne peut être atteinte que s'il existe un consensus de la communauté scientifique, des autorités de régulation et des opérateurs sur le fait qu'il est possible de garantir une utilisation sûre de l'IA dans des systèmes critiques. Il apparaît donc nécessaire d'initier une discussion collégiale, d'une part, pour questionner la proportionnalité des contraintes existantes au vu des spécificités des systèmes d'IA et, d'autre part, pour étudier les moyens techniques et procédures qui permettraient d'atteindre un niveau de confiance suffisant. En l'absence de maîtrise parfaite de tels systèmes, il est naturel d'en exclure l'utilisation pour des fonctions critiques.

206. **La nécessaire préservation de la capacité d'innovation des entreprises.** Néanmoins, persiste la question de savoir si le corpus actuel est compatible avec une dynamique d'innovation. En dépit de son exhaustivité, est-il en capacité de s'adapter suffisamment rapidement pour permettre l'utilisation de nouveaux systèmes sûrs bien que différents des logiciels classiques ? Ou bien agit-il comme un carcan dissuasif et empêche-t-il les opérateurs d'installations critiques d'utiliser de nouvelles techniques, fussent-elles plus sûres ? L'innovation technologique, notamment dans le domaine des technologies numériques, est synonyme d'optimisation des processus en facilitant des calculs ou plus généralement en permettant la réalisation de tâches jusqu'alors trop complexes pour être réalisées par des logiciels classiques. Bien dirigée, l'innovation peut donc également conduire à réduire des risques[439]. Par exemple, des systèmes d'IA peuvent être employés pour automatiser de nombreuses tâches simples et répétitives, si bien que le risque humain lié à la répétition d'un geste disparaîtrait. Or, dans le secteur de l'électricité, notre étude de la règlementation applicable révèle qu'elle exclut l'utilisation de systèmes d'IA pour plusieurs raisons (absence de standards de sécurité adaptés aux modèles d'apprentissage automatique, procédure de

[439] Pour des exemples dans l'industrie de la construction : O.S. Abioye, L.O. Oyedele, L. Akanbi, *et al.*, « Artificial intelligence in the construction industry: A review of present status, opportunities and future challenges », *Journal of Building Engineering*, 2021, vol. 44, 103299.

certification inadaptée…)[440]. Deux types d'initiatives peuvent être menés pour pallier cette situation.

207. **La promotion de l'innovation par la mise en place de mécanismes juridiques incitatifs.** Premièrement, il est possible d'utiliser des mécanismes juridiques pour créer un environnement favorable à l'innovation. La création de bacs à sable règlementaires, comme le pratiquent l'ARCEP[441] ou la CNIL[442] dans leur domaine, pourrait être transposée à la sécurité des systèmes d'informatique critiques. Or, ces cadres d'expérimentation ne sont possibles que parce qu'ils entrent dans le cadre des missions des autorités de contrôle prévues par la loi[443]. Sur le même modèle, l'ANSSI pour les SIIV et l'ASN pour les systèmes utilisés dans les centrales nucléaires pourraient ainsi mettre en place des appels à projets d'innovation dans les systèmes critiques et permettre à des opérateurs de développer et tester des outils logiciels en bénéficiant de souplesses règlementaires, avec la supervision rapprochée des autorités.

208. **La promotion de l'innovation par le levier du financement.** Deuxièmement, la stimulation de l'innovation peut passer par le financement d'ambitieux programmes de recherche ou par la commande publique. Aux États-Unis, le département de l'énergie est très proactif dans la recherche d'applications d'IA dans la production nucléaire d'électricité, que ce soit en menant ses propres recherches[444] ou en finançant des projets sélectionnés pour leur caractère disruptif[445]. Ce domaine n'est que peu investi en France, laissant à penser que

[440] Voir Supra, 157-201.
[441] ARCEP, « Bac à sable règlementaire », *Site officiel de l'ARCEP*, 28 juin 2022, disponible en ligne : <https://www.arcep.fr/professionnels/startups-entrepreneurs/bac-a-sable-reglementaire.html>, consulté le 22 septembre 2022.
[442] CNIL, « La CNIL propose un nouveau « bac à sable » pour accompagner l'innovation numérique dans le domaine de l'éducation », *Site officiel de la CNIL*, 18 janvier 2022, disponible en ligne : <https://www.cnil.fr/fr/la-cnil-propose-un-nouveau-bac-sable-pour-accompagner-linnovation-numerique-dans-le-domaine-de>, consulté le 22 septembre 2022.
[443] La loi pour une République numérique ayant modifié l'article L42-1 du Code des postes et communication pour les pouvoirs de l'ARCEP et la loi informatique et libertés (article 11) pour la CNIL.
[444] C. Nunez, « How artificial intelligence could lower nuclear energy costs », *Argonne National Laboratory (blog)*, disponible en ligne : <https://www.anl.gov/article/how-artificial-intelligence-could-lower-nuclear-energy-costs>, consulté le 22 septembre 2022.
[445] U.S. OFFICE OF SCIENCE, « Department of Energy Announces $5.7 Million for Research on Artificial Intelligence and Machine Learning (AI/ML) for Nuclear Physics Accelerators and Detectors », *Office of Science (blog)*, 2 décembre 2021, disponible en ligne : <https://www.energy.gov/science/articles/department-energy-announces-57-million-research-artificial-intelligence-and>, consulté le 22 septembre 2022.

l'utilisation de l'IA dans des environnements critiques tels que des centrales nucléaires est tabou, ce qui est regrettable au vu des potentiels bénéfices que les opérateurs pourraient en tirer.

209. **Conclusion et transition.** En somme, il est possible de créer un environnement juridique plus favorable à l'innovation dans les infrastructures critiques de l'énergie sans assouplir la règlementation directement. Une discussion d'ampleur entre les parties prenantes (exploitants, régulateurs, universitaires, grand public) devrait avoir lieu pour étudier l'ensemble de ces pistes et questionner la proportionnalité du cadre existant au regard des bénéfices que l'IA pourrait apporter en termes de sécurité ou d'optimisation du fonctionnement dans les systèmes critiques. L'assouplissement de la règlementation ne peut intervenir qu'en dernier recours, lorsqu'il existe un consensus scientifique sur les moyens de garantir la sûreté des systèmes d'IA, ce qui peut passer par la création de standards techniques harmonisés ou le recours à des méthodes de certifications adaptées.

B/ Le recours à des méthodes de certification adaptées

210. **Plan.** Dans un domaine aussi technique que celui des systèmes logiciels utilisés dans des fonctions critiques il est normal que l'essentiel du cadre juridique soit constitué de normes techniques, contraignantes ou non. Ainsi, l'adaptation de l'ordre juridique aux spécificités des systèmes d'IA pour permettre leur utilisation dans des infrastructures critiques peut passer par la modernisation des standards techniques faisant aujourd'hui référence ou par la création de nouveaux. Ces deux options seront illustrées dans la suite de notre propos. D'une part, une modernisation de la norme relative à la méthode de qualification des systèmes logiciels par l'ASN apparaît possible (**1**). D'autre part, il existe une littérature scientifique abondante sur les principes qui pourraient fonder de nouvelles normes de certification applicables et adaptées aux systèmes d'IA (**2**). Leur concrétisation permettrait de surmonter les difficultés actuelles relatives à l'utilisation de l'IA dans les systèmes critiques.

1. Une possible modernisation de la doctrine de l'ASN relative à la qualification des logiciels de sûreté

211. **Adaptation de l'étape de vérification.** Afin de garantir la sécurité d'un système, la phase de vérification pourrait se concentrer sur le contrôle du bon fonctionnement du modèle d'IA entraîné, notamment en vérifiant que les résultats donnés sont cohérents par rapport aux

données d'entrée, et sur l'absence d'anomalies dans le code et les données utilisées. À ce stade, le choix du mode d'apprentissage employé et des données d'entraînement ne devrait pas être questionné. La phase de vérification se concentrerait ainsi, tout comme dans le processus actuel, sur la performance du modèle en soi et non sur sa pertinence au regard de l'objectif poursuivi.

212. **Adaptation de l'étape de validation.** Une phase de validation adaptée aux techniques d'apprentissage automatique pourrait, elle, consister dans le contrôle du choix de la technique d'apprentissage ainsi que des choix réalisés dans la sélection des données d'entraînement et de test : le mode d'apprentissage choisi (supervisé, non supervisé, par renforcement) est-il le plus adapté à la situation et celui capable de produire les résultats les plus précis ? Comment le jeu de données d'apprentissage a-t-il été constitué ? Était-il suffisamment représentatif et exempt de biais ?

213. **Adaptation de l'étape de quantification des incertitudes.** La phase de quantification des incertitudes pose moins de problème au regard des modèles d'apprentissage puisqu'il s'agit seulement de comparer les prédictions du logiciel avec des données expérimentales du monde réel. En revanche, il peut être plus difficile de conduire une étude de sensibilité. En effet, certaines techniques d'IA sont opaques, ce qui rend impossible de savoir le poids de certains paramètres dans le fonctionnement du système. La constitution d'un jeu de données de validation suffisamment conséquent et sélectionné avec attention sera nécessaire pour estimer la performance du modèle et ainsi garantir, le cas-échéant, qu'il est digne de confiance.

214. **Conclusion sur l'adaptation de la méthode VVQI.** Repenser le processus de qualification pour les logiciels nucléaires nécessiterait également que l'autorité de régulation modifie sa doctrine pour prendre en compte ces adaptations. En France, c'est le guide de l'ASN dédié à cette procédure[446] qu'il conviendrait d'amender en conséquence. Toutefois, il appartient à la communauté scientifique de parvenir à un consensus sur le fait de savoir si l'adaptation de la méthode VVQI suffirait à garantir la sécurité d'un logiciel fondé sur des techniques d'apprentissage automatique. Dans l'attente, il est plus prudent et raisonnable de ne pas avoir recours à ces techniques dans un contexte aussi critique que celui de la sûreté nucléaire.

---

[446] ASN, *Qualification des outils de calcul scientifique utilisés dans la démonstration de sûreté nucléaire – 1re barrière*, Guide pratique, 2017, n° 28.

2. Vers une création de nouvelles normes relatives à la certification des systèmes d'IA

215. **La pertinence de la création de méthodes de certification adaptées aux spécificités de l'IA.** Le corpus normatif relatif à la sécurité informatique des infrastructures critiques a été pensé pour l'utilisation de logiciels déterministes. Pour s'en convaincre, il suffit de lire n'importe quelle procédure de certification logicielle, imposant de contrôler le code et les règles implémentées dans le système sans jamais parler de données, pourtant au cœur du fonctionnement de l'IA. De nombreux travaux scientifiques sont menés pour pallier cette lacune en créant de nouvelles normes adaptées aux techniques d'apprentissage machine bien qu'il n'existe aucun consensus à l'heure actuelle. Cette question, éminemment technique, s'éloigne du champ juridique de notre étude. C'est pourquoi nous nous contenterons de donner ici quelques exemples de pistes de réflexion présentes dans la doctrine, sans avoir la prétention de pouvoir juger de leur pertinence.

216. **Un exemple de projet de recherche français sur la certification de l'IA.** En France d'abord, la chaire de recherche ANITI[447] dispose d'un axe dédié à la certification de l'IA. Dans un livre blanc de 2021, plusieurs de ses chercheurs recommandent la création d'une norme relative à la sûreté des algorithmes d'apprentissage automatique, sur le modèle des normes DO-178C/ED-12C (aviation civile), EN50128 (systèmes ferroviaires) ou encore ISO 26262 (automobile)[448]. Selon eux, cette norme pourrait reposer non pas sur la méthode « Vérification & Validation » étudiée précédemment mais sur une nouvelle méthode découpée en trois étapes : une analyse détaillée du processus de conception du système d'IA (du choix des données et de l'algorithme jusqu'à son utilisation), l'analyse des similitudes et différences entre les techniques employées et celles utilisées dans des logiciels déjà certifiés, ainsi qu'une analyse rétrospective visant à identifier au cours de l'utilisation ou dans le passé l'existence de potentielles vulnérabilités[449]. Le groupe de chercheurs a également identifié huit propriétés qui,

[447] *Artificial and Natural Intelligence Toulouse Institute*.
[448] F. Mamalet, E. Jenn, G. Flandin, *et al.*, *Machine learning in certified systems*, livre blanc, ANITI – IRT Saint-Exupéry, 2021, p. 17 et s.
[449] *Ibid.*, p. 21.

si elles sont détenues par un système d'IA, devraient permettre de le qualifier comme sûr et digne de confiance[450] :

- La capacité à être audité ou à être vérifié par un examen indépendant ;
- La qualité des données, correspondant au fait que les données utilisées pour l'apprentissage, les tests et le fonctionnement doivent être autant que possible exemptes d'erreurs et présenter des caractéristiques pertinentes au regard de la finalité du système ;
- L'explicabilité ou le fait que le fonctionnement du système puisse être compris par des humains ;
- La maintenabilité ou la possibilité de faire évoluer le système en conservant sa conformité aux exigences de sûreté ;
- La résilience ou la capacité du système à continuer à fonctionner malgré erreur ou défaillance ;
- La robustesse, qui peut être globale (capacité du système à fonctionner correctement même en présence de données d'entrées anormales ou inconnues) ou locale (capacité du système à produire continuellement les mêmes résultats pour les mêmes entrées) ;
- La capacité à être spécifié : le système doit pouvoir être décrit précisément par une liste d'éléments objectifs, notamment ses fonctions, ses performances…
- La capacité à être évalué sur la base de critères objectifs.

Ces travaux constituent un point de départ intéressant bien que chaque axe doive maintenant faire l'objet de recherches plus spécifiques, par exemple pour définir des critères pour évaluer le degré d'explicabilité d'un système d'IA ou pour déterminer quelles doivent être les mesures à mettre en œuvre pour assurer la « qualité des données ». Ce n'est que lorsque chacun de ces

[450] *Ibid.*, p. 22.

axes sera décliné sous la forme de critères objectifs qu'une normalisation de la certification sera envisageable.

217. **Un exemple de projet de recherche américain sur la certification de l'IA.** Ensuite, des travaux similaires ont été menés aux États-Unis, sous un angle plus théorique, allant jusqu'à questionner la pertinence même du processus de certification. En effet, pourquoi la garantie de la sûreté logicielle ne pourrait être apportée que par une méthode bureaucratique consistant à vérifier un certain nombre de critères à un instant T, suivant une méthode harmonisée et unique pour tous les logiciels ? Ne pourrait-on penser une nouvelle forme de démonstration de la sûreté logicielle[451] ? Les travaux de la Professeure Nancy Leveson, pionnière en matière de sécurité des systèmes et logiciels, dans son ouvrage *Engineering a safer world*[452] contiennent à cet égard de nombreuses pistes de réflexion très intéressantes. Par exemple, il est nécessaire selon elle de mettre l'accent sur la complémentarité homme-machine pour superviser le fonctionnement du système d'IA plutôt que de chercher à tout prix à atteindre un risque zéro, ce qui ne sera jamais possible. Aussi, il faut, d'après elle, prévoir la maintenabilité de la sécurité du système dès la conception, en prévoyant les moyens techniques pour identifier d'éventuelles défaillances et réaliser les mises à jour nécessaires pour les résoudre[453].

218. **La récente impulsion européenne sur le sujet de la certification de l'IA.** Enfin, au sein de l'Union européenne, le projet de règlementation sur l'IA a poussé la communauté scientifique[454] et les régulateurs[455] à accélérer leur recherche de moyens d'auditer et de certifier

---

[451] Voir à ce sujet : A. Rudolph, S. Voget, J. Mottok, « A consistent safety case argumentation for artificial intelligence in safety related automotive systems », *ERTS 2018*, janvier 2018, Toulouse, France, hal-02156048 ; A. Wassyng, T. Maibaum, M. Lawford, H. Bherer, « Software Certification : Is There a Case against Safety Cases ? », in *Foundations of Computer Software. Modeling, Development, and Verification of Adaptive Systems*, dir. R. Calinescu, E. Jackson, Monterey Workshop 2010, Lecture Notes in Computer Science, Springer, 2011, vol. 6662.
[452] N.G. Leveson, *Engineering a Safer World: Systems Thinking Applied to Safety*, MIT Press, 2012, Cambridge, MA, USA.
[453] L'ensemble des pistes de réflexion de N.G. Leveson pour parvenir à des systèmes d'IA sécurisés sont présentées et commentées dans R. Dobbe, « System Safety and Artificial Intelligence », *2022 ACM Conference on Fairness, Accountability, and Transparency (FAccT '22) Proceedings*, Association for Computing Machinery, 2022, New York, USA, p. 1584.
[454] G. Falco, B. Shneiderman, J. Badger, *et al.*, « Governing AI safety through independent audits », *Nature Machine Intelligence*, 2021, vol. 3, n°7, pp. 566–571.
[455] A. Vitard, « Dans la perspective de la nouvelle réglementation sur l'IA, la Cnil se prépare à auditer les algorithmes », *L'usine digitale (blog)*, 22 septembre 2022, disponible en ligne : < https://www.usine-digitale.fr/article/dans-la-perspective-de-la-nouvelle-reglementation-sur-l-ia-la-cnil-se-prepare-a-auditer-les-algorithmes.N2046717>, consulté le 24 septembre 2022.

les algorithmes d'IA. Bien que ces travaux n'aient pas abouti à la reconnaissance d'une norme partagée pour le moment, tout porte à croire que l'entrée en vigueur prochaine de l'*AI Act* devrait donner la dernière impulsion dont la standardisation a besoin.

219. **Conclusion du §1 sur l'adaptation du corpus juridique applicable aux systèmes d'IA utilisés dans les infrastructures critiques.** En ce qui concerne la règlementation applicable à la sécurité des infrastructures critiques de l'énergie, l'adaptation de l'ordre juridique aux spécificités des systèmes d'IA peut emprunter plusieurs voies. De la création de bacs à sable règlementaires pour favoriser l'innovation, à l'adaptation de la doctrine des autorités sectorielles en passant par l'octroi de la juridicité à de nouvelles normes de certification, les mécanismes envisageables sont nombreux. Le choix à opérer relève néanmoins de questions éminemment techniques : les bénéfices que peuvent apporter les systèmes d'IA aux exploitants de centrales nucléaires justifient-ils qu'on fasse évoluer le cadre juridique existant ? Ne risque-t-on pas d'introduire de nouvelles vulnérabilités dans des infrastructures critiques pour la sécurité ? Seules les autorités de régulation, les experts techniques et les exploitants de ces systèmes peuvent répondre à ces interrogations. C'est pourquoi nous nous sommes contentés d'évoquer les solutions envisageables du point de vue juridique. Le reste de la décision relève du scientifique et du politique.

220. **Transition.** Néanmoins, si la conclusion de ce débat était que le cadre règlementaire ne devait pas évoluer pour des raisons de sécurité, cela sonnerait-il le glas de l'IA dans les centrales nucléaires ? Une autre voie est à explorer : à défaut d'adapter la règlementation des OIV et de la sûreté nucléaire à l'IA, ne peut-on pas adapter l'IA à l'ordre juridique ?

§2 : Une adaptation des systèmes d'IA à l'ordre juridique

221. **Plan.** L'assouplissement de la doctrine juridique à la faveur de l'utilisation de systèmes d'IA dans les infrastructures critiques du secteur de l'électricité ne doit pas conduire à autoriser l'utilisation de systèmes que l'on ne comprend pas ou que l'on ne maîtrise pas parfaitement. À l'heure actuelle, de nombreux systèmes d'IA, notamment ceux fondés sur des réseaux de neurones profonds, fonctionnent sans qu'il soit possible de comprendre et d'expliquer précisément les raisons pour lesquelles ils ont abouti à tel résultat à partir de telles données d'entrée. En fonctionnant de manière opaque (en « boîte noire » selon l'expression consacrée), il est difficile d'anticiper avec certitude la façon dont le système va se comporter dans la durée. Comment, dès lors, pourrait-on avoir confiance dans son fonctionnement ? Pour permettre le recours à l'IA dans des systèmes critiques, il est indispensable que celle-ci gagne en explicabilité (**A**). Il s'agit là d'un prérequis à tout assouplissement du corpus juridique. Une autre solution technique permettrait le recours à l'IA dans des conditions où la confiance et la sécurité sont cruciales. Elle consiste à combiner l'approche probabiliste de l'IA avec des approches symboliques plus classiques, fondées sur l'application de règles et donc beaucoup plus facilement certifiables (**B**). Dans ce cas-là, la technique d'IA n'a pas gagné en explicabilité mais l'introduction de briques auditables dans le système global permet d'avoir un haut niveau de confiance dans son fonctionnement.

A/ Vers des systèmes d'IA explicables ?

222. **L'indispensable explicabilité des systèmes d'IA utilisés dans des infrastructures critiques.** L'utilisation de systèmes d'IA dans des fonctions critiques pour la sûreté – qu'importe le secteur d'activités concerné – serait grandement facilitée si les techniques employées étaient facilement auditables et explicables. En effet, les principales limites de l'application des processus de certification logicielle aux systèmes fondés sur des techniques d'IA tiennent au fait que ces dernières ne sont pas aussi transparentes que des logiciels déterministes traditionnels, pour lesquels il suffit de contrôler la pertinence des règles logiques implémentées. Finalement, le degré de criticité des environnements dans lesquels l'IA doit pouvoir être utilisée dépend du degré de précision avec lequel son fonctionnement peut être expliqué. Partant de ce constat, on comprend aisément pourquoi la recherche est très active à

ce sujet[456]. De nombreuses pistes très prometteuses sont explorées notamment pour fournir des explications dites locales, consistant à expliquer les raisons ayant conduit à un résultat donné. Ces pistes peuvent consister, par exemple, dans le développement de petits applicatifs logiciels qui, accolés aux systèmes d'IA, collectent des données sur leur fonctionnement afin de pouvoir accompagner leur résultat d'une courte explication et, le cas-échéant, assurer la traçabilité de la décision (pouvoir remonter aux données et raisons ayant conduit à tel ou tel résultat)[457]. À l'échelle internationale, certaines autorités de régulation travaillent d'ailleurs activement pour proposer des standards d'explication adaptés à chaque technique d'IA[458]. En France, la CNIL a d'ailleurs expérimenté en 2022 les outils d'explication algorithmique développés par le chercheur Clément Hénin dans sa thèse soutenue en 2021[459]. Bien qu'il faille encore un peu de temps pour que les techniques permettant d'expliquer le fonctionnement de l'IA parviennent à maturité, leur développement permettra de pouvoir auditer ces systèmes et ainsi faciliter l'évaluation de leur robustesse technique.

223. **Les liens entre explicabilité, confiance et acceptabilité du recours à l'IA.** La question de la confiance dans les modèles d'apprentissage automatique ne peut être résolue qu'à travers un processus de certification ou un audit de leur fonctionnement. En effet, certaines applications peuvent également poser des problèmes d'acceptabilité et nécessiter que leur fonctionnement soit expliqué à différentes parties prenantes : utilisateurs, opérateurs ou régulateurs. Le contenu de l'explication à apporter peut donc être très différent suivant le contexte et son destinataire :

---

[456] K. Gade, S.C. Geyik, K. Kenthapadi, *et al.*, « Explainable AI in Industry », in *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (KDD '19)*, 2019, Association for Computing Machinery, New York, 3203–3204 ; R. Goebel, A. Chander, K. Holzinger, *et al.*, « Explainable AI : The New 42 ? », in *Machine Learning and Knowledge Extraction*, dir. A. Holzinger, P. Kieseberg, A. Tjoa, *et al.*, Springer, 2018, CD-MAKE 2018, conference proceedings, vol. 11015, pp. 295–303.

[457] Sur les méthodes permettant d'expliquer le fonctionnement de modèles d'apprentissage automatique, voir l'article de référence : S.M. Lundberg, G. Erion, H. Chen, *et al.*, « From local explanations to global understanding with explainable AI for trees », *Nature Machine Intelligence*, 2020, vol. 2, 56–67.

[458] Voir notamment les travaux très éclairants de l'autorité anglaise de protection des données : ICO et ALAN TURING INSTITUTE, *Explaining decisions made with AI : Part 1*, draft guidance, 2019, version 1.1., disponible en ligne : <https://ico.org.uk/media/about-the-ico/consultations/2616434/explaining-ai-decisions-part-1.pdf>, consulté le 23 septembre 2022.

[459] C. Henin, *Expliquer et justifier les systèmes de décisions algorithmiques*, thèse pour le doctorat en informatique, Université de Lyon, 2021, 176 p. ; A. Vitard, « Dans la perspective de la nouvelle réglementation sur l'IA, la Cnil se prépare à auditer les algorithmes », *L'usine digitale (blog)*, 22 septembre 2022, disponible en ligne : <https://www.usine-digitale.fr/article/dans-la-perspective-de-la-nouvelle-reglementation-sur-l-ia-la-cnil-se-prepare-a-auditer-les-algorithmes.N2046717>, consulté le 24 septembre 2022.

un utilisateur a seulement besoin d'une compréhension générale du fonctionnement et des limites du système alors qu'une autorité de contrôle comme l'ANSSI va au contraire avoir besoin d'une compréhension précise et beaucoup plus exhaustive. À ce sujet, des chercheurs de Telecom Paris considèrent que la forme de l'explication et le niveau de détail peuvent être dictés par quatre facteurs contextuels : l'audience visée, la pertinence de l'explication fournie, le contexte règlementaire et les contraintes opérationnelles liées aux conditions d'utilisation du système concerné[460].

224. **L'absence de consensus sur les modalités de l'explication des systèmes d'IA.** Toutefois, un long chemin reste encore à parcourir tant pour parvenir à un consensus sur les méthodes pouvant être employées pour expliquer le fonctionnement d'un système d'IA que pour déterminer avec précision les cas où une telle explication serait requise ainsi que son contenu.

225. **Transition.** À défaut de pouvoir comprendre dans le détail la façon dont fonctionnent les modèles d'apprentissage automatique, leur combinaison avec des approches déterministes classiques pourrait apporter plus de transparence au système global. En effet, l'hybridation des techniques est une piste très intéressante pour permettre l'utilisation de l'IA dans les infrastructures critiques, notamment dans le secteur de l'énergie.

B/ Vers une hybridation des approches symboliques et connexionnistes ?

226. **L'intérêt de l'hybridation.** Pour rappel, si l'on adopte une acception large, les techniques d'IA peuvent se répartir en deux grandes catégories. La première, symbolique, englobe les techniques reposant sur l'application de règles logiques par un algorithme. La grande majorité des logiciels et systèmes experts utilisés aujourd'hui, y compris dans des environnements critiques pour la sécurité, font partie de cette catégorie. La seconde catégorie correspond aux techniques dites « connexionnistes » fondées sur l'apprentissage automatique

[460] V. Beaudouin, I. Bloch, D. Bounie, *et al.*, « Identifying the "Right" Level of Explanation in a Given Situation », *Hal archives ouvertes*, 2020, hal-02507316, disponible en ligne : <https://hal.telecom-paristech.fr/hal-02507316>, consulté le 20 avril 2020.

(supervisé, non supervisé ou par renforcement, réseaux de neurones profonds…), sur lesquelles s'est concentrée la présente thèse jusqu'ici. Ce sont ces dernières qui présentent de grandes aptitudes pour le traitement massif de données, pour la réalisation de calculs complexes ou pour la simulation numérique. En revanche, ce sont également elles qui présentent un fonctionnement opaque pouvant soulever des difficultés au regard de la sécurité informatique. Si les systèmes d'IA symboliques sont bien appréhendés par le cadre règlementaire en matière de sécurité des systèmes critiques dans l'énergie, les approches connexionnistes apparaissent, elles, exclues de telles applications. Pourquoi, dès lors, ne pas chercher à combiner les deux approches pour bénéficier à la fois de la transparence de la logique symbolique et de la performance de l'apprentissage automatique ? Cette idée n'est pas nouvelle et fait déjà l'objet de plusieurs publications scientifiques[461], bien que l'on ne puisse pas encore parler de littérature abondante. Les développements qui suivent donnent un exemple concret d'application potentielle dans le domaine de la sûreté nucléaire et de ses atouts en matière de sûreté et de conformité avec le cadre règlementaire.

227. **Un exemple d'hybridation dans le domaine nucléaire.** Depuis plus de trente ans, plusieurs publications scientifiques ont démontré l'intérêt d'utiliser des algorithmes fondés sur des réseaux de neurones profonds afin d'optimiser le plan de rechargement du combustible dans les réacteurs nucléaires[462]. En effet, ce dernier a une durée de vie limitée, si bien qu'il faut le changer à intervalles réguliers. Le combustible communément utilisé dans les centrales nucléaires en France est constitué de pastilles de dioxyde d'uranium. Ces pastilles sont empilées dans des tubes en alliage de zirconium d'environ quatre mètres de longueur, aussi appelés « gaines ». L'ensemble pastilles-gaine constitue un « crayon »[463]. Plusieurs dizaines de crayons

[461] K. Chapi, V.P. Singh, A. Shirzadi, *et al.*, « A novel hybrid artificial intelligence approach for flood susceptibility assessment », *Environmental Modelling & Software*, 2017, vol. 95, pp. 229–245 ; J.M. Corchado, J. Aiken, « Hybrid artificial intelligence methods in oceanographic forecast models », *IEEE Transactions on Systems, Man, and Cybernetics*, novembre 2022, vol. 32, n°4, pp. 307–313.

[462] H.G. Kim, S.H. Chang, B.H. Lee, « Optimal fuel loading pattern design using an artificial neural network and a fuzzy rule-based system », *Nuclear Science Engineering*, 1993, vol. 115, pp. 152-163 ; A. Galperin, E. Nissan, « Application of a heuristic search method for generation of fuel reload configurations », *Nuclear Science Engineering*, 1988, vol. 99, pp. 343-352. Notons que d'autres exemples auraient pu être pris dans le domaine nucléaire (notamment sur la disposition du combustible usagé lors de son enfouissement en couche géologique : V. Solans, D. Rochman, C. Brazell, *et al.*, « Optimisation of used nuclear fuel canister loading using a neural network and genetic algorithm », *Neural Computer & Applications*, 2021, vol. 33, 16627–16639.

[463] « Combustible nucléaire », *Wikipédia*, disponible en ligne : <https://fr.wikipedia.org/wiki/Combustible_nucl%C3%A9aire>, consulté le 24 septembre 2022.

sont ensuite insérés dans un « assemblage combustible » qui sera plongé dans la cuve du réacteur afin de produire de l'électricité. Lors d'un rechargement de combustible, le réacteur est mis à l'arrêt pour que les exploitants puissent retirer les assemblages de combustible usagé et les remplacer par de nouveaux. La disposition des crayons à l'intérieur de l'assemblage fait l'objet de nombreux calculs et vérifications en amont car elle conditionne le bon déroulé de la réaction nucléaire ainsi que la sûreté de fonctionnement du réacteur lorsqu'il sera redémarré. Le principal problème dans la détermination de la disposition des crayons dans l'assemblage combustible est le grand nombre de combinaisons possibles. Plusieurs techniques ont été développées pour automatiser ces calculs, reposant pour la majorité d'entre elles sur des algorithmes déterministes[464]. D'autres reposent sur des techniques d'IA, notamment le recours à des réseaux de neurones profonds[465]. Qu'importe la méthode utilisée, le plan de rechargement résultant des calculs doit faire l'objet d'une vérification par un code informatique qualifié par l'ASN[466]. Bien que les techniques d'apprentissage automatiques présentent des résultats prometteurs dans ce contexte, leur mode de fonctionnement ne permet pas de les faire qualifier auprès de l'ASN comme cela a été démontré précédemment[467]. Ainsi, à défaut de faire évoluer le cadre règlementaire ou de parvenir à créer une méthode de certification adaptée à l'IA, une solution serait de conserver la phase de vérification par un code de calcul déterministe dûment qualifié. En somme, le modèle d'apprentissage automatique serait utilisé pour produire une proposition de plan de chargement, lequel serait ensuite injecté dans un logiciel déterministe qualifié pour le valider. Procéder de telle manière permettrait d'avoir recours à des systèmes d'IA non déterministes dans des environnements critiques, sans prendre aucun risque du point de vue de la sûreté et sans avoir besoin de faire évoluer le corpus juridique existant. À défaut d'une compréhension suffisante des modèles d'IA connexionnistes et dans l'attente de pouvoir

[464] Pour une liste des techniques algorithmiques employées pour l'optimisation du plan de rechargement combustible des réacteurs nucléaires, voir : E. Faria, C. Pereira, « Nuclear fuel loading pattern optimisation using a neural network », *Annals of Nuclear Energy*, 2003, vol. 30, pp. 603–613.
[465] *Ibid.*
[466] Certaines méthodes requièrent même le recours à certains codes informatiques qualifiés afin de calculer des paramètres du problème à résoudre : voir notamm. E. Faria, C. Pereira, *op. cit.* : *« The main idea is to train the ANN with information on the performance of a group of spatial configurations for the fuel assemblies, using them later for the generation of new configurations that will be evaluated according to the performance of the parameters obtained through simulation with the WIMS and CITATION codes »*.
[467] Voir Supra, 190-200.

les certifier avec autant de certitude que des logiciels déterministes, l'hybridation semble être une bonne solution de transition.

228. **Conclusion du Chapitre 1 relatif à l'incompatibilité manifeste entre l'IA et les contraintes de sécurité dans les systèmes critiques.** Les contraintes de sécurité sont fortes dans le secteur de l'électricité et l'IA s'en accommode mal. En effet, les régimes juridiques applicables aux OIV, dont font partie notamment les gestionnaires des réseaux de transport et de distribution, et aux exploitants de centrales nucléaires rendent très difficile l'utilisation de l'IA dans ces activités. Ce cadre apparaît disproportionné au vu des potentiels bénéfices que pourraient apporter de tels systèmes dans les infrastructures critiques de l'énergie, que ce soit en termes de sécurité ou de performance. Le présent Chapitre a permis de défendre l'idée qu'une discussion collégiale était nécessaire afin d'étudier la possibilité d'assouplir la règlementation pour permettre le recours à l'IA. Ce débat, nécessairement multipartite et pluridisciplinaire, doit réunir les parties prenantes que sont les opérateurs d'infrastructures critiques dans l'énergie, les autorités de régulation sectorielles, des experts indépendants ainsi que les citoyens pour se prononcer sur l'acceptabilité des solutions envisagées.

229. **Rappel des propositions.** Concernant les potentielles solutions, plusieurs propositions ont été développées, allant du recours à des mécanismes juridiques favorables à l'innovation (bacs à sable règlementaire…) à la reconnaissance de nouveaux standards de sécurité adaptés aux systèmes d'IA en passant par la promotion des approches hybrides, combinant la transparence des logiciels déterministes et les performances des techniques d'IA.

230. **Transition.** Toutefois, quand bien même les contraintes relatives à la sécurité venaient à être surmontées et qu'une utilisation sûre de l'IA dans les infrastructures critiques de l'énergie était possible, la conception de ces systèmes nécessitera toujours la mobilisation d'importantes quantités de données. La collecte et l'utilisation de ces dernières apparaissent pourtant encore trop contraintes par d'autres règlementations sectorielles, contrairement à l'objectif affiché d'ouverture des données du secteur de l'énergie.

Chapitre 2 : Une insuffisante ouverture des données dans le secteur de l'électricité

231. **L'importance de la disponibilité des données pour le développement de l'IA.** La donnée est au cœur de la conception et du fonctionnement des systèmes d'IA. L'essor de leur développement depuis les années 90 est dû, notamment, à trois facteurs : l'augmentation des capacités de calcul disponibles, l'apparition de nouvelles techniques d'apprentissage automatique sur des données et la multiplication du nombre de données mobilisables. En effet, la quantité de données produites par des objets connectés, générées par l'utilisation de services en ligne ou collectées par des compteurs communicants ne cesse de croître chaque année[468]. Elles constituent la matière première requise pour le développement de systèmes d'IA.

232. **La production de données dans le secteur de l'électricité.** Le secteur objet de notre étude n'échappe pas à l'explosion du volume de données. Chaque jour, ses acteurs – gestionnaires de réseaux, producteurs, fournisseurs d'électricité et de services associés – en génèrent ou récoltent de grandes quantités. Les clients finals participent également à cette production en tant qu'utilisateurs d'équipements connectés comme des compteurs communicants, de chaudières intelligentes ou même de bornes de recharge de véhicule électrique. Les données ainsi produites ou collectées sont très variées par leur nature ainsi que par les régimes juridiques qui s'y appliquent.

233. **Un sujet saisi par l'autorité de régulation sectorielle.** À cet égard, la Commission de régulation de l'énergie (CRE) a publié, le 18 mai 2017, un rapport relatif aux données dont disposent les gestionnaires de réseaux et d'infrastructures d'énergie[469]. Il établit, de façon inédite, une typologie tant juridique que technique des données dont disposent les différents opérateurs des infrastructures du réseau électrique, laquelle révèle leur importante volumétrie, leur complexité et leur hétérogénéité[470].

---

[468] Sur l'évolution de la quantité de données produites dans le monde depuis 2010, voir STATISTA RESEARCH DEPARTMENT, *Amount of data created, consumed, and stored 2010-2025, Statista Research Department*, site officiel, 18 mars 2022, disponible en ligne : <https://www.statista.com/statistics/871513/worldwide-data-created/>, consulté le 25 avril 2022.
[469] COMMISSION DE RÉGULATION DE L'ÉNERGIE (CRE), *Rapport du comité d'études relatif aux données dont disposent les gestionnaires de réseaux et d'infrastructures d'énergie*, rapport, 18 mai 2017, 116 p.
[470] COMMISSION DE RÉGULATION DE L'ÉNERGIE (CRE), *op. cit.*, p. 11 et s.

234. **Définition et typologie des données énergétiques.** Les données énergétiques – entendues ici comme l'ensemble des données produites ou collectées dans le cadre de l'une des activités composant le secteur de l'électricité[471] – englobent ainsi des données de consommation et de production, des données patrimoniales (relatives aux ouvrages constituant les réseaux), des données d'exploitation des réseaux et des infrastructures (mesure des flux d'énergie au niveau des ouvrages), des données de mesure de la qualité d'alimentation ou bien encore des données environnementales. Leur exploitation peut avoir de nombreux bénéfices, que ce soit pour le bon fonctionnement du marché de l'électricité *via* la mise en place d'offres individualisées[472], pour le développement des territoires en permettant aux collectivités de disposer des informations nécessaires à la conduite des politiques publiques[473], pour l'appropriation de la consommation d'énergie par les clients finals[474], ou encore pour garantir la transparence de l'action des opérateurs de services publics[475]. Toutefois, l'exploitation des données énergétiques peut aussi favoriser l'innovation et la conception de nouveaux services numériques[476], le cas-échéant fondés sur de l'IA. Le développement d'applications de maintenance prédictive pour les réseaux nécessite par exemple de disposer de nombreuses données relatives au fonctionnement des infrastructures et à la survenance de dysfonctionnements. Si ces données étaient exclusivement détenues par les gestionnaires de réseaux, alors aucun acteur tiers ne pourrait concevoir de telles applications. De la même manière, tous les systèmes d'IA pouvant être utilisés dans le cadre de la modernisation des réseaux vers le modèle du réseau bidirectionnel dit « intelligent » requièrent le traitement en temps réel des données de consommation[477]. En outre, des solutions d'analyse des consommations individuelles pourraient fournir aux clients finals des recommandations afin de

[471] La production, le transport, la distribution, la fourniture d'électricité et les services associés : voir Supra, 26-30.
[472] COMMISSION DE RÉGULATION DE L'ÉNERGIE (CRE), *op. cit.*, p. 60.
[473] *Ibid.*, p. 62.
[474] *Ibid.*, p. 65.
[475] *Ibid.*, p. 70.
[476] *Ibid.*, p. 66.
[477] INSTITUT MONTAIGNE, *Transition énergétique : faisons jouer nos réseaux*, rapport, 2019, disponible en ligne : <https://www.institutmontaigne.org/ressources/pdfs/publications/transition-energetique-faisons-jouer-nos-reseaux-rapport.pdf>, consulté le 25 avril 2022 ; P. Giannakaris, P. Trakadas, T. Zahariadis, *et al.*, « Using Smart Contracts in Smart Energy Grid Applications », *Proceedings of the International Scientific Conference*, Novi Sad, Serbia, Singidunum University, 2019, 597–602.

réduire leur facture et, ainsi, de contribuer aux efforts de sobriété énergétique[478]. Ainsi, pour promouvoir le développement des systèmes d'IA dans le secteur de l'électricité, il est nécessaire que les entreprises puissent avoir accès à des données de qualité et en quantité suffisante[479].

235. **La tendance à l'ouverture progressive des données énergétiques.** À ce stade, il convient de noter que, selon leur nature et le régime juridique qui leur est applicable, certaines données énergétiques font l'objet d'obligations légales de mise à disposition des clients finals, de leurs mandataires ou des fournisseurs[480]. Elles peuvent également être agrégées, géographiquement ou temporellement, et être mises à disposition des personnes publiques ou du public dès lors qu'elles sont libres de droit. S'est ainsi structuré, au fil des décennies, un véritable *open data* législatif des données énergétiques, concrétisé par la Loi pour une République numérique du 7 octobre 2016[481]. Ce contexte législatif est plutôt favorable à notre problématique puisqu'il permet de rendre disponibles d'importantes quantités de données. Encore faut-il que ces dernières soient de qualité et qu'elles soient mises à disposition des acteurs les plus à mêmes de concevoir des solutions innovantes et utiles dans le secteur de l'électricité.

236. **La tendance paradoxale de la protection de la confidentialité des données énergétiques.** En parallèle de la tendance législative à l'ouverture des données, certaines dispositions législatives ou règlementaires visent, dans une logique totalement inverse, à protéger la confidentialité de certaines données. À ce titre, le Code de l'énergie prévient la diffusion des données constituant des « informations commercialement sensibles », susceptibles d'affecter la concurrence si elles étaient connues par les acteurs du marché[482]. De plus, le cadre juridique relatif à la protection des données à caractère personnel s'applique à certaines données énergétiques dès lors qu'elles peuvent révéler *« une information se rapportant à une personne physique identifiée ou identifiable »*[483]. C'est le cas, par exemple,

[478] COMMISSION DE RÉGULATION DE L'ÉNERGIE (CRE), *op. cit.*, p. 68.
[479] F. Choné, « L'énergéticien du XXIe siècle : le numérique au service du consommateur et de la transition énergétique », *Annales des Mines - Responsabilité et environnement*, 2017, vol. 87, n°3, 43.
[480] C. Boiteau, « L'entreprise régulée », *RFDA*, 2018, 469.
[481] *Loi n° 2016-1321 du 7 octobre 2016 pour une République* numérique, publiée au JORF n°0235 du 8 octobre 2016, texte n° 1.
[482] Code de l'énergie, article L111-72 et s.
[483] RGPD, article 4.

des historiques de consommation électrique, appelés « courbe de charge ». Leur traitement – et *a fortiori* leur partage – ne peut donc se faire au mépris des principes du RGPD[484] ou de la loi Informatique et libertés[485] : licéité, loyauté et transparence ; limitation des finalités et de la durée de conservation ; exactitude ; intégrité et confidentialité. Ces dispositions garantissent un premier niveau de protection pour les individus dont les données de consommation sont collectées. Toutefois, ces dernières bénéficient en plus d'un statut particulier et d'une protection accrue par rapport aux autres données à caractère personnel. En effet, à l'occasion du déploiement des compteurs communicants en France, la CNIL a été conduite à s'interroger sur la nature des données de consommation des clients. Dans une délibération du 15 novembre 2012, l'autorité indépendante est venue poser des conditions particulièrement précises s'agissant des modalités de collecte de la courbe de charge et du consentement des personnes concernées[486]. Ce régime particulier est justifié par les informations intimes, relatives à la vie privée des individus, qu'il est possible de déduire de l'analyse de la courbe de charge. À titre d'exemple, la CNIL considère qu'une courbe de charge avec un pas de temps de 10 minutes permettrait d'identifier les heures de lever et de coucher, les heures ou périodes d'absence et d'autres indications touchant la personne occupant le logement[487]. Ces contraintes, bien que justifiées, rendent difficile l'exploitation des données issues des compteurs communicants qui pourraient pourtant permettre le développement de solutions intelligentes de performance énergétique (pour recommander aux consommateurs des actions pour réduire leur consommation, identifier des passoires thermiques ou mieux prédire la consommation des foyers notamment). Enfin, certains droits de propriété intellectuelle peuvent venir faire obstacle à la diffusion des données, comme le droit *sui generis* sur les bases de données produites par des gestionnaires de réseaux ou encore le régime légal du secret des affaires[488].

Sur l'articulation des régimes de protection des données à caractère personnel avec l'open data, voir notamm. : L. Maisnier-Boche, N. Botchorichvili, « Projet de loi « pour une République numérique » : quels impacts sur la protection des données personnelles ? », *RLDI*, janvier 2016, n° 3911.

[484] RGPD, article 5.

[485] *Loi n° 78-17 du 6 janvier 1978 relative à l'informatique, aux fichiers et aux libertés, dit « Loi Informatique et Libertés » ou « LIL »*, publiée au JORF du 7 janvier 1978, article 4.

[486] CNIL, *Délibération n° 2012-404 du 15 novembre 2012 portant recommandation relative au traitement des données de consommation détaillées collectées par les compteurs communicants*.

[487] *Ibid.*

[488] O. Beatrix, « Open data et secteur de l'énergie : le début de l'histoire », *RFDA*, 2018, 49.

237. **L'opposition entre les deux tendances.** On assiste ainsi à un double mouvement : ouverture des données d'une part, protection de la confidentialité d'autre part. Au regard de notre sujet, ces deux tendances sont parfaitement justifiées et souhaitables puisque, on l'a vu, il est nécessaire de rendre disponibles plus de données pour stimuler le développement de l'IA, tout en protégeant les droits et libertés des individus. La situation aboutit toutefois à un casse-tête pour les détenteurs de ces données, un *« exercice schizophrénique »* d'après certains acteurs[489], *« consistant à devoir assurer en même temps la protection des données et une ouverture de plus en plus large de ces données »*[490]. L'opposition entre ces deux mouvements n'est pas que théorique et se concrétise dans la pratique par des situations où des détenteurs de données refusent systématiquement de partager leurs données en invoquant les régimes de protection des données ou la protection du savoir-faire, ou encore par des situations où des fournisseurs cherchent par tous les moyens à recueillir le consentement des consommateurs[491]. Lorsque les dynamiques d'ouverture et de confidentialité entrent en confrontation, la question se pose de savoir laquelle doit primer : Faut-il aller plus loin dans l'ouverture des données énergétiques pour favoriser le développement de l'IA dans le secteur de l'électricité ou faut-il au contraire renforcer les garde-fous pour protéger les intérêts des individus et la sécurité des infrastructures dans un monde toujours plus numérisé et sujet aux attaques informatiques ?

238. **Une conciliation nécessaire.** L'ouverture des données énergétiques ne doit pas, selon nous, se faire au détriment de la protection des droits et libertés des individus. Certaines contraintes apparaissent toutefois disproportionnées et pourraient être allégées pour favoriser le développement de l'IA dans le secteur de l'électricité. Le présent Chapitre tentera d'apporter une réflexion *a coherentia*[492], visant à renforcer la cohérence des régimes applicables aux données énergétiques dans l'objectif susmentionné de conciliation entre innovation et sécurité des données.

---

[489] COMMISSION DE RÉGULATION DE L'ÉNERGIE (CRE), *op. cit.*, p. 14.
[490] *Ibid.*
[491] Parfois aux frontières de la légalité, voir notamment les faits constatés dans la mise en demeure CNIL, *Décision MED 2019-035 du 31 décembre 2019*.
[492] G. Tarello, « Sur la spécificité du raisonnement juridique », *Archives de philosophie du droit et de philosophie sociale*, 1972, n°7, p. 105.

239. **Plan.** Pour ce faire, il convient de noter que l'ouverture des données est une tendance bien établie (**Section 1**), un objectif ancré dans les politiques française et européenne relatives au secteur de l'énergie. Bien qu'elle soit établie, cette tendance est néanmoins contrainte (**Section 2**) et doit être amplifiée, ce qui requiert de lever les différents obstacles à sa mise en œuvre, tout en conservant un niveau élevé de protection des intérêts des individus.

Section 1 : L'ouverture des données énergétiques : une tendance établie

240. **L'*open data* de l'énergie : un objectif assumé.** L'ouverture des données dans le secteur de l'électricité est l'un des objectifs des politiques publiques en matière d'électricité que ce soit au niveau national ou européen. En effet, elle permettrait notamment aux consommateurs de mieux appréhender et agir sur leur consommation d'électricité, aux collectivités territoriales de mieux cibler les politiques publiques, ou encore aux opérateurs de service public de bénéficier des meilleures solutions techniques pour optimiser les coûts liés à la maintenance des infrastructures. L'ouverture des données est également un vecteur d'innovation. Par exemple, le recours croissant aux énergies renouvelables nécessite d'en prévoir plus finement la production, ce que permettent certains systèmes d'IA conçus à partir de l'analyse de nombreuses données relatives à l'état des installations, leur exposition, la météo, ou encore aux éventuelles opérations de maintenance planifiées. En outre, le développement de systèmes d'aide à la décision permettant aux autorités publiques de cibler les bâtiments nécessitant une rénovation énergétique requiert l'analyse de grandes quantités de données relatives auxdits bâtiments : année de construction, évolution de la consommation d'électricité, état du réseau électrique aux alentours. Enfin, d'autres nouveaux services innovants, notamment de pilotage de la consommation, pourraient être proposés aux clients finals si leurs données de consommation étaient plus facilement accessibles.

241. **Plan.** Ces constats ont motivé une politique volontariste à l'échelle européenne (**§1**), visant au déploiement massif de systèmes de comptage intelligents et à la facilitation de l'accès à certaines données énergétiques. Il s'agit là des premières pierres à l'édifice de l'ouverture des données, indispensables à qui souhaite encourager l'innovation dans le secteur de l'électricité. Cette impulsion européenne a par la suite été concrétisée en droit français (**§2**), qui organise un véritable *open data* des données énergétiques.

§1 : L'ambitieuse impulsion européenne pour la production de données dans le secteur de l'électricité

242. **La promotion bienvenue de la production de données dans le secteur de l'électricité par le déploiement des compteurs communicants.** Afin que des entreprises développent des systèmes d'IA souhaitables dans le secteur de l'électricité, elles doivent avoir accès à des données en quantité et de bonne qualité. La politique de l'Union européenne en matière d'énergie va dans ce sens. En effet, elle promeut le recours aux « systèmes intelligents de mesure » depuis la directive 2009/72/CE du 13 juillet 2009[493]. Plus récemment, la directive (UE) 2019/944 du 5 juin 2019 définit un tel système – aussi appelé communément « compteur communicant » – comme *« un système électronique qui est capable de mesurer l'électricité injectée dans le réseau ou l'électricité consommée depuis le réseau en fournissant davantage d'informations qu'un compteur classique, et qui est capable de transmettre et de recevoir des données à des fins d'information, de surveillance et de contrôle en utilisant une forme de communication électronique »*[494]. D'après les considérants de la directive de 2019, les compteurs communicants ont deux fonctions principales. D'une part, ils autonomisent les individus en leur permettant d'être informés de manière précise et en temps quasi réel sur leur consommation ou production d'énergie, de participer aux programmes de participation active de la demande, de bénéficier d'autres services de performance énergétique, ou de réduire leurs factures d'électricité. D'autre part, les compteurs communicants permettent aux gestionnaires de réseau d'avoir une meilleure vision de leurs réseaux et, par conséquent, de réduire leurs dépenses d'exploitation et de maintenance[495]. Les bénéfices attendus sont tels que les États membres sont enjoints à déployer ces systèmes de mesure auprès de 80% des clients finals d'ici 2024[496]. En plus de la charge du déploiement des compteurs communicants, les gestionnaires

[493] *Directive (UE) 2009/72/CE du Parlement européen et du Conseil du 13 juillet 2009 concernant des règles communes pour le marché intérieur de l'électricité et abrogeant la directive 2003/54/CE,* publiée au JOUE n°L211/55 du 14 août 2009.
[494] *Directive (UE) 2019/944 du Parlement européen et du Conseil du 5 juin 2019 concernant des règles communes pour le marché intérieur de l'électricité et modifiant la directive 2012/27/UE*, publiée au JOUE L158/125 du 14 juin 2019, article 2, 23).
[495] *Ibid.*, considérant 52 ; K. Huhta, « Smartening up while keeping safe? Advances in smart metering and data protection under EU law », *Journal of Energy & Natural Resources Law*, 2020, vol. 38, n°1, 5–22 ; S. Lavrijssen, A. Carrillo Parra, « Radical prosumer innovations in the electricity sector and the impact on prosumer regulation », *Sustainability*, 2017, vol. 9, 1207.
[496] *Ibid.*, annexe II.

de réseau de distribution ont l'obligation de rendre les données collectées aux clients finals ainsi qu'aux fournisseurs d'électricité, notamment aux fins de facturation. L'ensemble de ces dispositions ont par la suite été transposées en droit français, dans le Code de l'énergie.

243. **Des applications de l'IA rendues possibles par le déploiement des compteurs communicants.** L'exploitation des données collectées par les systèmes intelligents de mesure va permettre le développement d'innovations indispensables aux mutations du secteur de l'électricité dans un contexte de transition écologique. En particulier des systèmes d'IA conçus et fonctionnant à partir de ces données peuvent contribuer au développement des réseaux intelligents, aux mécanismes de participation active des clients finals à l'équilibrage du réseau, ainsi qu'à la création de plateformes d'achat et de vente d'énergie en pair à pair[497].

244. **L'utilité de l'IA dans la modernisation des réseaux électriques.** D'abord, les compteurs communicants jouent un rôle majeur dans l'établissement d'un réseau intelligent – bidirectionnel et au fonctionnement optimisé grâce aux technologies numériques – en ce qu'ils fournissent aux différents acteurs, et en particulier les gestionnaires de réseau, des informations détaillées sur la quantité d'électricité injectée ou soutirée[498]. La production locale d'électricité d'origine renouvelable donne une nouvelle dimension au comptage intelligent. En effet, les ménages peuvent générer leur propre énergie, par exemple en installant des panneaux photovoltaïques sur leur toit. Dans ce contexte, les compteurs communicants peuvent mesurer la quantité d'énergie produite, produisant ainsi des données que peut utiliser le gestionnaire du réseau dans sa mission d'équilibrage production-consommation. L'exploitation de ces données peut s'avérer complexe en raison de leur nombre et de l'imprévisibilité de la production des énergies renouvelables. En parallèle, les systèmes d'IA sont particulièrement efficaces pour traiter de grandes quantités de données d'une part, et pour réaliser des prédictions à partir de multiples paramètres, d'autre part. Ainsi, le recours à des systèmes d'IA dans l'énergie va vraisemblablement être accéléré par le développement des compteurs et réseaux intelligents.

[497] S. Lavrijssen, B. Espinosa Apráez, T. ten Caten, « The Legal Complexities of Processing and Protecting Personal Data in the Electricity Sector », *Energies*, 2022, vol. 15, n°3, 1088.
[498] S.F. Bush, *Smart Grid: Communication-Enabled Intelligence for the Electric Power Grid*, John Wiley & Sons Ltd, West Sussex, UK, 2014, pp. 3–183.

245. **L'utilité de l'IA dans l'équilibrage de la production et de la consommation d'électricité.** Ensuite, si une personne physique ou morale est à la fois consommateur et producteur et qu'il produit plus d'électricité qu'il n'en consomme, alors il doit pouvoir la réinjecter sur le réseau en échange d'une compensation : c'est ce que l'on appelle la « participation active » des clients finals, ayant donné lieu au terme « *prosumer* » en anglais, mélange de « *producer* » et de « *consumer* »[499]. La traduction juridique de ce concept peut se trouver dans la directive européenne du 5 juin 2019 sous la dénomination de « client actif », défini comme *« un client final, ou un groupe de clients finals agissant conjointement, […] qui vend l'électricité qu'il a lui-même produite ou participe à des programmes de flexibilité ou d'efficacité énergétique, à condition que ces activités ne constituent pas son activité commerciale ou professionnelle principale »*[500]. De la même manière, les responsables d'équilibre doivent être en mesure d'anticiper le surplus d'autoproduction réinjectée sur le réseau afin d'adapter en conséquence la production et éviter des risques de surtension. Le recours à des systèmes d'IA permet d'optimiser cette gestion complexe.

246. **L'utilité de l'IA dans les transactions de pair-à-pair entre producteurs et consommateurs.** Enfin, lorsque les clients actifs décident de ne pas fournir leur surplus de production au marché mais directement à d'autres consommateurs, ils s'adonnent à des échanges en pair-à-pair. Ces échanges peuvent se dérouler sur des plateformes fondées sur la technologie de chaînes de blocs, dispensant de l'intervention de tiers dans les transactions (les fournisseurs d'énergie) [501]. La directive européenne en date du 11 décembre 2018 relative à la promotion de l'utilisation de l'énergie produite à partir de sources renouvelables est venue apporter une définition juridique à l'échange de pair-à-pair, devant être entendu comme *« la vente d'énergie renouvelable entre participants au marché sur la base d'un contrat contenant des conditions préétablies régissant l'exécution et le règlement automatiques de la transaction soit directement entre les participants au marché, soit indirectement par l'intermédiaire d'un*

[499] S. Lavrijssen, A. Carrillo Parra, « Radical prosumer innovations in the electricity sector and the impact on prosumer regulation », *Sustainability*, 2017, vol. 9, 1207.
[500] *Directive (UE) 2019/944 du Parlement européen et du Conseil du 5 juin 2019 concernant des règles communes pour le marché intérieur de l'électricité et modifiant la directive 2012/27/UE*, publiée au JOUE n°L158/125 du 14 juin 2019, article 2, 8).
[501] C. Zhang, J. Wu, C. Long, *et al.*, « Review of Existing Peer-to-Peer Energy Trading Projects », *Elsevier Energy Procedia*, 2017, vol. 105, 2563–2568.

participant au marché tiers certifié, par exemple un agrégateur »[502]. Le cas-échéant, l'IA peut permettre d'automatiser ces transactions ou d'intégrer des mécanismes de négoce avantageux pour les clients actifs, par exemple en déclenchant automatiquement une vente sur le marché classique lorsque la demande locale est insuffisante, ou au contraire en privilégiant la vente locale lorsque les prix du marché sont plus faibles. Les expérimentations en la matière sont déjà nombreuses[503].

247. **La promotion bienvenue du déploiement des compteurs communicants.** Le déploiement des compteurs communicants sous l'impulsion européenne est donc une très bonne chose pour le secteur de l'électricité puisqu'il permet le développement d'innovations indispensables à sa mutation dans un contexte de transition écologique. Ces dernières visent à faciliter le pilotage du réseau, l'intégration des énergies renouvelables, l'autoconsommation et l'échange local d'électricité. Chacune d'entre elles requiert l'utilisation de technologies numériques avancées, et notamment d'IA comme cela a été présenté dans les paragraphes précédents. Toutefois, cette « mise en données » du secteur de l'électricité ne peut se faire au détriment des clients finals qui doivent rester maîtres de leur consommation. C'est la raison pour laquelle l'Union européenne a choisi d'accompagner cette dynamique par la garantie, pour les clients finals, d'un accès facilité aux données produites grâce aux compteurs communicants.

248. **L'accès facilité aux données des compteurs communicants pour les clients finals.** À en croire la récente directive européenne du 5 juin 2019, l'Union européenne accorde beaucoup d'importance à la facilité d'accès aux données énergétiques par les consommateurs eux-mêmes. Si bien qu'elle est présentée comme la première finalité du déploiement des compteurs communicants dans son article 20 : *« les systèmes intelligents de mesure ont pour fonction de mesurer avec précision la consommation réelle d'électricité et sont capables de fournir aux clients finals des informations sur le moment réel où l'énergie a été utilisée. Les clients finals doivent pouvoir accéder facilement aux données validées relatives à l'historique de*

[502] *Directive (UE) 2018/2001 du Parlement européen et du Conseil du 11 décembre 2018 relative à la promotion de l'utilisation de l'énergie produite à partir de sources renouvelables,* publiée au JOUE n°L328/82 du 21 décembre 2018, article 2, 18).

[503] T. Cortade, J.-C. Poudou, « Les plateformes numériques d'échange d'électricité », *Annales des Mines – Enjeux numériques*, septembre 2021, n°15 ; « Blockchain dans le domaine de l'énergie : où en est-on ? », *ThinkSmartgrids (blog)*, 7 mars 2019, disponible en ligne : <https://www.thinksmartgrids.fr/actualites/blockchain-domaine-energie>, consulté le 27 avril 2022.

consommation et les visualiser facilement, de manière sécurisée, sur demande et sans frais supplémentaires. Les clients finals doivent également pouvoir accéder facilement aux données non validées relatives à la consommation en temps quasi réel et de manière sécurisée, sans frais supplémentaires, via une interface normalisée ou via un accès à distance, afin de favoriser les programmes automatisés d'amélioration de l'efficacité énergétique, la participation active de la demande et d'autres services »[504]. Les clients finals peuvent également demander la communication de ces données à un tiers agissant en leur nom[505]. En permettant aux clients finals d'avoir accès aux données de comptage sans coût additionnel[506], le cadre européen souhaite encourager le développement de services de gestion de l'énergie, de performance énergétique ou encore de conseil en économie d'énergie[507]. À ce titre, il facilite également la conception de systèmes d'IA poursuivant ces objectifs. Néanmoins, la multiplication du nombre de données produites et de leurs transferts participe à la création de nouveaux risques pour les données des clients, qui peuvent révéler des informations sur leur vie privée. À cela, le régime européen tente de répondre par des dispositions générales obligeant notamment les États membres à prévoir la mise en place des *« meilleures techniques disponibles pour garantir le plus haut niveau de protection en matière de cybersécurité, tout en gardant à l'esprit les coûts et le principe de proportionnalité »*[508] et à respecter *« les règles de l'Union applicables en matière de protection des données et de respect de la vie privée »*[509].

249. **Transition.** La démarche européenne est louable et a permis le développement d'un cadre favorable au développement de l'IA dans le secteur de l'électricité en organisant le déploiement de compteurs communicants et la mise à disposition des données produites, tout en protégeant leur sécurité et la vie privée des individus. Toutefois, ce cadre ne concerne que les données de consommation. Or, il existe également d'autres données énergétiques (patrimoniales, environnementales…) dont l'ouverture pourrait être utile. C'est à ce constat que tente de

[504] *Directive (UE) 2019/944 du Parlement européen et du Conseil du 5 juin 2019 concernant des règles communes pour le marché intérieur de l'électricité et modifiant la directive 2012/27/UE*, publiée au JOUE n°L158/125 du 14 juin 2019, article 20, a).
[505] *Ibid.*, article 20, e).
[506] *Ibid.*, article 23, 5.
[507] *Ibid.*, article 19, 1).
[508] *Ibid.*, article 20, b).
[509] *Ibid.*, article 20, c).

répondre le régime français, allant plus loin que le cadre européen en organisant un véritable *open data* des données énergétiques.

§2 : La concrétisation française de l'ouverture des données énergétiques

250. **Plan.** L'*open data* relève, en France, d'un processus législatif sophistiqué visant à promouvoir la mise à disposition et la libre réutilisation gratuite des données publiques. Le concept relève d'une ambition plus grande que les dispositions portées par les directives européennes dans le secteur de l'électricité. En matière d'*open data*, la France se classe même parmi les pays européens les plus avancés[510]. Les données énergétiques n'échappent pas à ce constat, si bien que s'est développé un véritable modèle français de l'*open data* des données énergétiques (**A**), dont le point d'orgue est la promulgation de la Loi pour une République numérique du 7 octobre 2016[511]. Cette dynamique législative s'est concrétisée dans la pratique avec la mise en place de nombreuses plateformes de mise à disposition de données par les acteurs du secteur de l'électricité (**B**), lesquels partagent parfois plus que ce à quoi ils sont légalement tenus. L'ensemble de ces efforts sont positifs pour le développement de l'IA dans le secteur et doivent être poursuivis.

[510] CAPGEMINI INVENT, *Open data maturity report 2021*, rapport commandité par la Commission européenne, 7ème ed., 17 décembre 2021.
[511] *Loi n° 2016-1321 du 7 octobre 2016 pour une République numérique*, publiée au JORF n°0235 du 8 octobre 2016, texte n° 1.

A/ Le modèle français de l'*open data* des données énergétiques

251. **Plan.** L'ouverture des données énergétiques est issue en France de plusieurs lois (**1**) puis a été organisée de façon plus précise dans le Code de l'énergie (**2**).

1. Une ouverture des données énergétiques organisée par la loi

252. **Les objectifs de l'*open data*.** Les données publiques sont régulièrement considérées comme des biens communs, devant échapper aux régimes propriétaires et pouvoir être réutilisées par les entreprises pour innover[512]. L'*open data* vise à réaliser cet idéal en érigeant en principe absolu l'ouverture des données publiques. Il crée le droit, pour le public, d'accéder à l'information, de l'analyser, de la contester et de pouvoir la réutiliser librement[513]. Dans son étude annuelle de 2014 consacrée au numérique, le Conseil d'État précise les objectifs de l'ouverture des données[514]. Le premier est de nature démocratique et consiste à donner aux citoyens un droit de regard sur les politiques publiques et leurs résultats. Le second, qui nous intéresse plus particulièrement, est d'ordre économique. Le Conseil d'État considère en effet que l'exploitation des données publiques constitue désormais le support de nombreux services créateurs de nouvelles possibilités pour les opérateurs économiques.

253. **Une accélération depuis la loi pour une République numérique.** La publication des données publiques a été grandement accélérée par la loi pour une République numérique[515], promulguée le 7 octobre 2016. Elle élargit l'obligation de publication en ligne et le principe de libre réutilisation à toutes les données produites ou détenues dans le cadre d'une mission de

[512] CONSEIL NATIONAL DU NUMÉRIQUE, *Ambition numérique – pour une politique française et européenne de la transition numérique*, rapport remis au Premier Ministre, juin 2015, p. 276.
[513] C. Bouchoux, *Refonder le droit à l'information publique à l'heure du numérique : un enjeu citoyen, une opportunité stratégique*, Rapport d'information n° 589 fait au nom de la mission commune d'information du Sénat sur l'accès aux documents administratifs, 5 juin 2014.
[514] CONSEIL D'ÉTAT, *Le numérique et les droits fondamentaux*, Etude annuelle 2014, La Documentation française, 64, p. 66.
[515] *Loi n° 2016-1321 du 7 octobre 2016 pour une République numérique*, publiée au JORF n°0235 du 8 octobre 2016, texte n°1.

service public industriel et commercial, et non plus seulement par les administrations[516]. Cette obligation n'est pas sans conséquence dans le secteur de l'électricité. En effet, le champ d'application de la loi inclut de nombreux opérateurs du secteur de l'énergie puisque sont visées les données détenues par les personnes de droit public ou de droit privé chargées d'une mission de service public[517]. EDF, Enedis ou RTE sont ainsi concernés dès lors qu'ils exercent une telle mission. Les concernant, l'article 23 de la loi dispose que doivent être mises à disposition du public par voie électronique, dans un format ouvert, aisément réutilisable et exploitable par un système de traitement automatisé sous une forme agrégée garantissant leur caractère anonyme, *« les données détaillées de consommation et de production issues de leur système de comptage d'énergie »*[518]. Le texte vient donc préciser les modalités d'ouverture des données déjà engagée par la loi relative à la transition énergétique pour la croissance verte[519].

254. **La nature des données du secteur de l'électricité concernées par l'*open data*.** L'étude de la doctrine de la Commission d'accès aux documents administratifs (CADA), l'autorité administrative indépendante chargée de veiller à la liberté d'accès aux documents administratifs ainsi qu'à la réutilisation des informations publiques, donne un aperçu du type de données concernées par l'*open data* dans le secteur de l'énergie. La CADA a par exemple rendu un avis favorable à la communication par le gestionnaire des réseaux de distribution ERDF (aujourd'hui devenu Enedis) d'une copie du plan du réseau électrique à un particulier souhaitant connaître les points de raccordement possibles pour un projet de construction[520]. La Commission a estimé que ce document était communicable à toute personne qui en fait la demande, *« sauf en ce qui concerne les éléments dont la communication porterait atteinte à la sécurité publique »*, offrant ici une exception mobilisable par les gestionnaires de réseaux pour s'opposer à la diffusion de

[516] V. notamm. L. Cluzel-Métayer, « La loi pour une République numérique : l'écosystème de la donnée saisi par le droit », *AJDA*, 2017, p. 340 ; M. Bourgeois, A. Bounedjoum, « Les apports de la loi pour une République numérique en matière d'accès et de réutilisation d'informations publiques », *La Semaine Juridique Administrations et Collectivités territoriales*, 5 décembre 2016, n° 48, 2307.
[517] V. les articles 17 et 18 de la loi pour une République numérique.
[518] Loi pour une République numérique, article 23, complétant l'article L111-73 du Code de l'énergie.
[519] *Loi n° 2015-992 du 17 août 2015 relative à la transition énergétique pour la croissance verte*, publiée au JORF n°0189 du 18 août 2015, texte n° 1, ayant créé les articles L111-72 et suivants du Code de l'énergie, étudiés ci-après.
[520] CADA, avis n°20141099, séance du 10/04/2014, *Copie du plan du réseau électrique afin de connaître les points de raccordement au réseau possibles pour le projet de construction de sa cliente sur le territoire de la commune de La Croix-Valmer*.

certaines données. Dans un avis du 3 novembre 2016, la CADA s'est opposée à la communication des résultats de l'outil « Apogée » utilisé par EDF dans le cadre de l'optimisation de sa production d'électricité[521]. Cet outil vise à l'optimisation financière des moyens de production d'électricité, puisque permettant à l'opérateur, en présence d'un besoin de fourniture d'électricité à destination de ses clients, de déterminer s'il est économiquement préférable de produire davantage d'électricité pour répondre à cette demande ou d'acheter de l'électricité sur les marchés. La Commission considère que les données sollicitées présentent bien le caractère de documents administratifs au sens de l'article L300-2 du Code des relations entre le public et l'administration (CRPA) en ce qu'elles retracent l'exécution par EDF de sa mission de service public. Toutefois, elle a estimé que leur communication *« révèlerait la stratégie financière mise en œuvre par EDF pour assurer l'équilibre économique de la mission de service public dont elle est chargée et est, par suite, de nature à porter atteinte au secret en matière industrielle et commerciale, protégé par les dispositions de l'article L311-6 du code des relations entre le public et l'administration »*. Une nouvelle fois, malgré la qualification de documents administratifs, l'obligation d'*open data* ne peut être rendue effective en raison de la sensibilité des informations concernées. Enfin, un dernier exemple de tentative avortée de communication de données issues du secteur de l'électricité concerne les bases de données relatives aux clients soumis au tarif règlementé de vente (TRV), issus du monopole historique d'EDF[522]. Dans son avis du 9 juin 2016, la CADA a émis un avis défavorable à la communication de cette base de données pour deux motifs. Le premier relève de la présence de données à caractère personnel dans les données concernées, le second des secrets industriels et commerciaux qui pourraient être déduits des données, en particulier relatives à la stratégie financière d'EDF. La Commission s'est ainsi opposée à la communication de ces bases en considérant que *« l'ampleur des données qu'il convient d'occulter […] priverait d'intérêt la communication »*. Ces exemples nous amènent à deux constats. Premièrement, on constate que

[521] CADA, avis n°20164200, séance du 03/11/2016, *Copie, par voie électronique, ou par dépôt électronique sur un serveur, des résultats des modèles journaliers de l'outil « APOGEE » pour la période couvrant le 1er janvier 2010 au 15 juillet 2016.*

[522] CADA, avis 20161147, séance du 09/06/2016, *Communication des tableaux statistiques (de type sondage au 100ème, répartition par centile/décile des clients TRV (tarifs réglementés de vente), base de données des clients TRV « à température normale », etc.), transmis à la Commission de régulation de l'énergie (CRE) depuis le 1er janvier 2010, dans le cadre des discussions relatives aux TRV en niveau et en structure.*

les données pouvant être recherchées par la voie de l'*open data* des documents administratifs organisé par le CRPA relèvent avant tout d'informations générales relatives à la gestion des entreprises opérant des missions de service public : plan du réseau, informations financières relatives au fonctionnement des entreprises ou à leur base de clients. Ces données ne permettraient pas le développement de systèmes d'IA utiles pour répondre aux grands enjeux du secteur de l'énergie. La conception de systèmes d'IA vertueux nécessiterait plutôt l'analyse de données de consommation ou de données patrimoniales relatives aux moyens de production ou à l'état du réseau par exemple, qui ne sont pas des informations pouvant être qualifiées de « publiques ». Deuxièmement, l'ouverture des données dans le secteur de l'énergie est confrontée à de nombreuses contraintes légitimes : protection de la vie privée, des secrets industriels et commerciaux ou encore de la sécurité publique. L'occultation des données sensibles à ces égards peut priver d'intérêt la mise à disposition des informations demandées par le public. Jusqu'à maintenant, les justifications de l'ouverture des données relèvent principalement de motifs de transparence de l'action publique et de préservation de la concurrence. Ces constats nous poussent à croire qu'un régime spécifique à l'énergie est nécessaire.

255. **Des difficultés pratiques à la mise en œuvre de l'ouverture des données dans le secteur de l'énergie.** À cet égard, Olivier Beatrix, ancien directeur juridique de Gaz Réseau Distribution France (GRDF), considère que les trois caractéristiques de l'*open data* (facilité d'accès aux données, liberté de réutilisation et interopérabilité des formats) dessinent les contours d'un nouveau droit objectif dont l'insertion dans le droit sectoriel de l'énergie n'est pas sans difficultés[523]. De nombreuses dispositions du Code de l'énergie tentent d'organiser un *open energy data*, mettant à disposition du public une grande quantité de données pouvant être mobilisées pour créer des services innovants, y compris des systèmes d'IA. Si l'intention est louable, les effets de ces dispositions sur l'innovation restent encore à prouver.

[523] O. Beatrix, « Open data et secteur de l'énergie : le début de l'histoire », *RFDA*, 2018, p. 49.

2. Une ouverture des données énergétiques organisée par le Code de l'énergie

256. **Une double obligation de partage des données.** La législation française contient deux dispositions imposant la mise à disposition des données détenues par les opérateurs du réseau électrique.

257. **Une mise à disposition des données énergétiques auprès des personnes publiques.** Premièrement, les articles L111-72 et L111-73 du Code de l'énergie, issus de la loi du 17 août 2015 relative à la transition énergétique pour la croissance verte[524], imposent aux gestionnaires des réseaux de transport et de distribution de mettre les données issues de leurs systèmes de comptage à disposition des personnes publiques[525]. Cette obligation concerne les données de transport, de consommation et de production, dont la liste a été dressée par décret[526] en 2016 et aujourd'hui codifiée aux articles D111-53 et suivants du Code de l'énergie. Sont visées les données *« utiles à l'accomplissement des compétences exercées par les personnes publiques »*[527] destinataires, en particulier *« pour l'élaboration et la mise en œuvre des plans climat-air-énergie territoriaux prévus à l'article L229-26 du Code de l'environnement »*[528]. Les personnes publiques concernées sont entendues largement[529], ce qui est souhaitable puisque les collectivités territoriales et les autorités de régulation ont un rôle essentiel dans la modernisation du secteur de l'électricité. En effet, les premières sont à l'origine de nombreux projets de villes intelligentes, optimisant notamment les consommations d'énergie grâce au recours à l'IA[530]. Les secondes sont, elles, très impliquées dans l'accompagnement de tels projets, à la fois pour

[524] *Loi n° 2015-992 du 17 août 2015 relative à la transition énergétique pour la croissance verte*, publiée au JORF n°0189 du 18 août 2015, texte n° 1, article 179.
[525] Code de l'énergie, articles L111-72 et L111-73.
[526] *Décret n° 2016-973 du 18 juillet 2016 relatif à la mise à disposition des personnes publiques de données relatives au transport, à la distribution et à la production d'électricité, de gaz naturel et de biométhane, de produits pétroliers et de chaleur et de froid*, publié au JORF n°0167 du 20 juillet 2016, texte n° 2.
[527] Code de l'énergie, article L111-72, alinéa 4.
[528] Code de l'énergie, article L111-73, alinéa 4.
[529] Code de l'énergie, article D111-55.
[530] L. Bellot, *De la smart city au territoire d'intelligence(s) : l'avenir de la smart city*, rapport au Premier ministre, avril 2017, disponible en ligne : <https://www.gouvernement.fr/sites/default/files/document/document/2017/04/rapport_smart_city_luc_belot_avril_2017_definitif.pdf>, consulté le 28 avril 2022 ; J. Haëntjens, *Smart city, ville intelligente : quels modèles pour demain ?*, La Documentation française, 2021, 180 p.

faire émerger les idées par des mécanismes d'expérimentation et pour garantir la protection des consommateurs dans leur réalisation[531].

258. **Une mise à disposition des données énergétiques auprès du public.** Deuxièmement, au-delà de la mise à disposition des données susmentionnées aux personnes publiques, le décret de 2016 précise également qu'elles peuvent être diffusées au public dans leur intégralité, s'inscrivant donc pleinement dans la politique d'*open data*[532]. En effet, au titre de l'article D111-55 du Code de l'énergie, les personnes publiques peuvent, sous certaines conditions, diffuser au public les données qui leur ont été communiquées[533]. Par transitivité, les données issues des systèmes de comptage peuvent donc être publiées en *open data*, non pas par les gestionnaires de réseau, mais indirectement par les autorités publiques. Il convient toutefois de noter que pour éviter tout risque lié à la publication de données à caractère personnel (telles que les courbes de charge), les données de consommation doivent systématiquement être agrégées et anonymisées avant publication. Ce dernier point, s'il est essentiel à la protection de la vie privée des individus, peut néanmoins freiner l'innovation. En effet, l'agrégation et l'anonymisation diminuent naturellement la précision des données alors même que la majorité des techniques employées ne sont pas infaillibles[534].

259. **Transition.** L'ensemble de ces obligations sont censées contribuer à la promotion de l'innovation dans le secteur de l'énergie et ont abouti, en pratique, à la création de plusieurs plateformes de mise à disposition des données énergétiques par les différents acteurs les détenant. La tendance est établie et se propage chaque année. Certains opérateurs vont même plus loin que leurs obligations légales et règlementaires, ce qui est souhaitable puisque les données concernées par la publication obligatoire ne sont pas très utiles du point de vue de l'innovation. Se sont ainsi développées de véritables initiatives d'*open data* volontaire. L'accroissement du volume de données disponibles, si tant est que leur réutilisation n'est pas

[531] COMMISSION DE RÉGULATION DE L'ÉNERGIE (CRE), « Dispositif d'expérimentation réglementaire », *Site officiel (blog)*, 29 juillet 2021, disponible en ligne : < https://www.cre.fr/Transition-energetique-et-innovation-technologique/dispositif-d-experimentation-reglementaire>, consulté le 28 avril 2022.
[532] O. Beatrix, *op. cit.*
[533] Code de l'énergie, article D111-55, VI.
[534] Pour un état des lieux des techniques permettant de réidentifier des individus à partir de bases de données anonymisées, voir J. Henriksen-Bulmer, S. Jeary, « Re-identification attacks—A systematic literature review », *International Journal of Information Management*, 2016, vol. 36, n° 6, pp. 1184–1192.

contrainte par d'autres régimes juridiques, doit permettre à terme de favoriser le développement de nouveaux services innovants grâce, notamment, au recours à des systèmes d'IA.

B/ La participation des acteurs à l'*open data* des données énergétiques

260. **Origine et typologie des plateformes de mise à disposition de données énergétiques.** Le rapport de la CRE de 2017 dédié aux données énergétiques dresse un état des lieux des plateformes de partage de données créées par des opérateurs du secteur de l'électricité. De nombreux acteurs ont en effet entrepris de mettre à disposition de nombreuses données de deux façons différentes d'après la CRE : *« à l'attention des utilisateurs des réseaux ou consommateurs de services, d'une part, en leur donnant l'accès aux informations qui les concernent spécifiquement et, d'autre part, en mettant à disposition du public un certain nombre d'informations générales, ou agrégées et anonymisées »*[535]. Ces plateformes ont été créées à la suite de l'entrée en vigueur des obligations légales et règlementaires issues de la loi relative à la transition énergétique pour une croissance verte puis de la loi pour une République numérique. Leur efficacité mérite d'être saluée car les bases de données ainsi mises à disposition peuvent être utilisées par des entreprises tierces, y compris des start-ups innovantes qui souhaitent concevoir de nouveaux services à destination des opérateurs régulés, des collectivités locales, des fournisseurs d'énergie ou des clients finals. Les développements qui suivent recensent des exemples de mise à disposition de données, constituant autant de ressources pour bâtir des systèmes d'IA dans le secteur de l'électricité, par les différents acteurs : gestionnaires des réseaux de distribution et de transport, fournisseurs et producteurs d'électricité. Loin de prétendre à l'exhaustivité, l'énumération des initiatives vise surtout à illustrer par l'exemple les effets concrets de la politique française d'ouverture des données énergétiques, justifiant sa pertinence pour stimuler l'innovation.

261. **La mise à disposition de données énergétiques par Enedis, le principal gestionnaire de réseaux de distribution d'électricité.** Pour se mettre en conformité avec les obligations

[535] COMMISSION DE RÉGULATION DE L'ÉNERGIE (CRE), *Rapport du comité d'études relatif aux données dont disposent les gestionnaires de réseaux et d'infrastructures d'énergie*, rapport, 18 mai 2017, p. 35.

légales et règlementaires lui incombant, le principal gestionnaire de réseaux de distribution d'électricité, Enedis, a déployé plusieurs mécanismes de mise à disposition de données[536]. Ces initiatives consistent d'abord en des mises à disposition ciblées. Enedis a ainsi mis en place un portail web[537] et une application mobile[538] à destination des clients finals, permettant de visualiser ses consommations brutes, d'activer un certain nombre de services permis par les fonctionnalités du compteur Linky, ou encore de disposer d'informations relatives aux pannes électriques et horaires de rétablissement. Le gestionnaire de réseaux de distribution rend également accessibles aux fournisseurs d'énergie les données nécessaires à la facturation des clients, les bilans de consommation, ou autres données utiles à leur activité. Enedis partage également de nombreuses données avec les producteurs d'électricité[539], les collectivités locales[540] ou les autorités concédantes, que ce soit par des procédés *ad hoc* ou *via* la création d'une plateforme dédiée[541]. Enfin, Enedis a entrepris la publication de grandes quantités de données en *open data* (à destination de tous, dans un format accessible et sans limite de réutilisation) directement sur son site institutionnel[542] et sur un site spécifique[543]. La plupart des activités du gestionnaire de réseaux sont concernées puisque les données ainsi publiées ont trait aux infrastructures exploitées (données patrimoniales sur la longueur des lignes électriques, le nombre de poste de distribution), aux installations raccordées aux réseaux, au système électrique (bilans de consommation par secteur géographique, par puissance installée et par secteur d'activité), ou encore au niveau de service (durée moyenne des coupures de courant)[544]. L'exploitation de ces données, le cas-échéant en utilisant des systèmes d'IA, permet d'avoir

[536] *Ibid.*, p. 34 et s.
[537] Disponible en ligne : <https://mon-compte-client.enedis.fr/>, consulté le 29 avril 2022.
[538] Application appelée « Enedis à mes côtés ».
[539] Voir notamment <https://www.disporeseau-enedis.fr>, plateforme d'échanges de données entre producteurs et Enedis, où les premiers partagent des données relatives aux indisponibilités envisagées des installations et où Enedis, réciproquement, fournit des informations relatives aux travaux prévus sur le réseau. Cet échange réciproque de données permet notamment d'optimiser la planification des travaux et réduire leur impact pour les producteurs.
[540] D'après la CRE dans son rapport de 2017, la mise à disposition des données par Enedis aux collectivités territoriales, notamment des agrégats de consommation à la maille territoriale ou des données géographiques sur les implantations de bornes de recharge électrique, était antérieure à la promulgation de la loi relative à la transition énergétique pour la croissance verte.
[541] Voir COMMISSION DE RÉGULATION DE L'ÉNERGIE (CRE), *op. cit.*, pp. 35–37.
[542] Disponible en ligne : <http://www.enedis.fr/open-data/>, consulté le 29 avril 2022.
[543] Disponible en ligne : <https://data.enedis.fr>, consulté le 29 avril 2022.
[544] COMMISSION DE RÉGULATION DE L'ÉNERGIE (CRE), *op. cit.*, p. 45.

une vision très précise de l'état des infrastructures publiques et des consommations à l'échelle locale. Elle est particulièrement utile aux collectivités locales et à l'État pour identifier les leviers d'action pour réduire les consommations et optimiser le fonctionnement du secteur de l'électricité, dans une logique de transition énergétique et de sobriété.

Des initiatives similaires ont été entreprises par le gestionnaire du réseau de transport d'électricité, RTE.

262. **La mise à disposition de données énergétiques par RTE, le gestionnaire du réseau de transport d'électricité.** En réponse à la multiplication des obligations légales et règlementaires visant à l'ouverture des données détenues par les opérateurs de service public, la société RTE s'est dotée d'une stratégie numérique[545]. À ce titre, l'opérateur a pour objectifs d'utiliser les données afin d'accroître la performance du système électrique, de fournir à l'écosystème de l'énergie des éléments de décision pour les politiques énergétiques, ainsi que de soutenir la recherche et l'innovation[546]. Sa principale réalisation en termes d'*open data* est la plateforme eCO2mix permettant au public de disposer d'un aperçu de la consommation d'électricité de la France, des émissions de CO2 associées à chaque filière de production, des volumes échangés à chaque interconnexion avec les pays voisins, ou encore des prix sur les marchés[547]. Les données contenues dans la plateforme sont très utiles du point de vue de l'information des citoyens et de la pédagogie vis-à-vis des émissions de gaz à effet de serre. En revanche, leur potentiel pour d'éventuelles applications de l'IA permettant d'optimiser l'équilibrage du réseau, par exemple, est relativement faible. Il semble que RTE soit le seul acteur véritablement en mesure de concevoir des systèmes d'IA visant à améliorer la gestion ou la performance du réseau de transport.

Outre les gestionnaires des réseaux de transport et de distribution, les fournisseurs et les producteurs d'électricité ont également mis en place des dispositifs de partage de données pouvant être utiles aux entreprises souhaitant développer des systèmes d'IA à destination du secteur de l'énergie.

---

[545] COMMISSION DE RÉGULATION DE L'ÉNERGIE (CRE), *op. cit.*, p 36.
[546] *Ibid.*
[547] Disponible en ligne : <https://www.rte-france.com/eco2mix>, consulté le 29 avril 2022.

263. **La mise à disposition de données énergétiques par les producteurs d'électricité.** Tout d'abord, certains producteurs d'électricité mettent à disposition des données relatives à leurs installations : capacité installée, indisponibilités, lieux d'implantation, émissions de CO_2[548]… Leur exploitation grâce à l'IA peut s'avérer utile dans un contexte de transition énergétique. En effet, par exemple, le croisement des données de production avec des données géographiques et météorologiques permet de construire des systèmes de prévision de la production future de certains territoires. De telles applications de prévision sont utiles aux opérateurs chargés de l'équilibre du réseau, ainsi que pour les fournisseurs en leur permettant de mieux anticiper les besoins d'achats d'énergie sur le marché européen.

264. **La mise à disposition de données énergétiques par les fournisseurs d'électricité.** Ensuite, ces fournisseurs peuvent eux-mêmes publier des données en *open data*, ce qu'ils ne font encore que très peu. À ce titre, le groupe EDF a lancé le 7 décembre 2020 une plateforme unique sur laquelle est publiée *« une sélection de données publiques de nature environnementale, industrielle, financière et sociale en permettant notamment leur visualisation, leur analyse et leur ré-usage »*[549]. Les jeux de données sont répartis en quatre catégories : *« Groupe EDF (données financières et sociales), Production, Consommation, Transition énergétique »*. À son lancement, la plateforme ne comptait qu'une dizaine de jeux de données, étant précisé que la quantité et la nature des données seront enrichies à mesure des nouvelles actions conduites par l'entreprise[550]. Toutefois, force est de constater que le catalogue ne s'est pas élargi outre mesure, puisqu'en avril 2022, 18 bases de données y étaient publiées. Ces dernières n'en restent pas moins utiles et intéressantes à des fins d'information du public.

265. **La mise à disposition de données énergétiques par les établissements publics et autorités sectorielles.** Enfin, conformément aux dispositions de la loi pour une République numérique, les établissements publics et les autorités de régulation agissant dans le secteur de

[548] Voir par exemple la plateforme *open data* d'Engie, dédiée aux énergies renouvelables, disponible en ligne : <https://opendata-renewables.engie.com/>.
[549] EDF, « Le groupe EDF lance sa plateforme d'open data pour faciliter la compréhension de son action environnementale, industrielle et sociale », *Communiqué de presse*, 7 décembre 2020, disponible en ligne : <https://www.edf.fr/groupe-edf/espaces-dedies/journalistes/tous-les-communiques-de-presse/le-groupe-edf-lance-sa-plateforme-d-open-data-pour-faciliter-la-comprehension-de-son-action-environnementale-industrielle-et-sociale>, consulté le 29 avril 2022.
[550] *Ibid.*

l'énergie se sont eux aussi attelés à la publication des données publiques qu'ils détiennent ou se font communiquer. Ainsi l'agence de la transition écologique (ex-ADEME) publie sur son site pas moins de 115 jeux de données, pour un total d'environ 8 millions d'entrées[551]. On y trouve majoritairement des données relatives aux bâtiments, notamment à leur performance énergétique, ou aux projets d'aménagement du territoire. La CRE a également mis en place une plateforme *open data*, principalement pour répondre aux obligations de publication des données publiques des administrations[552].

266. **Conclusion de la Section 1 relative au cadre juridique favorable à l'ouverture des données énergétiques.** À voir le nombre d'initiatives de publication des données, tout porte à croire que le cadre juridique incitant à l'ouverture des données énergétiques a porté ses fruits. Pourtant, l'on pourrait regretter que les données publiées aient avant tout une fonction d'information du public, sous la forme de tableaux visuels, facilement accessibles. Rares sont les jeux de données mobilisables pour développer des systèmes d'IA qui seraient utiles pour répondre aux enjeux du secteur de l'électricité. Les données véritablement utiles dans une logique d'innovation sont, par exemple, les courbes de charge à l'échelle d'un foyer, d'un quartier ou d'une ville ; les courbes de tension sur le réseau électrique permettant d'identifier les injections ou soutirages effectués ; ou encore les productions en temps réel de chaque moyen de production renouvelable. Cependant, toutes les données doivent-elles être partagées ? Si certains auteurs vont jusqu'à considérer que le régime actuel tend à créer un *« open data sans limite dans l'énergie »*[553], les garde-fous sont bien réels[554]. Certains sont pleinement justifiés, pour préserver la confidentialité d'informations sensibles ou la vie privée des individus par exemple. Sont-ils toujours proportionnés au vu de la nécessité d'ouvrir les données énergétiques pour stimuler le développement de l'IA dans le secteur de l'électricité ? N'existe-t-il pas des

[551] Disponible en ligne : <https://data.ademe.fr/>, consulté le 29 avril 2022.
[552] Disponible en ligne : <https://www.cre.fr/Pages-annexes/open-data>, consulté le 29 avril 2022.
[553] O. Beatrix, *op. cit.*
[554] V. notamm. A. De La Mure, S. Mathon, O. Fouqueau, « Quelles limites juridiques à la libération des données ? Comment concilier open data et protection des données à caractère personnel ? », *La Semaine Juridique Administrations et Collectivités territoriales*, 22 janvier 2018, n° 3, 2031 ; L. Cluzel-Métayer, « Les limites de l'open data », *AJDA*, 2016, p. 102.

solutions juridiques pour amplifier la tendance à l'ouverture des données, tout en préservant les droits des individus ?

Section 2 : L'ouverture des données énergétiques : une tendance contrainte

267. **Des contraintes liées aux régimes juridiques applicables aux données énergétiques.** La conception de systèmes d'IA repose sur l'exploitation de données, d'où la nécessité qu'elles soient disponibles en quantité et qualité suffisantes. Toutefois, dans le secteur de l'énergie, la mise à disposition et l'exploitation de données sont soumises à un certain nombre de règles contraignantes. Le régime juridique applicable dépend naturellement de la nature des données concernées. Les données de consommation électrique relèvent ainsi du régime des données à caractère personnel avec toutes les conséquences que cela implique. D'autres, relatives aux infrastructures publiques peuvent être qualifiées de « données publiques » au sens du CRPA[555]. En outre, certaines données détenues par les opérateurs de réseaux ou fournisseurs peuvent être protégées par des droits de propriété intellectuelle ou par le secret des affaires[556]. Enfin, la diffusion de certaines données pourrait avoir des conséquences sur le fonctionnement des marchés, notamment en termes de concurrence. En effet, certaines données peuvent contenir ce que le Code de l'énergie nomme des *« informations commercialement sensibles »* et qui sont protégées par une obligation de confidentialité renforcée[557]. L'ensemble de ces contraintes sectorielles entrent en contradiction avec la dynamique d'*open data* dans le secteur de l'énergie. Les concilier n'est pas chose aisée, si bien que leur rencontre s'est finalement faite, jusqu'à maintenant, au détriment de l'ouverture des données.

268. **L'absence d'intérêt du partage de données pour les entreprises.** De plus, les entreprises ont une tendance naturelle à refuser le partage des données puisqu'elles n'y sont pas contraintes juridiquement et n'ont aucune incitation économique à le faire. Au contraire, les

[555] Code des relations entre le public et l'administration, article L 312-3.
[556] O. Beatrix, *op. cit.*
[557] Code de l'énergie, article L111-71.

entreprises ont intérêt à accumuler et protéger leur patrimoine informationnel dans la mesure où elles peuvent valoriser leurs données à l'égard de tiers[558].

269. **Le nécessaire questionnement de la proportionnalité des contraintes à l'ouverture des données énergétiques.** La majorité des restrictions à l'exploitation des données énergétiques sont légitimes dans la mesure où elles servent à protéger la vie privée des individus, à garantir la sécurité des infrastructures ou à assurer le bon fonctionnement du marché. Toutefois, il est aussi légitime d'en questionner la proportionnalité. Les règles applicables aux données énergétiques génèrent de fortes contraintes à leur collecte, à leur diffusion ainsi qu'à leur exploitation[559]. Or, le développement de systèmes d'IA dans le secteur de l'électricité s'avère indispensable pour répondre aux enjeux de transition énergétique[560]. Il serait regrettable que la conception d'applications vertueuses soit entravée par une mauvaise articulation ou par un cumul injustifié de règlementations sectorielles. En outre, l'évolution des techniques de protection des données, telles que la cryptographie, peut remettre en cause la pertinence de certaines contraintes. En effet, la promotion de ces méthodes permettrait d'alléger le corpus existant pour faciliter la diffusion et l'exploitation des données énergétiques.

270. **Plan.** La tendance à l'ouverture des données énergétiques, si souhaitable pour promouvoir l'innovation dans le secteur de l'électricité, fait donc face à de multiples contraintes. L'articulation du principe d'*open data* avec les règlementations sectorielles de l'énergie est particulièrement complexe tant les concepts sous-jacents sont antagonistes : ouverture et libre réutilisation d'une part, confidentialité et protection de la vie privée d'autre part[561]. L'objectif de la présente Section n'est pas de lever des contraintes pleinement justifiées. Il s'agit plutôt d'identifier les solutions juridiques permettant de faciliter la diffusion et l'exploitation des données énergétiques, sans faire peser de risque sur le fonctionnement du

[558] Voir par exemple les services sur les données proposés par Enedis : <https://datahub-enedis.fr/>, consulté le 3 mai 2022.
[559] A. Fourmon, « Ouverture des données énergétiques et big data », *Énergie - Environnement – Infrastructures*, Avril 2018, n° 4, comm. 22 ; K. Huhta, « Smartening up while keeping safe? Advances in smart metering and data protection under EU law », *Journal of Energy & Natural Resources Law*, 2020, vol. 38, n°1, pp. 5–22.
[560] Voir Supra, 38-42.
[561] Voir notamm. T. Dautieu, E. Gabrié, « Analyse de l'apport de la loi pour une République numérique à la protection des données à caractère personnel (1re partie) : L'ouverture de l'accès aux données publiques et sa conciliation avec la protection des données à caractère personnel », *Communication Commerce électronique*, décembre 2016, n° 12, étude 22.

marché, la sécurité des infrastructures critiques ou la vie privée des individus. Pour ce faire, il convient d'identifier précisément les limites à l'ouverture des données, issues des règlementations sectorielles (**§1**), avant de proposer des pistes d'adaptation pour parvenir à un cadre équilibré (**§2**).

§1 : Des limites justifiées à l'ouverture des données énergétiques

271. **Plan.** La collecte, la diffusion et l'utilisation des données énergétiques se heurtent à de nombreuses contraintes règlementaires sectorielles. Elles sont principalement justifiées par la protection de la vie privée des individus (**A**), la préservation de la libre concurrence dans le secteur de l'électricité (**B**), ainsi que par la protection des secrets d'affaires et des droits de propriété intellectuelle (**C**).

A/ La protection de la vie privée

272. **Le traitement de données à caractère personnel dans le secteur de l'électricité.** Pour mener à bien leur activité, les fournisseurs d'énergie et les opérateurs de réseaux collectent et traitent une grande quantité de données à caractère personnel au sens de la Loi informatique et libertés ou du RGPD : identité des clients finals, profil, historique et lieu de consommation… Outre la fourniture d'électricité, l'exploitation de ces données permet également le développement de services innovants tels que la tarification dynamique, des applications de suivi de la consommation[562], ou encore de pilotage de la consommation à l'échelle d'un quartier ou d'un foyer. Parmi les données traitées pour concevoir ces systèmes, les données de consommation des clients font l'objet d'un régime juridique particulier, spécifique au secteur de l'électricité et plus contraignant que les règles contenues dans le RGPD.

---

[562] Voir par exemple l'application « EDF & moi » (disponible en ligne : <https://particulier.edf.fr/fr/accueil/economies-d-energie/>, consulté le 10 mai 2022). Une étude interne à EDF a démontré que les clients utilisant régulièrement l'application réalisaient jusqu'à 12% d'économie d'énergie par an grâce à la visualisation détaillée de leur consommation et aux recommandations d'actions pour économiser de l'énergie (source interne EDF, économie moyenne estimée sur la conso d'énergie par différence entre un échantillon témoin de 1 910 clients et un traité de 1 672 utilisateurs 01/06/15 au 30/06/17).

273. **Un régime spécifique et contraignant pour les données de consommation électrique.** En effet, la CNIL a été conduite à s'interroger sur la nature juridique de ces données à l'occasion du déploiement du compteur communicant Linky en 2012. À la suite de l'adoption de deux textes règlementaires organisant le déploiement des compteurs connectés en France[563], l'autorité administrative indépendante a adopté une délibération relative au traitement des données de consommation détaillées, collectées par les compteurs communicants[564]. Dans cette délibération, la CNIL s'intéresse particulièrement à la « courbe de charge », c'est-à-dire le relevé, à intervalles réguliers (le pas de mesure) de la consommation électrique de l'abonné. Selon elle, un tel relevé avec un pas de temps de 10 minutes permettrait *« d'identifier les heures de lever et de coucher, les heures ou périodes d'absence, ou encore, sous certaines conditions, le volume d'eau chaude consommée par jour, le nombre de personnes présentes dans le logement, etc »*[565]. Il s'agit donc d'une donnée à caractère personnel particulière, pouvant révéler des informations détaillées sur la vie privée – et même intime – des individus. C'est la raison pour laquelle la CNIL est venue encadrer les conditions de leur collecte et de leur utilisation.

274. **La limitation des finalités de traitement.** Premièrement, la recommandation du 15 novembre 2012 limite les finalités du traitement de la courbe de charge au nombre de trois : la maintenance et le développement du réseau de distribution par les gestionnaires de ce réseau, la mise en place de tarifs adaptés à la consommation des ménages par les fournisseurs d'énergie et la fourniture de services complémentaires par des sociétés tierces tels que des travaux d'isolation par exemple.

275. **L'encadrement des modalités de la collecte.** Deuxièmement, la CNIL encadre les modalités de la collecte de la courbe de charge. Les gestionnaires de réseau ne peuvent la collecter que lorsque des problèmes d'alimentation ont été effectivement détectés. Toute collecte systématique est considérée disproportionnée. En effet, d'après l'autorité de régulation,

---

[563] *Décret n°2010-1022 du 31 août 2010 relatif aux dispositifs de comptage sur les réseaux publics d'électricité,* publié au JORF n°0203 du 2 septembre 2010 ; A*rrêté du 4 janvier 2012 pris en application de l'article 4 du décret n° 2010-1022 du 31 août 2010 relatif aux dispositifs de comptage sur les réseaux publics d'électricité,* publié au JORF n°0008 du 10 janvier 2012.
[564] CNIL, *Délibération n° 2012-404 du 15 novembre 2012 portant recommandation relative au traitement des données de consommation détaillées collectées par les compteurs communicants*.
[565] *Ibid.*

lorsque le problème n'est pas spécifiquement localisé, les gestionnaires peuvent utiliser d'autres données moins sensibles au regard de la vie privée des personnes concernées (variations de tension et coupures de courant notamment). De plus, les fournisseurs d'énergie et les sociétés tierces souhaitant proposer des services complémentaires ne peuvent collecter la courbe de charge qu'avec le consentement exprès des personnes concernées. À cet égard, la CNIL précise que *« ce consentement doit être libre, éclairé et spécifique. Il doit donc être recueilli pour chaque prestation fournie par les fournisseurs d'énergie ou les sociétés tierces »*[566]. Elle recommande également que la collecte du consentement soit réalisée directement par le gestionnaire du réseau dans la mesure où ce dernier réalise *in fine* la collecte des données. Outre la collecte, le pas de temps de la courbe de charge est lui aussi encadré. En effet, la CNIL précise que les compteurs doivent pouvoir enregistrer la consommation selon trois pas de temps : 10 minutes, 30 minutes ou 60 minutes. La recommandation de 2012 précise que les compteurs doivent être configurés par défaut sur le paramétrage le plus protecteur, soit l'intervalle le plus long.

276. **Des obligations générales renforcées.** Troisièmement, le texte renforce les modalités d'application d'autres règles contenues dans la Loi informatique et libertés et dans le RGPD, relatives notamment à la durée de conservation, à l'information des personnes concernées et aux mesures de sécurité.

277. **Un encadrement précisé dans un référentiel sectoriel.** Ces recommandations ont été complétées par la publication d'un « pack de conformité » dédié aux compteurs communicants en mai 2014[567]. Bien que dépourvu de valeur juridique contraignante, ce document est qualifié par la CNIL de *« référentiel sectoriel »*[568] devant guider les responsables de traitement dans leur mise en conformité. Le pack de conformité distingue trois scénarios : le scénario « in-in », qui concerne la gestion des données collectées dans le logement sans communication vers l'extérieur ; le scénario « in-out », qui s'applique à la gestion des données collectées dans le logement et transmises à l'extérieur ; le scénario « in-out-in » qui concerne la gestion des

[566] *Ibid.*
[567] CNIL, *Pack de conformité sur les compteurs communicants*, mai 2014, disponible en ligne : <https://www.cnil.fr/sites/default/files/typo/document/Pack_de_Conformite_COMPTEURS_COMMUNICANTS.pdf>, consulté le 11 mai 2022.
[568] *Ibid.*, p. 2.

données collectées dans le logement et transmises à l'extérieur pour permettre un pilotage à distance de certains équipements du logement[569]. Pour chaque scénario, la CNIL définit un cadre exhaustif encadrant les finalités de traitement, les durées de conservation, les destinataires autorisés, l'information et les droits des personnes concernées ainsi que les mesures de sécurité. Il faut saluer la clarté de cet outil de conformité qui permet aux entreprises de connaître précisément les précautions à prendre, qu'elles agissent en tant que fournisseur d'énergie, gestionnaire de réseau déployant les compteurs communicants ou encore en tant que tiers prestataire de service sur les données. En revanche, il est regrettable que le cadre ainsi créé ne laisse que peu de place à l'innovation, si bien que la recherche et le développement de nouveaux services innovants n'apparaissent dans aucune des finalités autorisées, quel que soit le scénario concerné. En dépit de sa clarté, la règlementation sectorielle relative aux données produites par les compteurs communicants se veut donc particulièrement contraignante.

278. **Des contraintes justifiées.** L'ensemble de ces contraintes sur les données de consommation énergétique sont justifiées par le fait qu'elles peuvent révéler des informations précises sur la vie privée des individus. En particulier, la courbe de charge répond à un régime plus contraignant qu'une donnée à caractère personnel classique, sans pour autant relever d'une « catégorie particulière » au sens du RGPD[570].

279. **Conclusion.** La protection de la vie privée constitue une contrainte légitime à la collecte et à l'exploitation des données énergétiques, en particulier de la courbe de charge qui est susceptible de révéler des informations détaillées sur la vie privée des personnes concernées. Les régulateurs sectoriels, la CRE et la CNIL, ont pris la mesure du risque lié à l'exploitation de ces données et ont bâti une règlementation spécifique. Les contraintes ainsi créées limitent fortement la capacité d'innovation des entreprises, qu'il s'agisse d'acteurs historiques – liés par la limitation des finalités de traitement – ou de nouveaux entrants – dans l'incapacité d'accéder aux données de consommation pour concevoir leur produit.

---

[569] *Ibid.*, p. 5 ; COMMISSION DE RÉGULATION DE L'ÉNERGIE (CRE), *Rapport du comité d'études relatif aux données dont disposent les gestionnaires de réseaux et d'infrastructures d'énergie*, rapport, 18 mai 2017, p. 16.
[570] RGPD, article 8.

280. **Transition.** Néanmoins, les données énergétiques englobent une grande quantité de données et pas uniquement la courbe de charge. Certaines, d'ailleurs, ne relèvent pas de la qualification de donnée à caractère personnel, telles que les données relatives à la qualité de l'alimentation, à l'état du réseau, des sites de production, entre autres. Leur diffusion aux acteurs du marché de l'électricité peut influencer leur comportement, ou en privilégier certains. C'est la raison pour laquelle le code de l'énergie contient également des restrictions relatives aux données dont le partage pourrait avoir des conséquences sur la concurrence et le bon fonctionnement du marché.

B/ La concurrence et le bon fonctionnement du marché

281. **La confidentialité des informations commercialement sensibles.** Les articles L111-72 et L111-73 du Code de l'énergie disposent que les gestionnaires des réseaux de transport et de distribution d'électricité *« doivent préserver la confidentialité des informations d'ordre économique, commercial, industriel, financier ou technique dont la communication serait de nature à porter atteinte aux règles de concurrence libre et loyale et de non-discrimination »*[571]. Créées en France par la loi du 10 février 2000 relative à la modernisation et au développement du service public de l'électricité[572] et reprises dans la directive européenne sur le marché intérieur de l'électricité[573], ces obligations visent à empêcher que des informations commercialement avantageuses portant sur les activités des gestionnaires de réseaux ne soient divulguées de manière discriminatoire, notamment aux fournisseurs d'énergie. Ainsi, la confidentialité de ces informations est un moyen pour le législateur de préserver le caractère non discriminatoire de l'accès au réseau, en évitant de procurer un avantage concurrentiel à certains fournisseurs[574].

[571] Code de l'énergie, article L111-72 pour le gestionnaire du réseau de transport d'électricité et article L111-73 pour les gestionnaires des réseaux de distribution d'électricité.
[572] *Loi n° 2000-108 du 10 février 2000 relative à la modernisation et au développement du service public de l'électricité*, publiée au JORF n°35 du 11 février 2000, texte n° 1.
[573] *Directive (UE) 2009/72/CE du Parlement européen et du Conseil du 13 juillet 2009 concernant des règles communes pour le marché intérieur de l'électricité et abrogeant la directive 2003/54/CE*, publiée au JOUE n°L211/55 du 14 août 2009, notamm. les articles 16, 17 et 27 portant sur la confidentialité des « informations commercialement sensibles ».
[574] O. Beatrix, *op. cit.*

282. **Une articulation complexe avec les obligations d'ouverture des données.** L'esprit de ces dispositions est en contradiction avec la dynamique d'ouverture des données énergétiques qui a été engagée dans les deux dernières décennies. Pour autant, l'*open data* ne remet pas en cause les règles sectorielles relatives aux informations commercialement sensibles. Cela signifie que les gestionnaires de réseaux doivent vérifier, préalablement à la mise en œuvre de l'*open data* ou du partage de données en général, si les données concernées ne constituent pas des informations soumises à confidentialité. Fort heureusement, les articles R111-26 du Code de l'énergie précisent la nature des données protégées. Il s'agit des dispositions contractuelles et informations échangées dans ce cadre, des informations issues des comptages et autres mesures effectuées, ainsi que des informations relatives aux programmes d'appel, d'ajustement et de consommation[575]. Les deux dernières catégories peuvent correspondre à des données utiles à la conception de systèmes d'IA relatifs au pilotage du réseau (prévision des consommations, ajustement en temps réel de la production…) ou au développement de services intelligents d'analyse de la consommation pour identifier des pistes d'économies d'énergie. Le cadre législatif, aux articles L111-80 et suivants du Code de l'énergie, et règlementaire, aux articles R111-27 à R111-29, prévoient un certain nombre d'exceptions. Trois d'entre elles permettent à des tiers de se faire communiquer des données qualifiées de commercialement sensibles malgré l'obligation de confidentialité. La première permet à tout utilisateur des réseaux publics de transport ou de distribution d'autoriser les gestionnaires à communiquer directement à un tiers des informations relatives à sa propre activité. La deuxième autorise les opérateurs d'effacement, proposant aux consommateurs de réduire leur consommation à des moments précis en l'échange de tarifs avantageux, à demander aux gestionnaires la communication des données nécessaires à l'identification, à la comptabilisation et à la certification des effacements de consommation réalisés sur ces sites. La troisième exception à la confidentialité permettant d'obtenir des données pouvant être utilisées pour concevoir des systèmes intelligents vise naturellement les fournisseurs d'énergie. Ces derniers peuvent se voir notamment contraints de communiquer l'historique des données de consommation et des puissances souscrites[576].

[575] Code de l'énergie, article R111-26.

[576] Pour un panorama de l'ensemble des exceptions à la confidentialité des informations commercialement sensibles, voir les articles L111-80 à L111-83 du Code de l'énergie, ainsi que les articles R111-27 à R111-29 ; voir aussi COMMISSION DE RÉGULATION DE L'ÉNERGIE (CRE), *Rapport du comité d'études relatif aux*

Néanmoins, lorsque les données communiquées sont des données à caractère personnel, le destinataire sera toujours lié par la limitation des finalités et la nécessité de recueillir un consentement exprès, libre, spécifique et éclairé.

283. **Un régime défavorable aux acteurs extérieurs souhaitant innover dans le secteur de l'électricité.** Les règles sectorielles imposant la confidentialité des informations commercialement sensibles sont particulièrement contraignantes et concernent une grande quantité de données utiles au développement de l'IA dans le secteur de l'électricité, et notamment les données de consommation. Elles sont pourtant nécessaires à la préservation de la libre concurrence sur les marchés de l'électricité. La mise en œuvre de ces obligations limite le nombre d'acteurs pouvant innover dans le domaine des réseaux électriques et des solutions d'analyse des consommations. Les gestionnaires de réseaux sont les seuls à disposer de l'ensemble des données nécessaires à la conception de systèmes d'IA dans ces domaines, ce qui est regrettable. Il est légitime que ces derniers protègent leur patrimoine informationnel, la vie privée des consommateurs d'énergie et les informations sensibles sur l'état du réseau. En revanche, le cadre juridique applicable prive les acteurs extérieurs (start-up, entreprises de services numériques) de développer des systèmes qui pourraient aider les gestionnaires dans leur mission. Des voies juridiques existent pour « libérer » ces données : exceptions légales permettant à certains acteurs de demander la communication de données, agrégation et anonymisation des données partagées, contractualisation avec les gestionnaires de réseaux. Malgré ces possibilités, il ressort de l'analyse de la CRE en 2017 que le cadre reste complexe et déséquilibré en défaveur de l'ouverture des données et de l'innovation[577].

284. **Transition.** Une dernière contrainte peut venir ralentir les efforts d'ouverture des données énergétiques : le respect des droits, notamment de propriété intellectuelle, des détenteurs des données.

---

données dont disposent les gestionnaires de réseaux et d'infrastructures d'énergie, rapport, 18 mai 2017, p. 17 et s.

[577] COMMISSION DE RÉGULATION DE L'ÉNERGIE (CRE), *op. cit.*, p. 24.

C/ Le respect des droits des détenteurs de données

285. **Un facteur de risque contentieux important pour les entreprises.** Plusieurs droits reconnus aux détenteurs de données peuvent venir entraver la dynamique d'ouverture des données, qu'elle soit volontaire ou imposée[578]. Ils doivent nécessairement être pris en compte en amont de la publication de données, sous peine de s'exposer à des actions en justice par leurs titulaires.

286. **Les droits de propriété intellectuelle.** Les données produites et détenues par les gestionnaires de réseaux, les fournisseurs ou autres acteurs du secteur de l'électricité peuvent être protégées par des droits de propriété intellectuelle. En effet, les « bases de données », définies dans le Code de la propriété intellectuelle comme des recueils *« d'œuvres, de données ou d'autres éléments indépendants, disposés de manière systématique ou méthodique, et individuellement accessibles par des moyens électroniques ou par tout autre moyen »*[579], bénéficient d'une double protection : au titre du droit d'auteur lorsque la structure de la base est originale d'une part, et au titre du droit dit *« sui generis »* des bases de données d'autre part[580]. En particulier, le « producteur », défini à l'article L341-1 du Code de la propriété intellectuelle comme *« la personne qui prend l'initiative et le risque des investissements correspondants, bénéficie d'une protection du contenu de la base lorsque la constitution, la vérification ou la présentation de celui-ci atteste d'un investissement financier, matériel ou humain substantiel »*[581], bénéficie de droits exclusifs sur le contenu des bases de données produites. Ces derniers permettent notamment au producteur d'interdire l'extraction et la réutilisation du contenu de la base de données pour laquelle il a effectué des investissements substantiels[582]. Les juridictions administratives reconnaissent d'ailleurs la possibilité de s'opposer à la

---

[578] CRPA, art. L321-2 : *« Ne sont pas considérées comme des informations publiques, pour l'application du présent titre, les informations contenues dans des documents : […] sur lesquels des tiers détiennent des droits de propriété intellectuelle »*.
[579] Code de la propriété intellectuelle, article L112-3.
[580] A. Bensamoun, J. Groffe, « Création numérique », *Répertoire de droit civil*, Dalloz, 2013, 34–45.
[581] Code de la propriété intellectuelle, article L341-1 pour la définition du producteur de base de données.
[582] Sur l'étendue de la protection, voir les articles L342-1 à L342-6 du Code de la propriété intellectuelle.
Sur la notion d'investissement substantiel, et notamm. son interprétation jurisprudentielle par la Cour de Justice des communautés européennes, voir : CJCE, 9 novembre 2004 (quatre arrêts), *The british Horseracing Board Ltd c/ W. Hill Organization Ltd*, aff. C-203/02, *Fixtures Marketing Ltd c/ Oy Veikkaus Ltd*, aff. C-46/02, *Fixtures Marketing Ltd c/ Svenska Spel AB*, aff. C-338/02, *Fixtures Marketing Ltd c/ OPAP*, aff. C-444/02, obs. F. Pollaud-Dulian, RTD Com., 2005, p. 90.

réutilisation de données concernées par l'obligation de publication si ces dernières sont couvertes par la protection *sui generis* du Code de la propriété intellectuelle[583]. Cette solution est contestée en doctrine[584] et par la CADA[585] mais uniquement à l'égard des administrations. Les droits des « tiers » à l'administration sont bien visés par la lettre du CRPA, incluant *de facto* les entreprises exerçant des missions de service public. Or, dans le secteur de l'électricité, les gestionnaires de réseaux organisent, exploitent, publient ou mettent à disposition de grandes quantités de données collectées à partir d'équipements connectés et structurées sous la forme de bases. Cette structuration répond, pour eux, d'un investissement substantiel. Ils pourraient donc légitimement chercher à faire valoir leur droit *sui generis* des articles L341-1 et suivants du Code de la propriété intellectuelle. Pourtant, dans la pratique, cette voie n'est que peu exploitée et ce, pour deux raisons principales. D'une part, des hésitations jurisprudentielles sur la notion d'investissement substantiel et sur l'étendue de la protection ont poussé les producteurs de bases de données à chercher d'autres moyens juridiques pour protéger et valoriser leurs données. D'autre part, les détenteurs de données énergétiques disposent d'un autre moyen juridique bien plus efficace pour s'opposer à leur divulgation : l'exception liée aux secrets industriels et commerciaux.

287. **La protection du savoir-faire et des secrets d'affaires.** En effet, au-delà de l'opposabilité des droits de propriété intellectuelle sur les bases de données produites par les opérateurs de l'énergie, la question plus générale de la conciliation de la politique d'*open data* et du savoir-faire des opérateurs s'est également posée[586]. À ce titre, la loi pour une République numérique a créé une exception spécifique, dispensant les opérateurs du secteur de l'énergie en charge d'une mission de service public de leur obligation de publication des données lorsque cette communication porterait atteinte au secret en matière industrielle et commerciale. L'article 6 de la loi précise que cette qualification concerne *« le secret des procédés, des informations*

[583] Voir notamm. CAA Bordeaux, 26 févr. 2015, n° 13BX00856, Société Notre famille.com : JurisData n° 2015-006245, JCP A 2015, 2239.
[584] M. Bourgeois, A. Bounedjoum, « Les apports de la loi pour une République numérique en matière d'accès et de réutilisation d'informations publiques », *La Semaine Juridique Administrations et Collectivités territoriales*, 5 décembre 2016, n° 48, 2307.
[585] CADA, avis n° 20144578, 8 janvier 2015.
[586] O. Beatrix, *op. cit.*

économiques et financières et des stratégies commerciales ou industrielles »[587]. La définition est large et se traduit en pratique par une mobilisation très fréquente de l'argument par les gestionnaires de réseaux pour s'opposer à la publication de leurs données. On pourra regretter que ces dispositions ne fassent pas référence au régime du « secret des affaires » en matière commerciale. En effet, au sens de l'article L151-1 du Code de commerce, le secret des affaires protège les informations qui ne sont pas généralement connues ou aisément accessibles, revêtant une valeur commerciale du fait de leur caractère secret et ayant fait l'objet de mesures de protection raisonnables par son détenteur légitime. Cette qualification ne pourrait-elle pas elle aussi être invoquée pour justifier la rétention de certaines données énergétiques par les opérateurs du secteur de l'énergie ? Cet argument reste pour le moment inutilisé, la protection du Code de l'énergie étant suffisamment large et n'est conditionnée ni à la valeur commerciale des données secrètes, ni à la mise en place de mesures visant à préserver leur confidentialité.

288. **Transition.** Si les limites à l'*open data* sont nécessaires, force est de constater qu'elles sont aujourd'hui à la faveur des opérateurs historiques du secteur de l'électricité, ce qui ne permet pas à de nouveaux entrants de disposer des données nécessaires au développement de services innovants, pourtant souhaitables. Il est donc nécessaire d'équilibrer ce cadre juridique. La mise en place de mécanismes de régulation innovants ou le recours accru à de nouvelles techniques peuvent être utilisés pour stimuler le partage des données tout en protégeant les données dont la diffusion pourrait porter atteinte aux intérêts des détenteurs de données, à la concurrence ou à la vie privée des individus.

[587] *Loi n° 2016-1321 du 7 octobre 2016 pour une République numérique*, publiée au JORF n°0235 du 8 octobre 2016, texte n°1.

§2 : La nécessaire recherche d'un équilibre relatif aux contraintes à l'exploitation des données énergétiques

289. **Un équilibrage souhaitable et nécessaire pour préserver l'innovation.** Les règles sectorielles encadrant la collecte, la diffusion et l'utilisation des données énergétiques limitent la capacité d'innovation des entreprises du secteur de l'électricité ainsi que des tiers souhaitant y développer leur activité. Dans un contexte où la transition numérique est une condition à la transition écologique[588], il convient de penser un nouvel équilibre de la règlementation, en faveur de l'innovation mais sans sacrifier les garanties pour les droits et libertés des individus. Ce nouvel équilibre doit également prendre en compte les évolutions technologiques et notamment les nouvelles mesures techniques permettant de protéger la confidentialité des données.

290. **L'assouplissement des contraintes à l'exploitation de la courbe de charge.** Tout d'abord, il apparaît possible d'assouplir les contraintes juridiques à l'exploitation de la courbe de charge pour stimuler l'innovation relative à la prévision de la consommation, aux mécanismes de flexibilité et d'effacement ou encore à la maison connectée, sans faire peser de risques disproportionnés pour la vie privée des individus. Dans le régime commun de la protection des données à caractère personnel, l'élargissement de la base légale de « l'intérêt légitime » aux finalités de recherche et développement dans un but écologique (conception de systèmes visant à réaliser des économies d'énergie ou d'aide à la décision pour le pilotage du réseau par exemple) est une première piste. Concernant la règlementation sectorielle, il convient d'adapter les finalités limitatives pour lesquelles la courbe de charge peut être traitée conformément à la délibération de la CNIL en date du 15 novembre 2012. L'utilisation de la courbe de charge ayant pour finalité l'expérimentation et la recherche de solutions vertueuses du point de vue écologique devrait être considérée comme proportionnée, à partir du moment où des mesures de sécurité adéquates ont été mises en place. Dans la mesure où ce sont les gestionnaires de réseaux qui procèdent à la collecte des données, il est pertinent que ces

[588] Les institutions européennes parlent des transitions « jumelles » pour parler des transitions écologique et numérique (traduction de l'expression « *twin transition* ») : voir notamm. COMMISSION EUROPÉENNE, *Digitalising the Energy System*, Consultation publique pour un plan d'action sur la numérisation du secteur de l'énergie, octobre 2021 ; voir aussi B. Bertrand, « The Twin Digital and Green Transition », *RTD eur.*, 2022, p. 619.

expérimentations se déroulent sous leur contrôle et que les données ne soient pas transférées sans mesures de protection à n'importe quelle autre entreprise souhaitant concevoir des services innovants. En particulier, la place centrale d'Enedis et de RTE dans le fonctionnement du réseau justifie leur implication, à la fois pour se positionner en fer de lance des innovations énergétiques et pour accompagner les territoires dans ces expérimentations[589]. L'assouplissement des conditions d'exploitation de la courbe de charge pourrait être réalisé par la CNIL en amendant sa recommandation de 2012 et le pack de conformité y relatif. Ainsi, les fournisseurs d'énergie qui auraient obtenues les courbes de charge des clients avec leur consentement pourraient également utiliser ces données afin de concevoir des services innovants dans une finalité écologique en ayant recours à la base légale de l'intérêt légitime. Cette adaptation permettrait de libérer le potentiel des données en permettant aux personnes ayant déjà accès à la courbe de charge de clients de les exploiter pour une finalité écologique. Là aussi, il conviendra de cadrer la notion de finalité écologique en faisant référence aux économies d'énergie ou encore à la facilitation du pilotage du réseau. L'exploitation de la courbe de charge ainsi permise ne doit cependant pas se faire sans garde-fous techniques pour préserver la sécurité des données et notamment pour garantir qu'il n'existe pas de risque de réidentification à partir de la solution créée. À ce sujet, la recherche en cryptographie progresse rapidement. Il est temps que le droit en prenne la mesure et s'adapte pour les prendre en considération.

291. **La promotion des techniques innovantes de protection des données.** La directive européenne du 5 juin 2019 concernant des règles communes pour le marché intérieur de l'électricité et modifiant la directive précédente de 2012 a introduit l'obligation de tenir compte *« des meilleures techniques disponibles »* dans les mesures de sécurité à mettre en œuvre en matière de compteurs communicants[590]. Ces techniques sont définies dans le cadre de la

[589] M. Derdevet, « Repenser la mission d'ERDF à l'heure de la transition énergétique », *Énergie – Environnement – Infrastructures*, janvier 2016, n° 1, dossier 5.

[590] *Directive (UE) 2019/944 du Parlement européen et du Conseil du 5 juin 2019 concernant des règles communes pour le marché intérieur de l'électricité et modifiant la directive 2012/27/UE,* publiée au JOUE n°L158/125 du 14 juin 2019, article 20, b) : *« la sécurité des systèmes intelligents de mesure et de la communication des données respecte les règles de l'Union applicables en matière de sécurité en tenant dûment compte des meilleures techniques disponibles pour garantir le plus haut niveau de protection en matière de cybersécurité, tout en gardant à l'esprit les coûts et le principe de proportionnalité ».*

protection des données et de la sécurité dans un environnement de compteurs intelligents comme *« les techniques les plus efficaces, avancées et adaptées dans la pratique pour constituer, en principe, la base sur laquelle s'appuyer pour respecter les règles de l'Union en matière de protection des données et de sécurité »*[591]. Il serait pertinent de conditionner les souplesses règlementaires proposées précédemment à l'utilisation de mesures techniques très protectrices. Les avancées de la recherche, notamment en cryptographie, ont permis de développer des techniques d'anonymisation bien plus avancées qu'il y a quelques années. Le chiffrement homomorphe ou la confidentialité différentielle en sont deux exemples très prometteurs. Il est également possible de promouvoir l'utilisation de données fictives mais présentant les mêmes caractéristiques que les données originales, sans être attachées à des personnes ou infrastructures réelles. On parle alors de « données synthétiques »[592]. La CRE pourrait ainsi inciter les gestionnaires de réseaux à publier des jeux de données synthétiques présentant les mêmes caractéristiques que les données originales sans révéler aucune information sur des personnes ou infrastructures réelles. La publication de données fictives relatives à la consommation finale ou à l'état du réseau permettrait à des entreprises innovantes de modéliser numériquement le fonctionnement d'un réseau électrique et de pouvoir concevoir des services innovants à partir de ce « jumeau numérique ». Puisqu'elle serait contrainte, la création des données et leur publication devraient être supervisées par la CRE qui s'assurerait notamment de la qualité des données et de leur interopérabilité. Ces travaux pourraient être financés par les fonds publics issus de la taxe d'utilisation des réseaux publics d'électricité (TURPE) puisqu'ils ont pour vocation de faire émerger des systèmes d'IA utiles à la maîtrise de la demande en énergie, au dimensionnement précis des réseaux publics et à la meilleure planification énergétique des territoires[593]. Une autre solution envisageable serait de faire financer ces travaux directement par les entreprises innovantes ayant un intérêt à la publication

[591] *Ibid.*, article 2, 27).
[592] V. Bolón-Canedo, N. Sánchez-Maroño, A. Alonso-Betanzos, « A review of feature selection methods on synthetic data », *Knowledge and Information System*, 2013, vol. 34, 483–519 ; J.M. Abowd, L. Vilhuber, « How Protective Are Synthetic Data ? », in *Privacy in Statistical Databases*, dir. J. Domingo-Ferrer, Y. Saygın, Springer, Berlin, 2008, 5262 ; A. Gupta, A. Vedaldi, A. Zisserman, « Synthetic Data for Text Localisation in Natural Images », *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, pp. 2315-2324.
[593] COMMISSION DE RÉGULATION DE L'ÉNERGIE (CRE), *Rapport du comité d'études relatif aux données dont disposent les gestionnaires de réseaux et d'infrastructures d'énergie*, rapport, 18 mai 2017, p. 110, recommandation 6 portant sur l'intérêt des technologies de l'information pour le pilotage du réseau.

des données synthétiques, par exemple dans le cadre d'un partenariat d'entreprises ou de la création d'une chaire de recherche par la CRE. Toutefois, si l'idée est séduisante, il convient de garder à l'esprit que toutes les techniques de protection des données ne sont pas infaillibles. La Professeure Lucie Cluzel-Métayer évoque à cet égard les « risques de réidentification » des données après avoir été anonymisées[594]. Selon elle, les techniques classiques d'anonymisation présentent de tels risques lorsqu'elles sont testées à l'aune des critères posés par le G29[595], à savoir l'individualisation (est-il toujours possible d'isoler l'individu ?), la corrélation (est-il possible de relier entre eux des ensembles de données distincts concernant un même individu ?) et l'inférence (peut-on déduire de l'information sur un individu ?)[596]. Des constats similaires peuvent être dressés au sujet de l'utilisation de données synthétiques[597]. En fin de compte, tout dépendra de la sécurité des outils et techniques de chiffrement utilisés. Pour guider les entreprises, plusieurs moyens juridiques peuvent être utilisés, et notamment le recours à des normes de droit souple, non contraignantes. Pour inciter les entreprises à publier des données anonymisées ou fictives, il est important de leur procurer un cadre où elles ne risquent pas d'engager leur responsabilité en cas de risques de réidentification. La publication de standards techniques, conçus à partir d'un consensus entre le régulateur et les acteurs concernés, permettrait de répondre à cette problématique[598]. La clé de la réflexion semble être la coopération entre l'autorité de régulation sectorielle et l'ensemble des parties prenantes : gestionnaires de réseaux, fournisseurs d'énergies, start-ups innovantes, entreprises de services numériques…

292. **La coopération des acteurs et la supervision accrue du régulateur sectoriel.** Consciente de l'importance du partage des données énergétiques, notamment pour la

[594] L. Cluzel-Métayer, « Les limites de l'open data », *AJDA*, 2016, p. 102.
[595] GROUPE DE TRAVAIL « ARTICLE 29 » SUR LA PROTECTION DES DONNÉES, *Avis 05/2014 sur les Techniques d'anonymisation*, 0829/14/FR, WP216.
[596] L. Cluzel-Métayer, *op. cit.* ; voir aussi L. Maisnier-Boché, « Anonymisation : que faire pour sortir de l'impasse ? », *Expertises des systèmes d'information*, septembre 2016, n° 416, 296–300.
[597] J.M. Abowd, L. Vilhuber, « How Protective Are Synthetic Data ? », in *Privacy in Statistical Databases*, dir. J. Domingo-Ferrer, Y. Saygın, Springer, Berlin, 2008, 5262.
[598] Sur la pertinence du recours aux standards techniques pour réguler l'environnement numérique, voir notamm. D.L. Burk, « Legal and Technical Standards in Digital Rights Management Technology », *Fordham Law Review*, 2005, vol. 74, 537 ; K. Werbach, « Higher Standards Regulation in the Network Age », *Harvard Journal of Law & Technology*, 2009, vol. 23, n°1, 179 ; J.L. Contreras, « Technical standards and "ex ante" disclosure: results and analysis of an empirical study », *Jurimetrics*, 2013, vol. 53, n°2, pp. 163-211.

conception de systèmes innovants, la CRE a proposé dans son rapport de 2017 la création d'une plate-forme mutualisée de mise à disposition des données[599]. Pensée comme une véritable « agence de services numériques partagée », cette plateforme devrait proposer à tous les acteurs du marché un ensemble de services sur les données énergétiques, par exemple pour simplifier l'accès aux données qui peuvent être détenues par plusieurs opérateurs. Dans sa proposition, la CRE détaille un ensemble de principes devant guider le développement de cette plateforme. Parmi eux se trouvent le co-financement de la plateforme par les différents opérateurs et par les tarifs de réseaux et d'infrastructures publics[600], la flexibilité et l'adaptabilité concernant la nature des données manipulées, les volumes de données exploités, les types d'acteurs concernés[601], ou encore la compatibilité avec les initiatives déjà lancées par des collectivités territoriales[602]. Faute de suivi, la recommandation de la CRE est finalement restée sans effet et est tombée dans l'oubli. Sa concrétisation serait pourtant une réelle avancée pour l'innovation dans le secteur de l'électricité, en particulier si la plateforme mutualisée permet aux start-ups et entreprises de services numériques d'accéder plus facilement à des données à partir desquelles elles pourront innover. Cette facilité d'accès et d'exploitation passe en grande partie par l'homogénéisation des formats de données, leur interopérabilité et leur mise à jour régulière, autres priorités identifiées par la CRE[603]. Le rôle du régulateur sectoriel est donc primordial dans le travail d'harmonisation et d'organisation du partage des données, participant à l'équilibrage des contraintes à l'utilisation des données énergétiques en faveur de l'innovation.

293. **Conclusion de la Section 2 relative aux contraintes juridiques entravant l'ouverture des données énergétiques.** De nombreuses règles sectorielles encadrent la collecte, la diffusion et l'utilisation des données issues des activités composant le secteur de l'électricité. Les régimes applicables dépendent de la qualification des données, par nature très variées. La courbe de charge constitue ainsi une donnée à caractère personnel particulière, répondant à une règlementation spécifique ; les données comprenant des informations commercialement

[599] COMMISSION DE RÉGULATION DE L'ÉNERGIE (CRE), *op. cit.*, p. 110, recommandation 10.
[600] COMMISSION DE RÉGULATION DE L'ÉNERGIE (CRE), *op. cit.*, p. 110, recommandation 10, point 4.
[601] *Ibid.*, point 5.
[602] *Ibid.*, point 6.
[603] COMMISSION DE RÉGULATION DE L'ÉNERGIE (CRE), *op. cit.*, p. 109, recommandations 2, 4 et 5.

sensibles sont soumises à une obligation de confidentialité renforcée ; et certaines données patrimoniales sont protégées au titre du secret industriel ou de droits exclusifs de propriété intellectuelle. Ces contraintes sont justifiées dans la mesure où les objectifs poursuivis, à savoir la protection de la vie privée, la préservation de la libre concurrence et la protection des intérêts économiques des détenteurs des données, sont légitimes. Pourtant, leur application aboutit à des conséquences négatives pour l'innovation. D'abord, ces règles empêchent les opérateurs d'utiliser les données dont ils disposent pour d'autres finalités – fussent-elles vertueuses – que celles pour lesquelles les données ont été collectées. Ensuite, les règlementations sectorielles limitent grandement le champ des données pouvant être publiées en *open data*. Enfin, elles conduisent de façon contre-intuitive à une accumulation des données par certains opérateurs qui n'ont pas forcément les moyens de les exploiter pleinement[604]. Ces conséquences pourraient être évitées en adaptant le corpus juridique sectoriel pour tenir compte des avancées de la technique, notamment en promouvant le recours aux techniques de protection des données telles que les nouvelles méthodes de chiffrement.

294. **Conclusion du Chapitre 2 relatif à la promotion de l'ouverture des données énergétiques.** Le développement de l'IA dans le secteur de l'électricité dépend de la disponibilité des données qui sont issues des activités qui le composent[605]. Le législateur a bien saisi cet enjeu en initiant un mouvement en faveur de l'*open data* des données énergétiques. Si cette tendance est bien engagée dans le secteur de l'électricité, son articulation avec les nombreuses règles sectorielles peut s'avérer difficile. D'un côté, les pouvoirs publics souhaitent favoriser le partage des données pour stimuler l'innovation tandis qu'en parallèle les règlementations sectorielles organisent la confidentialité de nombreuses données pour des raisons légitimes – protection de la vie privée, concurrence, secrets industriels. La conciliation de ces deux mouvements s'est effectuée jusqu'à aujourd'hui en faveur de la confidentialité et au détriment de l'ouverture et de l'innovation.

[604] M. Derdevet, « Repenser la mission d'ERDF à l'heure de la transition énergétique », *Énergie – Environnement – Infrastructures*, janvier 2016, n° 1, dossier 5.
[605] COMMISSION EUROPÉENNE, *Digitalising the Energy System*, Consultation publique pour un plan d'action sur la numérisation du secteur de l'énergie, octobre 2021.

Le présent Chapitre a d'abord permis de mesurer l'ambition de la politique d'*open data* dans le secteur de l'énergie, que ce soit à l'échelle européenne ou nationale. Nous avons ensuite pu identifier les différentes règles sectorielles encadrant les différents types de données énergétiques et pouvant constituer des obstacles au partage des données. Enfin, les derniers développements nous ont permis d'aborder des pistes de réforme et d'adaptation afin d'équilibrer le cadre de régulation actuel. Les solutions envisagées concernent principalement l'allègement de certaines contraintes disproportionnées, la promotion des techniques innovantes de protection des données ou la nécessité d'une coopération accrue entre les acteurs du secteur et l'autorité de régulation sectorielle.

Il convient toutefois de garder à l'esprit qu'aux contraintes juridiques s'ajoutent des obstacles techniques qui complexifient encore l'ouverture des données. Se pose par exemple la question du format des données publiées. Pour que ces dernières soient facilement réutilisables, il est important qu'elles soient partagées dans un format « interopérable », c'est-à-dire exploitable par d'autres systèmes d'information[606]. De plus, si l'objectif de l'*open data* dans le secteur de l'électricité est de stimuler l'innovation, il faut que les données publiées soient d'une qualité irréprochable, sans quoi les systèmes d'IA conçus sur leur fondement seront moins performants. La diffusion et l'exploitation des données énergétiques se heurtent donc également à des contraintes, mais d'ordre technique cette fois-ci. Là encore le droit peut apporter des solutions, comme en témoigne le développement de normes de droit souple relatives à l'interopérabilité[607]. Une réflexion juridique globale est donc nécessaire afin de repenser le cadre existant pour tenir compte à la fois de la nécessaire promotion de l'innovation et des nouvelles techniques disponibles, sans sacrifier les droits et libertés des individus. Pour ce faire, le droit lui-même doit se rapprocher de la technique, que ce soit pour comprendre les risques posés par l'exploitation des données au moyen de systèmes d'IA, les contraintes

[606] B. Ahlgren, M. Hidell, E. Ngai, « Internet of Things for Smart Cities: Interoperability and Open Data », *IEEE Internet Computing*, 2016, vol. 20, n°6, pp. 52-56.

[607] Voir par exemple les objectifs du consortium européen Gaia-X visant à créer un écosystème souverain, sûr et interopérable de partage de données : GAIA-X, *Policy Rules Document*, 21 avril 2022, disponible en ligne : < https://gaia-x.eu/sites/default/files/2022-04/Gaia-X_Policy%20Rules_Document_v22.04_Final.pdf>, consulté le 3 mai 2022.
Pour un exemple plus ancien, voir le Cadre d'Interopérabilité des Systèmes d'Information de Santé (CI-SIS) : AGENCE DES SYSTEMES D'INFORMATION PARTAGES DE SANTÉ, *Cadre d'interopérabilité des SIS*, référentiel technique, 13 novembre 2012, V1.3.1, 24 p.

auxquelles sont confrontées les entreprises détenant des données, mais aussi les mesures permettant de garantir la sécurité des données. Ce n'est que par la voie de la technicisation du droit, tant par la forme des instruments juridiques mobilisés que par leur contenu, qu'il sera possible d'aboutir à un cadre de régulation proportionné.

295. **Conclusion du Titre 2 relatif à l'adaptation des règlementations sectorielles de l'énergie.** Le secteur de l'électricité fait l'objet de nombreuses régulations : fonctionnement du marché, règles de concurrence, sécurité des infrastructures, encadrement de la collecte de l'utilisation des données de consommation et bien d'autres. Ce constat est justifié par le fait que l'électricité est un bien de première nécessité : son approvisionnement est indispensable à la vie de tout pays développé. En effet, tout défaut d'approvisionnement fait peser des risques sur le bon fonctionnement de l'économie ou sur les individus. Malgré leur légitimité, ces régulations peuvent avoir des effets de bords. En particulier, elles ont donné lieu à de nombreuses règlementations pouvant contraindre le développement de l'IA, objet de notre propos. Deux exemples ont été donnés dans le présent Titre.

D'une part, la règlementation applicable en matière de sécurité des infrastructures critiques vient exclure, de fait, l'utilisation de l'IA dans les systèmes d'information critiques du réseau électrique ou des centrales de production. Le manque de contrôle sur le fonctionnement autonome de certains modèles d'IA et leur potentielle opacité rendent cette exclusion parfaitement légitime. Toutefois, ces nouvelles techniques peuvent également servir la sécurité en réduisant l'aléa humain ou en permettant de traiter de plus grandes quantités de données pour détecter d'éventuelles défaillances. Le corpus règlementaire ne devrait pas conduire à exclure de telles applications vertueuses si elles ne font pas peser de risques insurmontables sur le fonctionnement des infrastructures. Ainsi, la pertinence et la proportionnalité des règles existantes mériteraient d'être questionnées. Au vu de la technicité et de la sensibilité du sujet, elle pourrait utilement faire l'objet d'une discussion collégiale et pluridisciplinaire, réunissant autorités de contrôle, experts et autres parties prenantes.

D'autre part, les règles applicables aux données énergétiques apparaissaient à première vue comme plutôt favorables à la dynamique d'innovation puisque ces dernières sont soumises au principe de l'*open data*. Néanmoins, de nombreuses barrières à la collecte, au traitement et à la réutilisation des données issues des activités composant le secteur de l'électricité persistent :

protection de la vie privée et des droits de propriété intellectuelle, ou encore confidentialité des informations commercialement sensibles.

Pour répondre aux besoins d'assouplissement ou d'adaptation du droit existant, de nombreuses pistes de solutions ont été présentées dans le présent Titre. Chacune d'entre elles vise à bâtir un cadre plus favorable au développement de l'IA, tout en essayant de préserver les garde-fous nécessaires à la protection des droits et libertés des individus.

296. **Conclusion de la Partie 1 relative à l'adaptation du droit aux spécificités des systèmes d'IA.** L'ordre juridique actuel est le fruit d'un long héritage et a donc été bâti dans des contextes technologiques bien différents. Bien que les corpus législatif et règlementaire soient en grande partie applicables aux systèmes d'IA, certaines des dispositions qu'ils contiennent peuvent venir freiner l'innovation de façon disproportionnée. En effet, les caractéristiques spécifiques des systèmes d'IA telles que leur dépendance à la donnée, leur autonomie ou leur potentielle opacité, peuvent gêner l'application des règles en vigueur. Certaines peuvent être rendues obsolètes. C'est le cas par exemple des régimes de responsabilité pour faute lorsqu'un dommage est causé par un système autonome. D'autres règles génèrent quant à elles une contrainte juridique injustifiée au regard des bénéfices potentiels de la technologie. Ce constat se vérifie malheureusement tant dans l'étude des régimes de droit commun que dans celle des règlementations sectorielles encadrant la production, le transport ou la commercialisation d'électricité, objet de notre thèse. Pour permettre le développement de cas d'usage sûrs et vertueux de l'IA dans le secteur de l'énergie électrique, il est crucial de rééquilibrer le corpus juridique actuel, agissant aujourd'hui comme un frein à l'innovation.

Dans cette quête de rééquilibrage entre la contrainte légitime et la promotion de l'innovation, le mot d'ordre doit être la proportionnalité. Notre première Partie a démontré qu'une telle adaptation était possible. D'une part, certains régimes de droit commun méritent d'être clarifiés sans être refondus complètement afin de garantir une plus grande sécurité juridique aux entreprises développant des systèmes d'IA. D'autre part, les règlementations du secteur de l'électricité pourraient être assouplies pour permettre le recours à l'IA lorsqu'il ne pose pas de risque majeur.

297. **Transition : l'insuffisance de l'adaptation du droit existant face à des risques nouveaux.** Toutefois, l'adaptation de l'existant suffira-t-elle à garantir un usage raisonné et vertueux de l'IA dans le secteur de l'électricité ? Pour sûr, elle permet de couvrir la plupart des risques posés par le développement de l'IA (risque d'atteinte à la vie privée, réparation des dommages causés…). Or, il s'agit là uniquement des risques connus et communs à toutes les technologies. *Quid* des risques nouveaux, spécifiques aux systèmes d'IA ? Ces derniers posent en effet des questions inédites.

D'abord, l'IA pose des problématiques éthiques importantes : Comment garantir que la technologie ne remplace pas l'humain sur toutes ses tâches, y compris les plus importantes ? Comment « garder la main » sur les algorithmes pour reprendre l'expression consacrée par la CNIL dans son rapport de 2017[608] ? Comment s'assurer que le fonctionnement des systèmes d'IA ne conduise pas à la reproduction de biais discriminatoires, à la défaveur des minorités ? Autant de questions auxquelles le corpus juridique existant ne permet pas de répondre.

Ensuite, les caractéristiques des systèmes d'IA font naître de nouveaux risques de sécurité liés à la manipulation des données, à la reproduction d'erreurs figurant dans les données d'entraînement ou encore à l'incomplétude du processus d'apprentissage. Rien, dans les règles existantes, n'est prévu pour encadrer la conception de ces systèmes afin de prévenir la survenance de ces risques.

Enfin, bien que la partie logicielle des systèmes d'IA soit immatérielle, son impact sur l'environnement est bien réel : consommation énergétique des centres de données nécessaires à leur fonctionnement, besoins croissants en capacité de calcul, fabrication de capteurs communicants produisant des données exploitées par l'IA… Le droit de l'environnement est à ce stade mué sur la question de l'empreinte du numérique et en particulier de l'IA, qui est pourtant un enjeu majeur dans le contexte actuel de transition écologique.

Le corpus actuel fait défaut sur l'ensemble de ces problématiques et la communauté internationale semble en avoir tiré les conséquences en proposant activement la création d'un

---

[608] CNIL, *Comment permettre à l'homme de garder la main ? Les enjeux éthiques des algorithmes et de l'intelligence artificielle*, Synthèse du débat public animé par la CNIL dans le cadre de la mission de réflexion éthique confiée par la loi pour une République Numérique, 15 décembre 2017.

cadre juridique spécifique. L'utilisation de l'IA peut être bénéfique ou maléfique suivant l'usage : l'ordre juridique devrait également avoir pour rôle d'orienter l'innovation vers des usages vertueux. Ainsi, un cadre juridique spécifique s'avère nécessaire pour prévenir la survenance des risques non couverts par le corpus existant et pour orienter l'innovation vers des systèmes d'IA vertueux.

Partie 2 : *De lege ferenda*, la nécessaire création d'un droit spécifique

298. **L'opposition doctrinale sur l'opportunité de créer un cadre juridique pour l'IA.** La réflexion sur la création d'un droit de l'IA n'est pas nouvelle en doctrine mais fait l'objet de farouches oppositions. Certains appellent légitimement à la prudence en recommandant de s'appuyer prioritairement sur le corpus juridique en vigueur, jugé suffisamment adaptable pour répondre aux problématiques soulevées par le développement de l'IA[609]. L'autre partie de la doctrine argue en faveur de la création d'un ensemble de règles spécifiques, un droit de l'IA, fondé sur des principes directeurs tels que la préservation de l'autonomie individuelle[610]. Sans surprise, la solution la plus raisonnable se situe quelque part entre ces deux positions. À ce titre, la présente thèse promeut une approche hybride fondée sur l'adaptation de l'existant qui est pleinement applicable aux systèmes d'IA (**Partie 1**) et la création d'un nouveau cadre pour répondre aux risques nouveaux qui ne trouvent pas de réponse dans les règles juridiques en vigueur (**Partie 2**).

299. **La réalité des risques liés au développement de l'IA.** En effet, l'utilisation de l'IA génère de nombreux risques qui se matérialisent dans les cas d'usage du secteur de l'électricité bien qu'ils n'y soient pas spécifiques. Notre étude a permis d'identifier cinq risques principaux liés aux caractéristiques spécifiques des systèmes d'IA. D'abord, ils peuvent menacer la sécurité des personnes et des biens en cas de perte de contrôle ou de défaillance, notamment lorsqu'ils sont utilisés dans le pilotage d'infrastructures critiques. Ensuite, leur dépendance à la donnée les rend vulnérables à de nouvelles formes d'attaques informatiques tels que l'empoisonnement des données d'apprentissage ou la manipulation du modèle entraîné. De plus, les grandes performances des techniques d'apprentissage automatique pour l'exploitation des données incitent les entreprises à leur collecte massive, ce qui peut conduire à des abus et par conséquent

---

[609] V.-L. Benabou, « Un droit vivant. Manifeste pour des juristes incarnés et sensibles à l'heure de l'intelligence artificielle », *Mélanges Michel Vivant*, 2020, pp. 715-730 ; B. Mathis, « Proposition de règlement européen sur l'intelligence artificielle : le regard d'un praticien », *RLDI*, 2022, n°192, 4179, p. 44 ; voir aussi, sur l'adaptation des régimes juridiques existant : A. Bensamoun, G. Loiseau, « L'intégration de l'intelligence artificielle dans l'ordre juridique en droit commun : questions de temps », *Dalloz IP/IT*, 2017, p. 239.

[610] S. Merabet, *Vers un droit de l'intelligence artificielle*, Thèse de doctorat en droit privé, Dalloz, coll. Nouvelles bibliothèque de thèses, 2020, vol. 197, 1ère ed., 592 p. ; G. Guegan, *L'élévation des robots à la vie juridique*, Thèse pour le doctorat en droit privé, Université Toulouse 1 Capitole, 2016, 368 p. ; Y. Meneceur, *L'intelligence artificielle en procès : Plaidoyer pour une règlementation internationale et européenne*, Bruylant, Paris, 2020, 209 p.

causer des atteintes à la vie privée des individus. En outre, le recours à des systèmes d'IA soulève d'importantes problématiques morales et éthiques, relatives à l'autonomie individuelle ou à la reproduction algorithmique de biais présents dans la société par exemple. Enfin, les techniques d'IA nécessitent souvent la mobilisation d'importantes ressources informatiques : capacités de calcul, stockage des données, objets connectés générant des données notamment. Les systèmes qui y ont recours présentent à ce titre une empreinte environnementale non négligeable. Cette dernière est liée à la consommation d'électricité nécessaire au fonctionnement du système d'IA, aux émissions générées par toutes les étapes de son cycle de vie, telles que l'extraction des matières premières pour la fabrication des matériaux informatiques ou leur recyclage, ainsi qu'à l'utilisation qui en est faite. En effet, les systèmes logiciels fondés sur des techniques d'IA peuvent être utilisés pour résoudre un grand nombre de problèmes. Les conséquences environnementales de leur utilisation dépendront donc de la finalité à laquelle ils sont attachés : une IA peut tout autant être utilisée pour analyser des habitudes de consommation électrique et recommander des actions pour minimiser cette dernière que pour améliorer la productivité des centrales à charbon, fortement émettrices de CO2. À ce titre, le développement de l'IA et la protection de l'environnement entretiennent une relation paradoxale. Il est crucial de pouvoir mesurer les risques et bénéfices attendus de la technologie afin de pouvoir mettre en balance les intérêts divergents que sont la prévention des risques et la promotion de l'innovation.

300. **L'insuffisance de l'adaptation du corpus existant.** Le droit a naturellement un rôle à jouer dans cette mise en balance. Les risques générés par l'utilisation de l'IA appellent à une réponse juridique pour protéger les droits et libertés des individus. Cette réponse repose en premier lieu sur le corpus existant puisque l'IA n'évolue pas dans un vide juridique. Notre première Partie a permis d'évoquer un certain nombre de ces régimes, dont quelques-uns méritent d'être adaptés aux spécificités des systèmes d'IA ou pour lever des contraintes injustifiées sur le développement de la technologie. Toutefois, l'adaptation du corpus existant ne suffira pas à couvrir l'intégralité des nouveaux risques générés. En effet, certains sont trop spécifiques ou les instruments juridiques pour y répondre n'existent simplement pas encore, notamment dans le cas de l'empreinte environnementale lié au développement de l'IA. Hypothétiquement, il peut y avoir autant d'applications vertueuses que d'applications présentant un danger pour les droits et libertés des individus. Dans pareille situation, il est important de réguler le développement de la technologie pour orienter l'innovation vers les applications souhaitables.

301. **Pour une régulation du développement de l'IA.** Le terme de régulation renvoie principalement à la régulation économique, définie par le Professeur Gérard Cornu comme *« l'action mi-directive mi-corrective d'orientation, d'adaptation et de contrôle exercée par des autorités (dites de régulation) sur un marché donné (à considérer par secteur, régulation financière, boursière, énergétique, etc.) qui, en corrélation avec le caractère mouvant, divers et complexe de l'ensemble des activités dont l'équilibre est en cause, se caractérise par sa finalité (le bon fonctionnement d'un marché), la flexibilité de ses mécanismes et sa position à la jointure de l'économie et du droit en tant qu'action régulatrice elle-même soumise au droit et à contrôle juridictionnel »*[611]. Cette notion se distingue de la règlementation, définie comme *« un mode d'encadrement de l'économie qui s'entend, en la matière de deux façons : d'une part, l'encadrement unilatéral des conduites par l'édiction de normes juridiques à l'objet varié ; d'autre part, de façon plus étroite, l'encadrement de l'intervention des entreprises sur certains marchés économiques »*[612]. La règlementation consiste donc en l'assujettissement d'une activité économique à des règles contraignantes. Elle peut à ce titre être une composante d'une action de régulation économique. Le développement de l'IA englobe un ensemble d'activités constituant un « marché » dont le fonctionnement est défaillant puisqu'il aboutit à la production de systèmes d'IA non souhaitables pour la société et qui ne respectent pas toujours les droits et libertés des individus. La présente thèse, sans entrer dans le champ de la théorie de la régulation économique, argue qu'une action régulatoire est nécessaire pour orienter le développement de l'IA vers des applications vertueuses et prévenir les risques que son utilisation peut générer. Nous nous intéresserons uniquement à la composante juridique de la régulation en défendant, dans les développements qui suivent, l'opportunité de créer un cadre juridique spécifique à l'IA. Les actions régulatoires purement économiques telles que la mise en place de subventions pour la conception de certains systèmes ou des initiatives visant au développement des compétences en IA sur le territoire européen, par exemple, ne seront pas étudiées. En revanche, le cadre juridique proposé dans cette Partie ne doit pas constituer une entrave économique disproportionnée au développement de l'IA. C'est pourquoi il convient de rechercher un équilibre dans la régulation entre prévention des risques et promotion de l'innovation.

---

[611] G. Cornu, *Vocabulaire juridique*, PUF, 2014, 10ème ed., p. 884.
[612] J.P. Colson, P. Idoux, *Droit public économique*, LGDJ, 4e éd., 2008, p. 185.

302. **Plan.** Ainsi, face aux risques nouveaux et non couverts par le corpus existant malgré son adaptation, plusieurs questions apparaissent : faut-il réellement créer un cadre spécifique à l'IA ? Quels sont les moyens à mettre en œuvre pour parvenir à une régulation proportionnée ? Ou encore comment concrétiser ces réflexions théoriques dans notre contexte institutionnel et géopolitique ? La présente Partie entend apporter des éléments de réponse à chacune de ces questions en démontrant en premier lieu que la création d'un cadre juridique adapté et fondé sur une approche de régulation proportionnée permettrait de réaliser la balance des intérêts entre ce que peut apporter un système d'IA à la société et les dangers qu'il peut à l'inverse représenter (**Titre 1**). Une fois l'opportunité d'un tel cadre justifiée, il conviendra de démontrer comment il pourrait être mis en œuvre en s'intéressant notamment aux projets de règlementation actuellement en construction (**Titre 2**).

Titre 1 : L'opportunité d'une régulation *sui generis* adaptée à l'IA

303. **Objectif du Titre 1.** Le présent Titre propose une réflexion théorique sur l'opportunité de créer un cadre juridique spécifique à l'IA ainsi que, le cas-échéant, sur la forme et le contenu que ce dernier devrait adopter.

304. **La nécessaire réflexion juridique sur les modalités de la régulation de l'IA.** Le sujet de l'IA est éminemment technique et reste encore aujourd'hui l'apanage des ingénieurs et autres philosophes des sciences. Bien que la doctrine juridique se soit mobilisée ponctuellement au gré des modes et publications officielles des institutions, la réflexion juridique sur la régulation de l'IA est finalement assez pauvre. Seule la thèse du Professeur Samir Merabet aborde la question en profondeur sous un angle théorique, avec une grande hauteur de vue. Pourtant, nous croyons qu'il est effectivement pertinent de prendre du recul sur la situation avant toute intervention législative. Cette prise de recul doit permettre de répondre à deux séries de questions.

305. **La justification des objectifs poursuivis par la création d'un cadre juridique pour l'IA.** D'une part, si l'on considère que la création d'un cadre juridique spécifique à l'IA est nécessaire, il faut être en mesure de pouvoir le justifier par des motifs objectifs (**Chapitre 1**). Il s'agira alors de démontrer qu'une régulation est requise pour répondre aux risques, liés aux utilisations de l'IA, que l'adaptation du corpus existant ne saurait prévenir. L'atteinte de cet objectif légitime ne doit néanmoins pas se faire au détriment de l'innovation, souhaitable pour le développement de notre société.

306. **La justification des moyens juridiques de la régulation de l'IA.** D'autre part, une fois la nécessité justifiée et les objectifs de la régulation identifiés, il convient de s'interroger sur les moyens juridiques qui permettraient, théoriquement, de bâtir un cadre juridique proportionné et équilibré (**Chapitre 2**).

Chapitre 1 : Des objectifs légitimes en faveur de la création d'un cadre juridique spécifique à l'IA

307. **Plan.** La création d'un cadre de régulation est nécessaire pour garantir un développement harmonisé, sûr et durable de l'IA puisque l'application du droit existant ne suffit pas à couvrir tous les risques que ses applications peuvent faire peser sur la société. La technologie doit se développer dans un cadre assurant sa compatibilité avec les droits et libertés des individus, c'est pourquoi il convient de réfléchir à la façon de prévenir efficacement les risques liés à l'utilisation de l'IA, tous secteurs confondus. Toutefois, au vu des potentiels bénéfices que l'IA peut apporter en termes de performance, de sécurité ou encore de protection de l'environnement, la contrainte juridique doit toujours chercher à concilier la prévention des risques et la promotion de l'innovation (**Section 1**). Pour atteindre cet équilibre, les pouvoirs publics vont devoir réaliser des choix parmi un large spectre d'approches de régulation présentant chacune des avantages et des inconvénients au regard de l'objectif poursuivi (**Section 2**).

Section 1 : La recherche d'un équilibre entre prévention des risques et préservation de l'innovation

308. **Une régulation pour mettre en balance les risques et bénéfices de la technologie.** Le développement des nouvelles technologies est ambivalent puisqu'il fait naitre à la fois de grands espoirs de bénéfices pour la société (optimisation des processus, gains de temps ou en efficacité, calculs et résolution de problèmes jusqu'alors très complexes) et la crainte des pires dérives. Comme souvent, la réalité est à trouver quelque part entre ces deux extrêmes. Par exemple, un système d'IA peut être utilisé au sein d'une maison connectée pour optimiser la consommation énergétique du logement en déclenchant certaines actions aux moments où l'électricité est la moins chère (heures creuses) ou la plus disponible (par exemple une nuit venteuse où l'éolienne locale tourne à perte). Une telle application est vertueuse à bien des égards puisqu'elle permet de réduire la consommation du foyer, donc sa facture et ses émissions de gaz à effet de serre, tout en optimisant le fonctionnement du réseau en adaptant la consommation à la production et en augmentant la rentabilité des sources de production d'énergie renouvelable. Toutefois, cette application n'est pas sans risques. Elle nécessite notamment la connexion de nombreux équipements électroménagers du foyer et la collecte d'importantes quantités de données relatives aux habitudes de consommation et à la consommation en temps réel. Ces données sont

nécessairement transférées vers un serveur pour être analysées et traitées aux fins de fournir les services d'optimisation promis. Aucun système informatique n'étant infaillible, un tel système crée donc de nouveaux risques pour la sécurité des données à caractère personnel des individus qui, si elle n'est pas bien gérée, pourrait conduire à des usurpations d'identité ou des utilisations frauduleuses. Faut-il pour autant l'interdire complètement ? Les bénéfices potentiels que cette application peut amener compensent-ils les risques supplémentaires qu'elle génère ? Non, assurément. Il est au contraire important de gérer convenablement ces risques, tout en gardant à l'esprit les apports de l'innovation pour la société. Cet équilibre est difficile à trouver mais la régulation de l'IA doit au moins en avoir l'ambition.

309. **Plan.** Ainsi, le cadre juridique à construire doit poursuivre un double objectif à savoir l'assujettissement du développement de l'IA à la règle de droit (**§1**), d'une part, et la préservation de la capacité d'innovation des entreprises et des États dans ce domaine (**§2**), d'autre part.

§1 : Un nécessaire assujettissement du développement l'IA à la règle de droit

310. **Un cadre juridique nécessaire pour garantir le développement d'usages vertueux de l'IA.** Le succès d'une technologie émergente et la capacité des individus à l'accepter dépendent d'une part de son apport à la société, les « bénéfices attendus », et d'autre part de la façon dont sont gérés les risques y relatifs, s'ils existent[613]. La littérature est unanime sur l'existence de

[613] Les critères de l'acceptabilité sociale d'une nouvelle technologie et ce que l'on appelle la « théorie de l'acceptation du risque » ont fait l'objet d'une littérature académique abondante. Il est désormais communément admis que le bénéfice attendu (« *perceived benefit* ») et le risque perçu (« *perceived risk* ») sont les principaux déterminants de l'acceptabilité d'une technologie (M. Siegrist, « The influence of trust and perceptions of risks and benefits on the acceptance of gene technology », *Risk Analysis*, 2000, vol. 20, n°2, pp. 195-203). D'autres critères, tels que l'expérience utilisateur ou le contexte social, peuvent également avoir une influence (I. Il, K. Yongbeom, H. Hyo-Joo, « The effects of perceived risk and technology type on users' acceptance of technologies », *Information & Management*, 2008, vol. 45, n°1, pp. 1-9). En tout état de cause, la confiance (« *trust* ») dans l'organisation ayant conçu ou déployant la technologie en question augmente les bénéfices ressentis et diminue les risques perçus par la population, ce qui a donc un effet positif sur son acceptabilité (M. Siegrist, *op. cit.* ; P.A. Pavlou, « Consumer Acceptance of Electronic Commerce: Integrating Trust and Risk with the Technology Acceptance Model », International Journal of Electronic Commerce, 2003, vol. 7, n°3, pp. 101-134 ; L. Ji-Hwan, S. Chi-Hoon, « Effects of trust and perceived risk on user acceptance of a new technology service », *Social Behavior and Personality : an international Journal*, 2013, vol. 41, n°4, pp. 586-597). Pour une

risques particuliers lorsque des systèmes d'IA sont utilisés, qu'il s'agisse de risques de biais algorithmiques, de manipulation des données, ou encore de sécurité. Certains de ces risques ne sont pas suffisamment couverts par le Droit existant qui, appliqué par défaut, présente des lacunes. Pour garantir le développement d'usages vertueux de l'IA, en assurer l'acceptabilité par le grand public, et garantir sa compatibilité avec la règle de Droit, sa régulation par les pouvoirs publics apparaît désormais incontournable. L'adaptation du droit existant et la construction d'une régulation adaptée de l'IA sont aujourd'hui des conditions à la libération de son potentiel au service des grands défis de l'Humanité, en particulier des objectifs de développement durable de l'ONU[614]. Dans le secteur de l'électricité, son utilisation permettrait, par exemple, de réduire les émissions de gaz à effet de serre par la réalisation d'économies d'énergie, d'accélérer l'innovation technologique au service de la transition énergétique ou encore d'augmenter la sécurité des travailleurs évoluant dans des situations à risques. Toutefois, ces usages vertueux ne sont possibles qu'à la condition d'une régulation adéquate, à la hauteur des enjeux moraux, éthiques, sociétaux et environnementaux soulevés par l'IA.

311. **La construction d'un cadre juridique à la hauteur des enjeux juridiques et écologiques du développement de l'IA.** Les technologies ayant profondément changé la société humaine ont chacune soulevé d'importants débats : impact de l'exploitation minière sur la santé des travailleurs[615], oppositions au programme nucléaire civil depuis sa création dans les années 40[616], réseaux sociaux et addictions numériques[617]… Le développement de l'IA n'échappe pas à ce constat et, on l'a vu, génère des craintes et réticences justifiées. La réflexion sur la régulation de l'IA doit donc permettre de répondre à plusieurs grands enjeux. Au-delà de

étude générale sur l'importance de « l'éthique » dans l'acceptation sociale d'une technologie et sur la pertinence de son utilisation en tant que fondement à des règlementations : T. Aven, « On the ethical justification for the use of risk acceptance criteria », *Risk Analysis*, 2007, vol. 27, n°2, pp. 303-312.

[614] Pour des études exhaustives du rôle de l'IA dans l'atteinte des objectifs de développement durable fixé par l'ONU, V. notamm. R. Vinuesa, H. Azizpour, I. Leite, *et al.*, « The role of artificial intelligence in achieving the Sustainable Development Goals », *Nature Communications*, 2020, 11, n°233 ; I. Palomares, E. Martinez-Camara, R. Montes, *et al.*, « A panoramic view and swot analysis of artifical intelligence for achieving the sustainable development goals by 2030 : progress and prospects », *Applied Intelligence*, Springer, 2021, 51, 6497-6527.

[615] J. Rainhorn, *Santé et travail à la mine XIXème-XXIème siècles*, Presses Universitaires du Septentrion, 2014, 306 p.

[616] C. Lewandowski, *Le nucléaire*, Presses Universitaires de France, coll. Que sais-je ?, 2021, 128 p., spec. pp.101-113.

[617] M.-P. Fourquet-Courbet, D. Courbet, « Anxiété, dépression et addiction liées à la communication numérique », *Revue française des sciences de l'information et de la communication*, 2017, 11.

la dimension éthique et juridique de la régulation de l'IA, son enjeu est également sociétal, au sens où les risques générés par certaines de ses applications, en particulier dans des systèmes critiques, sont portés par la société dans son ensemble. Ces risques ne sont pas portés uniquement par les individus, entreprises ou pays qui décident du recours à l'IA, mais bien souvent par le grand public. Par exemple, le défaut d'un système d'IA utilisé pour la maintenance du réseau électrique peut avoir pour conséquence une panne localisée à l'échelle d'un quartier et causer un préjudice à de nombreux riverains. C'est la raison pour laquelle l'enjeu de l'encadrement de l'IA est également politique : des choix doivent être réalisés pour encadrer efficacement la technologie et pour que les risques qu'elle génère soient acceptables par la société. La régulation de l'IA comporte également un enjeu géopolitique majeur puisque les grandes puissances mondiales se disputent une véritable course à l'innovation[618]. L'équilibre entre la prévention des risques et la promotion de l'innovation est un équilibre difficile à trouver, que bien des règlementations ont échoué à réaliser[619]. Enfin, le temps de la régulation de l'IA s'inscrit dans un contexte de transition écologique. Cette circonstance nécessite de réfléchir à la fois à l'empreinte environnementale que pourrait avoir un déploiement massif de tels systèmes et, en même temps, de réfléchir aux moyens de mettre la technologie au service de ladite transition, et en particulier de la lutte contre le réchauffement climatique. Cet enjeu est d'autant plus important dans le secteur de l'électricité qui est l'un des plus importants contributeurs aux émissions de gaz à effet de serre à l'échelle globale[620]. La conjonction des enjeux du développement de l'IA, éthiques, sociétaux, politiques et environnementaux, justifie la nécessité d'une règlementation à la hauteur, qui se doit d'être contraignante.

---

[618] Sur les enjeux géopolitiques de l'IA, V. notamm. G. Koenig, *La fin de l'individu : voyage d'un philosophe au pays de l'intelligence artificielle*, L'Observatoire, coll. De Facto, 189 p.

[619] B. Amable, L. Demmou, I. Ledezma, « L'impact de la règlementation sur l'innovation : une analyse des performances selon la proximité à la frontière technologique », *Economie et prévision*, 2011, 1-2, 197-198, pp. 1-19 ; P. Aghion, A. Bergeaud, J. Von Reenen, « The Impact of regulation on innovation », *Document de travail de la Banque de France*, janvier 2021, n°804, disponible en ligne : <https://publications.banque-france.fr/sites/default/files/medias/documents/wp804.pdf>, consulté le 17 janvier 2022.

[620] Le secteur de l'énergie (production d'électricité et autres sources) est responsable d'environ 47% des émissions globales de CO2 en 2020 selon les chiffres de l'Agence internationale de l'énergie (IEA) : IEA, « Greenhouse gas emissions from energy », *Site officiel de l'IEA*, chiffres de 2020, disponible en ligne : <https://www.iea.org/data-and-statistics/data-browser/?country=WORLD&fuel=CO2%20emissions&indicator=CO2BySector>, consulté le 20 janvier 2022.

312. **Une règlementation au service de la régulation éthique de l'IA.** La régulation de l'IA pour éviter le développement d'usages contraires aux droits humains ou à l'intérêt public, y compris à l'égard de la préservation de la planète, est nécessaire. On entend souvent la notion « d'éthique de l'IA » ou de « régulation éthique ». Si l'on s'en tient à l'étymologie, les termes « morale » et « éthique » ont la même origine : le grec « *ethos* », traduit en latin en « *moralis* », et signifiant « *qui a trait aux mœurs, aux coutumes et aux règles de conduites et à leur justification* »[621]. Généralement, la morale réfère à un ensemble de valeurs et principes permettant de justifier l'acceptable de l'inacceptable et auxquels il faudrait se conformer, tandis que l'éthique consiste en une réflexion argumentée en vue du « bien-agir »[622]. En ce sens, les principes européens[623] et internationaux[624] pour une IA « de confiance » (le contrôle humain, la protection de la vie privée…) seraient des principes moraux, tandis que l'éthique de l'IA consisterait plutôt en la réflexion sur la façon d'appliquer ces principes, concrètement comment agir en pratique pour que l'utilisation de l'IA soit « morale ». Ainsi, réfléchir à la régulation éthique de l'IA, c'est également réfléchir à la façon dont les usages de l'IA doivent être encadrés (et s'il faut les encadrer en premier lieu) afin d'être compatible avec des principes moraux sur lesquels un consensus commence à naître. L'objet de notre propos n'est pas de refaire le débat sur les principes moraux qui doivent guider le développement de l'IA. La morale étant par nature dépendante de chaque peuple[625], cette réflexion nous paraît vaine et contreproductive. En revanche, construire une régulation de l'IA basée sur une éthique (une « ligne de conduite »), visant au respect des droits fondamentaux et à sa compatibilité avec les objectifs de développement durable de l'ONU, nous semble absolument nécessaire. Ces corpus contiennent de grands principes et droits, et sont à notre sens suffisamment partagés à l'échelle

---

[621] J. Lagarrigue, G. Lebe, « Ethique ou morale », *Recherche & Formation*, 1997, 24, pp. 121-130.
[622] COMMISSION DE L'ETHIQUE EN SCIENCE ET EN TECHNOLOGIE (CEST), « Quelle est la différence entre éthique et morale ? », *Site gouvernemental du CEST (blog)*, disponible en ligne : <https://www.ethique.gouv.qc.ca/fr/ethique/qu-est-ce-que-l-ethique/quelle-est-la-difference-entre-ethique-et-morale/>, consulté le 29 septembre 2021.
[623] GROUPE D'EXPERTS DE HAUT NIVEAU EN IA, *Lignes directrices du 8 avril 2019 pour une IA digne de confiance*, 8 avril 2019.
[624] OCDE, *Recommandation du Conseil de l'OCDE sur l'Intelligence Artificielle*, 22 mai 2019 ; V. aussi UNESCO, *Avant-projet de Recommandation sur l'éthique de l'intelligence artificielle*, *Bibliothèque numérique de l'UNESCO*, 7 septembre 2020, p. 2, disponible en ligne : <https://unesdoc.unesco.org/ark:/48223/pf0000373434_fre>, consulté le 01/04/2021.
[625] E. Durkheim, *De la division du travail social*, PUF, Paris, 1962, p. 262 : « *Chaque peuple a sa morale, qui est déterminée par les conditions dans lesquelles il vit* ».

internationale pour fonder la base d'une régulation de l'IA. À ce titre, la règlementer, c'est-à-dire assujettir ses usages à des règles, apparaît comme le meilleur moyen pour garantir que son utilisation ne contrevienne pas à ces grands principes, qu'il ne nous faut pas réinventer.

313. **Une règlementation nécessairement prudentielle.** Les questionnements autour de la régulation de l'IA ont ramené sur le devant de la scène l'éternel débat de la relation entre l'éthique et le Droit. La multiplication des chartes éthiques, émises par des institutions, des entreprises ou des associations militantes, est un fait et a déjà été longuement analysée par la doctrine[626]. La grande majorité de ces chartes ne sont pas contraignantes et rentrent dans la catégorie de la « *soft law* »[627]. En l'absence d'intervention des pouvoirs publics, on parle « d'autorégulation » des entreprises, qui génèrent elles-mêmes leurs propres codes de bonne conduite[628]. En revanche, cette approche ne nous semble pas suffisante au regard des risques et enjeux du développement de l'IA. En effet, certains de ces risques appellent, selon nous, une réponse urgente et sans compromis. Trois raisons nous poussent à considérer que la règle de Droit, contraignante par nature, doit primer sur la règle morale ou éthique pour réguler l'IA[629]. D'abord, le caractère extérieur de la règle de Droit permet de surmonter les divergences culturelles sur ce que devrait être une utilisation « morale » ou « éthique » de l'IA. Le respect des droits fondamentaux et les objectifs de développement durable ne sont pas négociables ou sujets à l'interprétation de chacun. L'assujettissement de nouvelles technologies à ces mêmes standards ne devrait également pas l'être. Ensuite, la finalité de la règle de Droit est la régulation de la vie en société, ce qui est tout l'objectif de la régulation de l'IA. Au contraire, la règle

[626] A. Jobin, M. Ienca, E. Vayena, « The Global Landscape of AI Ethics Guidelines », *Nature Machine Intelligence*, septembre 2019, vol. 1, n°9, 389-399 ; Y. Meneceur, « Analyse des principaux cadres supranationaux de régulation de l'intelligence artificielle », 31 mai 2021, disponible en ligne : <https://lestempselectriques.net/ANALYSE_IA.pdf>, consulté le 20 janvier 2022.

[627] Pour une introduction à la notion de « soft law » et une présentation de ses avantages et inconvénients, notamment en comparaison avec la « hard law » : K. Abbott, D. Snidal, « Hard and Soft Law in International Governance », *International Organization*, 2000, vol. 54, n°3, 421-456.

[628] J. Villafranco, « Self-Regulation in the Big Data and AI Space », *The Judges' Journal*, 2020, vol. 59, n°1, 32-35.

[629] Sur les domaines du droit, de l'éthique et de la morale, différenciés suivant la justification des contenus de leurs règles et la nature des sanctions applicables en cas de leur violation, voir D.A. Laprès, « Le droit, la morale et l'éthique dans la gestion des entreprises », *Revue Lamy droit des affaires*, 1er juin 2012, nº 72, citant notamm. E. Kant, *Philosophy of Law*, in *General Introduction to the Metaphysics of Law*, Section 1, reprinted in Clarence Morris (dir.), *The Great Legal Philosophers*, U. Pennsylvania, 1979, p. 240 ; J.S. Mill, *De la liberté*, 1859, spec. Chap. IV, disponible en ligne : <www.utilitarianism.com/ol/five.html>, consulté le 21 janvier 2022.

morale ou éthique poursuit une finalité individuelle : rendre l'homme « bon » ou meilleur en tant qu'individu. Enfin, le caractère coercitif de la règle de Droit nous apparaît indispensable pour assurer une régulation efficace de l'IA. Les risques que son utilisation soulève, en particulier dans le secteur de l'électricité, appellent une réponse urgente, laquelle peut être apportée par l'édiction de règles contraignantes dont le respect sera contrôlé et assuré par les pouvoirs publics dont c'est le rôle[630].

§2 : Une nécessaire préservation de la capacité d'innovation des entreprises

314. **La relation complexe entre le Droit et l'innovation.** La relation entre le Droit, ou plus précisément la règlementation, et l'innovation a toujours été ambivalente. Par nature, la règlementation est une contrainte imposée sur une activité, en ce qu'elle consiste à l'assujettir à des « règles »[631]. De plus, l'expérience de la Silicon Valley aux États-Unis a montré que le libéralisme juridique était une condition favorable à l'innovation, contrairement aux approches européenne et asiatique ayant, dans le même temps, opté pour des règles strictes en matière de propriété intellectuelle ou de protection de la vie privée[632]. C'est la raison pour laquelle de nombreux acteurs, et en particulier les géants du numérique, cherchent à éviter l'édiction d'une régulation contraignante du numérique et de l'IA[633]. En effet, la course à l'innovation dans le domaine de l'IA soulève des enjeux à la fois de compétitivité des entreprises et de géopolitique.

---

630 N.A. Smuha, E. Ahmed-Rengers, A. Harkens, *et al.*, « How the EU Can Achieve Legally Trustworthy AI: A Response to the European Commission's Proposal for an Artificial Intelligence Act », *SSRN Electronic Journal*, 5 août 2021, spec. p. 5, disponible en ligne : < https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3899991>, consulté le 21 janvier 2022 : « *The most powerful features of legal standards concern the fact that – unlike ethical or technical standards – they are mandatory and legally binding: promulgated, monitored and enforced in the context of a system of institutions, norms and a professional community which work together to ensure that laws are properly interpreted, effectively complied with and duly enforced. […] This institutional structure differentiates law from both ethics and technical best practices.* »

631 G. Cornu, *Vocabulaire juridique*, PUF, 14ème ed., 2022, p. 883 ; pour un exemple de l'impact de la règlementation sur l'innovation, voir l'étude française : P. Aghion, A. Bergeaud, J. Van Reenen, « The impact of regulation on innovation », Document de travail pour la Banque de France, janvier 2021, n°804, disponible en ligne : <https://publications.banque-france.fr/leffet-des-regulations-sur-linnovation>, consulté le 24 janvier 2022.

632 A. Chander, « How Law Made Silicon Valley », *Emory Law Journal*, 2014, vol. 63, 639.

633 R. Ochigame, « How Big Tech Manipulates Academia to Avoid Regulation », *The Intercept (blog)*, 20 décembre 2019, disponible en ligne : <https://theintercept.com/2019/12/20/mit-ethical-ai-artificial-intelligence/>, consulté le 23 janvier 2020.

D'une part, le Droit ne doit pas venir entraver la capacité des entreprises européennes à concurrencer les entreprises étrangères[634], ni la capacité de l'Union européenne à rattraper son retard technologique accumulé sur les États-Unis, la Chine ou la Russie[635]. Pourtant, l'Union n'a pas opté, une nouvelle fois, pour le libéralisme et a fait le choix de la promotion d'une IA de confiance « à l'européenne ». En effet, la confiance dans une nouvelle activité ou technologie peut également être une condition à son acceptabilité par le grand public[636]. C'est pourquoi de nombreuses entreprises et institutions se disent en faveur d'une IA « digne de confiance » ou « éthique »[637]. L'Union européenne s'inscrit dans cette dynamique et espère se démarquer de ses « concurrents » internationaux par un encadrement rigoureux d'une technologie qu'elle souhaite éthique à tous niveaux[638]. À cet égard, la régulation peut également être un atout pour le développement de l'IA, du moins c'est ce qu'espèrent les décideurs européens. Ainsi, face au risque d'entraver l'innovation par des règles trop contraignantes, il est nécessaire de penser une régulation proportionnée et de mettre en balance les risques et bénéfices liés aux usages de l'IA.

315. **De la nécessaire mise en balance des risques et bénéfices liés aux usages de l'IA.** Rares sont les applications de l'IA qui apportent des bénéfices à la société sans soulever aucun risque, que ce soit pour la sécurité des personnes, leur vie privée ou l'environnement. Si la régulation de l'IA doit venir en limiter les risques pour en garantir l'acceptabilité, il faut, en même temps, qu'elle permette à la société d'en tirer tous les bénéfices attendus. À titre

---

634 C. Castets-Renard, « Enjeux géopolitiques et juridiques de l'intelligence artificielle : quelle stratégie pour l'Union européenne ? », in *Enjeux internationaux des activités numériques : entre logique territoriale des États et puissance des acteurs privés*, dir. C. Castets-Renard,V. Ndior, L. Rass-Masson, Larcier, septembre 2020, 202 p.

635 L'émergence de géants européens du numérique et le rattrapage du retard technologique font d'ailleurs partie des priorités de l'Union européenne : M. Pollet, « PFUE : la France attendue au tournant sur le volet numérique », *Euractiv*, 14 décembre 2021, disponible en ligne : <https://www.euractiv.fr/section/economie/news/pfue-la-france-attendue-au-tournant-sur-le-volet-numerique/>, consulté le 24 janvier 2022.

636 E. Magrani, « New Perspectives on Ethics and the Laws of Artificial Intelligence », *Journal of Internet Regulation*, 2019, vol. 8, n°3, 19.

637 Thales, *Charte éthique du numérique*, document officiel, septembre 2021, disponible en ligne : <https://www.thalesgroup.com/sites/default/files/2021-10/Charte%20%C3%A9thique%20du%20num%C3%A9rique.pdf>, consulté le 11 octobre 2021 ; IBM, *IBM's Principles for trust and transparency*, 30 mai 2018, disponible en ligne : <https://www.ibm.com/blogs/policy/wp-content/uploads/2018/06/IBM_Principles_SHORT.V4.3.pdf>, consulté le 20 octobre 2019 ; A. Jobin, M. Ienca, E. Vayena, *op. cit.*

638 COMMISSION EUROPÉENNE, *Livre blanc du 19 février 2020 sur l'Intelligence Artificielle – Une approche européenne axée sur l'excellence et la confiance*, 19 février 2020, COM(2020) 65.

d'exemple, dans le secteur de l'électricité, une application d'IA visant à piloter de façon optimale un réseau de distribution d'électricité à l'échelle locale permettrait de réaliser des économies d'énergie conséquentes, et donc autant d'émissions de gaz à effet de serre évitées. Toutefois, la performance d'un tel système serait intimement liée à la capacité de son opérateur à disposer des données précises de consommation des individus, centralisant leurs données et les exposant donc à de potentielles attaques informatiques. Dans cet exemple, le risque est à la hauteur du bénéfice attendu, mais ne peut-on pas encadrer juridiquement cette situation pour sécuriser les données, réduisant le risque et faisant donc basculer la balance du côté des bénéfices ? C'est à l'aune de ce paradoxe que la pertinence des règles et mécanismes de régulation proposés devraient être analysées. Cette pertinence peut être évaluée notamment grâce à une analyse économique du droit, consistant à « *analyser la manière dont les agents économiques appréhendent l'environnement juridique, afin de comprendre l'émergence des règles de droit et d'évaluer leur pertinence* »[639]. Le présent Titre n'a pas vocation à présenter une réflexion sur l'économie de la régulation de l'IA mais certains principes, notamment ceux relatifs à la mise en balance des intérêts ou à la réflexion sur le coût social des règles de droit, nous semblent particulièrement utiles à la construction d'une régulation proportionnée de l'IA.

316. **Des compromis à réaliser pour atteindre une régulation proportionnée de l'IA.** La proportionnalité d'une régulation au regard de la mise en balance des risques et bénéfices attendus d'une activité ou technologie passe par deux leviers. Le premier levier est le choix du mode de régulation, complété, le cas-échéant, des choix réalisés dans le contenu de la règlementation. En effet, la règle de droit contraignante n'est pas la seule option de régulation et d'autres approches peuvent être utilisées : autorégulation et droit souple, approches « par le marché » *via* la mise en place de mécanismes d'incitation financière ou de subvention… Chaque approche présente des avantages et inconvénients et une analyse doit être menée pour déterminer laquelle serait la plus pertinente pour la construction d'une régulation proportionnée de l'IA. Enfin, un deuxième levier est la réalisation de compromis dans la régulation. Cette problématique n'est pas nouvelle et s'est par exemple illustrée dans les débats sur la lutte contre la haine en ligne par les plateformes et réseaux sociaux. Dans cet exemple, se posait la question

[639] Y. Gabuthy, « Analyse économique du droit : présentation générale », *Economie & Prévision*, 2013, n°202-203, pp. 1-8.

des compromis à réaliser entre la liberté d'expression, un droit fondamental, et la nécessité de limiter l'accès (notamment des mineurs) à des contenus illicites ou simplement choquants[640]. Dans notre cas, des compromis doivent être réalisés entre d'un côté le potentiel de l'IA pour la transition écologique du secteur de l'électricité et les risques que ses applications peuvent soulever pour des droits fondamentaux tels que la vie privée. De même, des mesures de régulation visant à assouplir le Droit existant pour encourager le développement d'usages vertueux de l'IA pourraient elles-mêmes présenter des risques pour certains droits. Il convient, systématiquement, d'analyser les conséquences des mesures de régulation et du contenu des règlementations à cet égard afin de s'assurer de leur proportionnalité. L'optimum ne se trouve pas dans la suppression totale du risque, mais dans une optimisation du coût social de la régulation basée sur la réalisation de compromis entre différents droits, risques et objectifs[641].

317. **Transition.** Il est légitime de s'interroger sur les moyens juridiques les plus adaptés pour atteindre la proportionnalité dans la régulation de l'IA. L'idée d'une règlementation pour encadrer la conception, le développement, l'utilisation et la commercialisation des systèmes d'IA fait son chemin et apparaît à de nombreux égards utile et nécessaire. Toutefois, la régulation de l'IA doit réaliser des compromis entre des objectifs contradictoires : la prévention des risques, d'une part, et la promotion de l'innovation, d'autre part. Ainsi, le Parlement européen reconnaissait déjà en 2017 « *qu'il y a lieu d'adopter, au niveau de l'Union, une approche graduelle, pragmatique et prudente, [...] en ce qui concerne toute future initiative relative à la robotique et à l'intelligence artificielle, de façon à ne pas mettre un frein à l'innovation* »[642]. La proportionnalité doit être une priorité dans le choix des moyens juridiques pour réguler l'IA (**Section 2**). En effet, la proportion nous semble pouvoir être apportée par la combinaison entre une règlementation contraignante pour les risques les plus importants et une autorégulation sectorielle pour les autres.

---

[640] W. Maxwell, *A Method to Assess Regulatory Measures Designed to Limit Access to Harmful Content on the Internet*, Thèse pour le doctorat en sciences économiques, Telecom ParisTech, 2016.
[641] *Ibid.*
[642] *Résolution 2015/2103(INL) du Parlement européen du 16 février 2017 contenant des recommandations à la Commission concernant des règles de droit civil sur la robotique*, publiée au JOUE n°C252/239 le 18 juillet 2018, pt. X.

Section 2 : Un équilibre à trouver dans le choix du mode de régulation

318. **Plan.** Une fois dépassée la réflexion sur les objectifs de la régulation, se pose la question des moyens à disposition pour la réaliser. Pour y répondre, il convient de s'intéresser aux options de régulation à la disposition des autorités publiques, qui se révèlent être diverses (**§1**). Chacune de ces options, allant de la règlementation la plus stricte à l'autorégulation totale, présente des avantages et des inconvénients. Pour choisir parmi elles l'approche, ou la combinaison d'approches, susceptible d'aboutir à un cadre de régulation équilibré, il apparaît utile d'étudier la façon dont d'autres secteurs ou domaines ont pu être régulés par le passé. Cette expérience de la régulation doit permettre aux pouvoirs publics de déterminer avec plus de justesse le mode de régulation le plus adapté pour répondre aux risques générés par le développement de l'IA (**§2**).

§1 : La diversité des modes de régulation envisageables

319. **Les enjeux du choix du mode de régulation.** Nous convergeons sur la nécessité d'une intervention des pouvoirs publics pour réguler une technologie qui soulève des risques pour les droits fondamentaux et qui peut, en même temps, contribuer à rendre la société meilleure. Partant, il convient de s'interroger sur les moyens juridiques disponibles pour réaliser cette régulation. Si l'exercice n'est pas usuel pour un juriste non spécialiste de la théorie, de l'histoire ou de la philosophie du droit, il n'en reste pas moins intéressant et essentiel. Ainsi, les pouvoirs publics doivent faire un choix quant au(x) mode(s) de régulation à retenir pour atteindre le double objectif de prévention des risques et de promotion de l'innovation. La contribution du juriste, notamment issu de la sphère académique, peut consister en l'éclairage sur les options de régulation à la disposition des décideurs, leurs effets attendus, et, le cas-échéant en des recommandations sur les choix qui semblent les plus à-mêmes de remplir les objectifs poursuivis. C'est là tout l'objet du présent Paragraphe.

320. **Les options quant au modèle de régulation.** Bien que la décision finale soit souvent politique, le choix d'un modèle de régulation plutôt qu'un autre devrait reposer sur une analyse

et des critères objectifs. Les options sont en effet nombreuses et il convient de les distinguer[643], avant de pouvoir discuter de leur opportunité :

- **Régulation contraignante ou non-contraignante** : cette distinction peut notamment conduire à devoir distinguer entre ce qui relève du domaine de la loi et ce qui relève de l'éthique ou de la bonne conduite, là où la morale est censée guider les comportements. Elle peut également amener à se poser la question de la mise en œuvre de la régulation : supervision par une autorité étatique ou délégation de la conformité aux acteurs visés par la règlementation.

- **Régulation générale (horizontale) ou sectorielle (verticale)** : la première permet l'harmonisation des règles sur un territoire et limite les effets de « mille-feuilles administratif », tandis que la seconde conduit à l'accumulation de règles, adaptées à des enjeux spécifiques mais parfois trop nombreuses. La logique voudrait qu'un bon modèle de régulation parvienne à trouver un équilibre entre les deux approches. Cet équilibre impose d'abord de capitaliser sur l'existant, de connaître et de comprendre les régulations sectorielles en place. Des mécanismes fonctionnant dans certains secteurs pourraient alors être dupliqués dans d'autres, ou être utilisés comme source d'inspiration pour une régulation plus générale. En revanche, par application d'un principe de subsidiarité, il convient, autant que possible, de ne pas remettre en cause des mécanismes éprouvés dans le temps. La création d'un cadre de régulation horizontal ne doit pas conduire à dégrader l'efficacité de certaines régulations sectorielles, sous couvert « d'harmonisation ». Ainsi, une initiative de régulation générale, comme celle de la Commission européenne[644], ne doit pas consister en l'ajout frénétique de nouvelles règles mais en des choix méticuleux pour éviter la création de fardeau règlementaire purement bureaucratique. C'est dans cet objectif que notre thèse propose une réflexion

[643] Sur les options de régulation à disposition des pouvoirs publics, V. notamm. : M. Ebers, « Regulating AI and Robotics: Ethical and Legal Challenges » in *Algorithms and Law*, dir. M. Ebers, S. Navas Navarro, Cambridge University Press, 2019, pp. 37-99.

[644] COMMISSION EUROPÉENNE, *Proposition de règlement du Parlement européen et du Conseil établissant des règles harmonisées concernant l'intelligence artificielle (législation sur l'intelligence artificielle) et modifiant certains actes législatifs de l'Union*, 21 avril 2021, 2021/0106 (COD).

objective sur les moyens juridiques à mettre en œuvre pour réguler l'IA de façon proportionnée[645].

- **Régulation nationale, européenne ou internationale** : le choix du niveau de la régulation est bien souvent dicté par des considérations d'ordre politique. Dans le cas des enjeux sociétaux et éthiques liés au développement des systèmes d'IA dans tous les domaines, la régulation doit se faire au plus large niveau possible. Les initiatives d'instruments normatifs lancées par l'OCDE[646] ou encore l'UNESCO[647] en témoignent. Toutefois, dans le cadre de notre thèse, nous nous focalisons sur le niveau européen puisque l'UE est aujourd'hui la communauté internationale la plus avancée sur le sujet, avec un projet de règlement d'application directe. De plus, le partage des valeurs européennes, et notamment du respect des droits fondamentaux, par les États membres de l'UE, nous semble être un terreau fertile pour bâtir une vision commune de la régulation de l'IA.

Dans la suite de nos développements, nous nous intéresserons principalement aux deux premiers champs d'interrogation : le choix du modèle de régulation et l'articulation entre les règlementations horizontales et sectorielles. En effet, comme cela vient d'être évoqué, le choix du niveau de la régulation répond avant tout de considérations politiques.

321. **Les options quant au contenu des mesures de régulation.** Suivant le modèle de régulation retenu, dont le choix peut être guidé par des réflexions économiques, juridiques ou philosophiques, il persistera encore la question du contenu des mesures. Dans une démarche de règlementation d'une activité ou d'une technologie, on pourra ainsi encadrer la conception par des normes contraignantes (« *design regulation* »[648]), créer des mécanismes de certification *ex-ante* ou des règles de responsabilité *ex-post*[649]. Au contraire, dans une démarche « incitative »

[645] Voir Infra, Chapitre 2.
[646] OCDE, *Recommandation du Conseil de l'OCDE sur l'Intelligence Artificielle*, 22 mai 2019.
[647] UNESCO, *Avant-projet de Recommandation sur l'éthique de l'intelligence artificielle*, *Bibliothèque numérique de l'UNESCO*, 7 septembre 2020, p. 2, disponible en ligne : <https://unesdoc.unesco.org/ark:/48223/pf0000373434_fre>, consulté le 01/04/2021.
[648] N. Leveson, « The use of safety case in certification and regulation », *MIT ESD Working paper series*, novembre 2011, n°2011-13, spec. pp.1-2 « *Types of regulation* ».
[649] Sur la nécessité de mesures de régulation à la fois *ex-ante* et *ex-post* en matière d'IA, V. notamm. C. Castets-Renard, « Comment construire une intelligence artificielle responsable et inclusive ? », *Recueil Dalloz*, 2020, 225 ;

de régulation volontariste, les pouvoirs publics pourraient avoir recours à des mécanismes tournés vers le marché (« *market based regulation* »[650]) de subventions conditionnées à la réalisation de certains objectifs ou du respect de normes[651], la création de codes de bonne conduite[652] ou encore des obligations de transparence pouvant avoir des effets sur le marché[653]. Au cours de nos recherches, il nous est apparu que toutes ces options n'ont que très peu été analysées par la doctrine française, en particulier dans le contexte de la régulation de l'IA[654]. À ce titre, l'étude du droit de la régulation et des expériences de régulation passées, y compris en droit public, nous apparaissent indispensables pour identifier les meilleurs mécanismes juridiques à mobiliser pour construire un cadre de régulation adapté à l'IA.

322. **La possible combinaison d'approches de régulation.** Il existe plusieurs approches de la régulation, allant de la règlementation la plus contraignante à l'autorégulation totale des entreprises, qu'il conviendra de présenter et d'analyser en détails dans nos développements. Toutefois, il faut garder à l'esprit que ces approches ne sont pas exclusives. Elles peuvent tout

---

J. Sénéchal, « Responsabilisation ab initio, régulation ex ante et responsabilités a posteriori : le coeur des débats européens sur les systèmes d'intelligence artificielle, hors et dans le secteur du commerce électronique », *Dalloz IP/IT*, 2020, p. 667. Plus spécifiquement sur la fonction normative des règles de responsabilité *ex-post*, V. notamm. M. Mekki, « Les fonctions de la responsabilité civile à l'épreuve du numérique : l'exemple des logiciels prédictifs », *Dalloz IP/IT*, 2020, p. 672.

[650] R.N. Stavins, « Market-Based Environmental Policies », in *Public Policies for Environmental Protection*, dir. P.R. Portney, R.N. Stavins, Resources for the Future, 2000, 2ème ed., spec. p. 1 : *« Market-based instruments are regulations that encourage behavior through market signals rather than through explicit directives regarding pollution control levels or methods »*.

[651] A titre d'exemple, l'article 45 loi de finances 2021 a conditionné le bénéfice du tarif réduit de la taxe sur l'électricité consommée par un data center à la mise en place de mesures ambitieuses de limitation de leur empreinte environnementale. V. SENAT, *Projet de loi de finances 2021 : Economie*, Avis n°139 au nom de la commission des affaires économiques sur le projet de loi de finances, adopté par l'Assemblée nationale, pour 2021, spec. II « Un nécessaire verdissement de la fiscalité applicable aux data centers, à concilier avec l'enjeu d'attractivité du territoire ».

[652] Pour un exemple de code de bonne conduite, V. par ex. le code établi par des acteurs européens du Cloud Computing, sur le fondement du RGPD : CISPE.CLOUD, *Code de Conduite CISPE relatif à la Protection des Données*, 9 février 2021, disponible en ligne : <https://www.cnil.fr/sites/default/files/atoms/files/code_de_conduite_des_fournisseurs_dinfrastructures_cloud_relatif_a_la_protection_des_donnees_-_cispe_-_version_francaise.pdf>, consulté le 25 janvier 2022.

[653] Sur l'effet normatif de l'obligation de reporting extra-financier, V. notamm. B. Boyer-Allirol, « Faut-il mieux réglementer le reporting extrafinancier ? », *Revue française de gestion*, 2013, vol. 8, n°237, pp. 73-95.

[654] Plusieurs auteurs français ont étudié l'application des régimes de droit privé au nouvel objet de droit qu'est l'IA (voir notamm. A. Bensamoun, G. Loiseau, « L'intelligence artificielle : faut-il légiférer », *Recueil Dalloz*, 2017, p. 581 ; A. Bensamoun, G. Loiseau, « L'intégration de l'intelligence artificielle dans l'ordre juridique en droit commun : questions de temps », *Dalloz IP/IT*, 2017, p. 239 ; J. Larrieu, « Robot et propriété intellectuelle », *Dalloz IP/IT*, 2016, p. 291) mais les sources étudiant les options de régulation sont plus rares (C. Castets-Renard, « Le Livre blanc de la Commission européenne sur l'intelligence artificielle : vers la confiance ? », *Recueil Dalloz*, 2020, n°15, p. 837).

à fait être combinées selon les problématiques à réguler. Dès lors, leur potentielle complémentarité devra être étudiée afin de limiter la contrainte juridique au strict nécessaire. De telles réflexions ont déjà été menées en doctrine sur le même sujet[655] ou d'autres[656], soulignant l'importance de la bonne combinaison d'approches de régulation suivant l'objectif poursuivi. Ces études constituent un bon point de départ dans la réflexion sur la construction d'une régulation proportionnée de l'IA, n'entravant pas l'innovation.

323. **Les expériences de la régulation comme boussoles.** Pour répondre à la question du choix du modèle de régulation à retenir, l'expérience de récentes régulations peut être très utile. En matière environnementale, la nature des risques combattus présente des similarités avec ceux engendrés par l'utilisation de l'IA[657]. Dans le secteur bancaire et financier, la mise en place d'une régulation prudentielle prouve que l'on peut réguler *ex ante* des comportements en déléguant la conformité aux acteurs les mieux placés pour mettre en place des mesures de prévention des risques. Enfin, la protection des données personnelles est un bon exemple de combinaison d'approches de régulation avec sa mise en œuvre basée sur une forte responsabilisation des acteurs, la reconnaissance de droits spécifiques aux individus et un contrôle *ex-post* par une autorité de supervision. Prendre en exemple les expériences récentes de régulation permettrait également d'identifier des approches innovantes, autres que les traditionnelles approches de règlementations verticales et prescriptrices qui ne produiraient pas forcément les effets escomptés compte tenu des enjeux spécifiques de la régulation de l'IA.

[655] M. Ebers, *op. cit.*
[656] U. Pagallo, P. Casanovas, R. Madelin, « The Middle-out Approach: Assessing Models of Legal Governance in Data Protection, Artificial Intelligence, and the Web of Data », *The Theory and Practice of Legislation*, 2 janvier 2019, vol. 7, n°1, 25.
[657] N.A. Smuha, « Beyond the Individual: Governing AI's Societal Harm », *Internet Policy Review*, 30 septembre 2021, vol. 10, n°3.

§2 : L'expérience de la régulation comme source d'inspiration

324. **La pertinence de l'étude d'autres activités régulées pour déterminer les moyens adaptés à la régulation de l'IA.** De nombreuses leçons peuvent être tirées de l'expérience de la régulation. En effet, d'autres activités présentant des risques spécifiques ont fait l'objet, par le passé, de mesures de régulation. La plupart du temps ces régulations ont fait appel, au moins partiellement, à la règlementation – l'édiction de règles – pour guider les comportements des opérateurs économiques, prévenir les risques et favoriser le bon fonctionnement du marché. Ces initiatives peuvent paraître très diverses tant les règlementations qui en découlent sont techniques et tant les secteurs concernés (télécommunications, énergie, finance…) sont différents. Pourtant, leur étude a permis de mettre en évidence des principes et objectifs communs, faisant émerger un « droit de la régulation »[658]. Elle permet également, selon nous, de fournir des exemples de modes de régulation expérimentés par le passé et de servir de source d'inspiration lorsque le législateur se retrouve face à une nouvelle activité à réguler. Dans notre cas, l'objet de la régulation est à ce stade bien établi : les systèmes d'IA tels que nous les avons définis en introduction[659]. Il nous reste à définir les mécanismes de régulation, y compris de règlementation, qu'il convient d'intégrer dans l'ordre juridique afin d'atteindre le double objectif de prévention des risques pour les droits fondamentaux des individus et de promotion de l'innovation. L'étude des expériences de régulation peut, selon nous, éclairer ce choix. En particulier, la régulation de l'IA pourrait s'inspirer de certaines approches ou mécanismes du droit de l'environnement[660], de la protection des données[661] ou de la régulation financière[662],

[658] M.A. Frison-Roche, « Le droit de la régulation », *Dalloz*, 2001, n°7, pp. 610-616.
[659] Voir Supra, 21.
[660] A l'échelle internationale, quelques auteurs seulement défendent la pertinence du modèle du Droit de l'environnement pour inspirer la régulation du numérique et de l'IA plus spécifiquement : V. notamm. en droit américain : D.D. Hirsch, « Protecting the Inner Environment: What Privacy Regulation Can Learn from Environmental Law », *Georgia Law Review*, 2006, vol. 41, n°1 ; en droit canadien : S.M. Smyth, « The Greening of Canadian Cyber Laws: What Environmental Law can Teach and Cyber Law can learn », *International Journal of Cyber Criminology*, 2014, vol. 8, n°2, 111-155 ; et en droit européen : N.A. Smuha, « Beyond the Individual: Governing AI's Societal Harm », *op. cit.*
[661] Pour un exemple américain où les mécanismes du RGPD sont utilisés comme modèle pour construire une règlementation de l'IA, voir M.E. Kaminski, J.M. Urban, « The Right to Contest AI », *Columbia Law Review*, 16 novembre 2021, vol. 121, n°7.
[662] O.J. Erdélyi, J. Goldsmith, « Regulating Artificial Intelligence: Proposal for a Global Solution », *Proceedings of the 2018 AAAI/ACM Conference on AI, Ethics, and Society*, 2018, 95-101.

puisque les risques adressés (**A**) et les objectifs poursuivis (**B**) présentent des similitudes avec notre sujet.

A/ Des similitudes dans la nature des risques à adresser par la régulation

325. **Des régulations répondant à des risques de commune nature.** Les risques générés par l'utilisation de systèmes d'IA convergent en de nombreux points avec les risques adressés par les régulations environnementale, financière et des données personnelles. Ce constat justifie que l'on s'inspire des instruments juridiques mis en œuvre dans ces domaines dans la création d'un droit de l'IA.

326. **Les similitudes avec le risque environnemental.** Premièrement, les risques créés par l'utilisation de systèmes d'IA convergent avec le risque environnemental en ce qu'il est nouveau, évolutif, de portée globale et d'une grande complexité. Les travaux de la chercheuse australienne Sara Smyth en 2014 montrent que ces quatre caractéristiques peuvent s'appliquer de façon similaire aux potentielles atteintes aux droits fondamentaux dans le cyberespace qu'aux dommages écologiques causés par des activités polluantes[663]. Elles s'appliquent de façon indifférenciée, selon nous, aux risques générés par l'utilisation de systèmes d'IA. Cette idée selon laquelle les préjudices liés à l'utilisation de technologies numériques présentent des similitudes avec les préjudices écologiques n'est pas nouvelle et se retrouve dans la doctrine américaine[664] et européenne[665]. En particulier la chercheuse Nathalie Smuha considère à cet égard que le risque de l'IA est, tout comme le risque environnemental, un risque « sociétal »[666]. Le risque devient sociétal lorsqu'il ne concerne pas uniquement un potentiel danger pour des intérêts individuels mais au contraire pour la société dans son ensemble[667]. Dans le cas de l'IA, cela se traduit, par exemple, par la disparition progressive de l'humain dans les processus de

[663] S.M. Smyth, *op. cit.*
[664] D.D. Hirsh, *op. cit.*, spec. p. 23.
[665] N.A. Smuha, « Beyond the Individual: Governing AI's Societal Harm », *op. cit.*, spec. pp. 9-12.
[666] *Ibid.*
[667] Dans le même sens voir : C. Kutz, *Complicity: Ethics and Law for a Collective Age*, Cambridge University Press, 2000, 1st edition, 344p.

prise de décision. Aujourd'hui, les préjudices individuels liés à l'automatisation sont relativement minimes – suppression d'emplois compensée par les emplois créés, marge d'erreur des systèmes automatisés plus faible que l'erreur humaine – mais ses conséquences à long terme pour l'ensemble de la société sont beaucoup plus importantes : perte de savoir-faire de l'humain dans toutes les industries complètement automatisées, dépendance à la machine, résilience systémique dépendante du bon fonctionnement des infrastructures numériques,… Ce caractère sociétal, que l'on retrouve également dans le risque environnemental, a par le passé justifié la mise en place des mécanismes visant à encadrer les activités concernées[668].

327. **Les similitudes avec les risques financiers.** Deuxièmement, les risques créés par l'utilisation de systèmes d'IA convergent avec les risques financiers en ce qu'ils font courir à la société un risque systémique[669]. Raison d'être de la règlementation prudentielle sur les marchés financiers[670], le risque systémique se caractérise par le fait que le comportement d'un acteur peut avoir des conséquences sur la stabilité du système financier dans son ensemble[671]. L'utilisation massive d'algorithmes pour réaliser du trading haute fréquence a déjà illustré le risque qu'il pouvait faire peser sur l'ensemble des marchés financiers. Si toutes les grandes banques et fonds d'investissement se dotent d'algorithmes pour automatiser leurs activités de trading, l'erreur commise par un seul de ces algorithmes pourrait envoyer un signal erroné au marché et ainsi causer un effet « boule de neige »[672]. D'ailleurs, l'erreur pourrait être suivie et

---

668 *Ibid.*, spec. p. 9 et p. 12 : « *Inspiration can be drawn from a legal domain that is specifically aimed at protecting a societal interest: (EU) environmental law. While a polluting practice or activity —whether undertaken by a public or private actor—is liable to cause individual harm, the interest to secure a clean and healthy environment is one that is shared by society at large. An individual living in city A might be unaware of, choose to ignore, or be indifferent to the pol-luting practice. Yet the adverse effects that will ensue from the practice are likely to go over and above the individual or collective level, as it can give rise to harm also for people living in city B, country C and region D, and to future generations* ».

669 W.A. Ben Youssef, « Les cyber risques : nature, étendue et moyens de couverture », *Droit et patrimoine*, 1er janvier 2020, n°298 : « *La transposition [des] trois critères [du risque systémique (taille, non-substituabilité, interconnexion)] aux cyberactivités révèle le caractère intrinsèquement systémique du cyber risque, notamment pour son degré d'interconnexion relativement élevé* ».

670 D. Blache, « Règles prudentielles européennes applicables aux établissements de crédit, entreprises d'investissement, établissements de paiement et établissements de monnaie électronique », *JurisClasseur Banque - Crédit - Bourse*, Fasc. 110 Droit bancaire et financier européen, spec. 28-32.

671 J.-F. Lepetit, *Rapport sur le risque systémique*, rapport commandité par Ministère de l'Économie, de l'Industrie et de l'Emploi, *Documentation française*, 2010, 108 p., spec. p. 12 : « *Le déclenchement d'une crise systémique naît d'un choc qui se propage à l'ensemble du secteur financier* ».

672 H. Poulenard, « L'encadrement juridique des algorithmes mis en œuvre sur les marchés financiers », *Revue Lamy droit des affaires*, 1er décembre 2018, nº 143.

reproduite autant par des algorithmes entièrement automatisés, que par un humain ayant une confiance aveugle dans le fonctionnement de la machine. Cet exemple illustre bien la nature systémique du risque sur les marchés financiers, et comment l'IA peut y contribuer. Ce raisonnement est applicable à tous les secteurs et marchés dont le fonctionnement est fondé sur des prises de décisions humaines. C'est le cas d'un grand nombre d'activités comme le secteur financier, déjà étudié, avec les décisions concernant les ordres de bourse, le secteur médical, avec les diagnostics de santé, ou encore le secteur de l'électricité, dont toutes les activités reposent sur des décisions humaines : conduite d'une centrale de production d'électricité, prévision de consommation et planification de l'offre anté-journalière, maintenance des infrastructures… Dans ces secteurs, dont la liste n'est pas exhaustive, les conséquences d'un défaut du système d'IA utilisé aux fins de prise de décision peuvent se répercuter sur l'ensemble d'un système complexe. Dans le secteur de l'électricité, le défaut du système d'IA utilisé pour la prévision de la consommation peut conduire les gestionnaires de réseau à mal ajuster la production d'électricité le lendemain et, ainsi, compromettre le fonctionnement de toute la chaîne de l'approvisionnement en électricité : les opérateurs de centrales de production recevront l'ordre de réduire leur production, les prix de l'électricité diminueront dans un premier temps (car il y aura une offre excédentaire) puis augmenteront massivement dans un second temps (quand le marché découvrira que l'offre est déficitaire par rapport à la demande), et la confiance dans les opérateurs de réseau pour garantir la sécurité d'approvisionnement en électricité sera gravement atteinte. La nature systémique du risque de l'IA, similaire en ce point au risque sur les marchés financiers, relève donc d'un enjeu de souveraineté humaine et de résilience face à la machine, tous secteurs confondus. Elle appelle, comme cela a été fait pour la finance, à l'édiction de règles pour minimiser la probabilité de survenance du risque : une règlementation prudentielle[673].

328. **Les similitudes avec les risques d'atteinte à la vie privée.** Troisièmement, les utilisations de systèmes d'IA peuvent atteindre à la vie privée des individus, comme tout traitement de données à caractère personnel. La nature de ce risque, qui a donc justifié la création des régimes de protection des données à travers le monde, est particulière et accentuée

[673] V. Dhar, « The future of Artificial Intelligence », Big Data, 2016, vol. 4, n°1.

par la nature même des systèmes d'IA. En effet, ils permettent le traitement de plus grandes quantités de données et leur conception elle-même peut être réalisée à partir de données. Ses particularités peuvent entrer en contradiction avec les objectifs de la régulation en matière de protection des données à caractère personnel, pensée initialement pour limiter les atteintes à la vie privée des individus dans l'environnement numérique[674]. Aussi, les régimes de protection des données s'appliquent déjà en grande partie aux systèmes d'IA[675]. Ils peuvent donc à la fois être pris en exemple pour identifier des moyens efficaces et proportionnés pour la régulation de l'IA, mais doivent aussi être questionnés quant à leur pertinence face à l'évolution des techniques.

329. **Conclusion et transition.** Ainsi, les natures sociétale, systémique et informationnelle des risques générés par l'utilisation des systèmes d'IA justifie l'intérêt à porter aux modèles de régulation environnementale, financière et de protection de la vie privée. Outre la similitude dans la nature des risques adressés, les objectifs poursuivis par chacun de ces modèles sont également transposables à la régulation de l'IA.

B/ Des similitudes dans les objectifs de la régulation

330. **Des régulations répondant à des objectifs de commune nature.** La régulation d'une activité peut poursuivre de nombreux objectifs : réduction des risques, ouverture d'un marché à la concurrence, garantie de la sécurité des produits… Les objectifs poursuivis par le Droit de l'environnement, les règlementations financières ou les régimes de protection des données à caractère personnel présentent des similitudes avec les objectifs de la régulation de l'IA proposée dans notre thèse. Ce constat confirme l'intérêt de s'inspirer des instruments juridiques mis en œuvre dans ces domaines pour bâtir une régulation proportionnée de l'IA.

331. **Un objectif commun de régulation économique.** La raison d'être de la régulation économique est de pallier la défaillance d'un marché[676]. Le Droit de l'environnement participe à la régulation économique en matière environnementale en ce qu'il participe à l'internalisation

[674] Voir Supra, 107-112.
[675] Voir Supra, 93 et s.
[676] G. Cornu, *Vocabulaire juridique*, PUF, 2014, 10ème ed., p. 884.

d'une externalité négative[677]. Cette dernière existe lorsque des individus ou des groupes d'individus utilisent une ressource mais qu'ils font peser le coût de cette utilisation sur d'autres individus ou groupes d'individus. Par exemple, certaines activités industrielles rejettent dans la nature des substances toxiques. Si aucune mesure de régulation économique n'est prise, ces rejets dans la nature n'auraient pas de conséquences pour les entreprises à leur origine. En revanche, ce sont les riverains, les agriculteurs, les pêcheurs et autres utilisateurs des ressources polluées qui subiront les effets néfastes de ce comportement. Le « coût » (au sens économique du terme) de l'externalité (le rejet) est donc porté par d'autres individus que celui qui en est à l'origine. Ce dernier n'a donc aucune incitation naturelle à en réduire la quantité[678]. La solution, communément acceptée en économie, est « d'internaliser » l'externalité, par des mécanismes de marché, des régulations individuelles ou par une intervention législative. Dans le cas de la régulation environnementale, ces mesures peuvent être d'ordre économique (quota d'émissions de gaz à effet de serre[679]) ou juridique (interdiction d'utiliser certains produits polluants, seuils à respecter sous peine d'amende[680], obligation de transparence pour guider les consommateurs vers les entreprises vertueuses[681]…).

332. **La nécessité d'une régulation économique du développement de l'IA.** Dans le cas de l'exploitation des données par des systèmes d'IA, il existe plusieurs externalités négatives à corriger. Lorsque des entreprises collectent et exploitent des données personnelles, la part de vie « privée » des individus concernés diminue. En l'espèce, les conséquences de la collecte ne sont subies que par les utilisateurs, elle est donc bien externe et les entreprises ne sont pas incitées (si ce n'est par l'instauration de règles sur la protection des données) à prendre en compte ce coût « caché ». De même, le fonctionnement des systèmes d'IA génère des émissions indirectes de gaz à effet de serre, liées à l'alimentation en énergie électrique ou à la fabrication

[677] D.D. Hirsch, « Protecting the Inner Environment: What Privacy Regulation Can Learn from Environmental Law », *Georgia Law Review*, 2006, vol. 41, n°1, spec. p. 23.
[678] J. Salzman, B.H. Thompson, *Environmental law policy*, Foundation Press, 3rd ed., 2010, 397p., spec. 17-18 ; R.H. Coase, « The nature of the firm », *Economica*, novembre 1937, vol. 4, n° 16, 386-405.
[679] Y. Petit, « Système d'échange de quotas d'émission de gaz à effet de serre dans la Communauté », *Répertoire de droit européen*, Dalloz, Environnement – Protection des milieux, janvier 2007, 220 et s.
[680] Pour des exemples d'interdiction de l'utilisation de certains produits ou de mise en place de seuils d'émission dans le cadre la lutte contre la pollution atmosphérique, voir : A. Mesnard, « Environnement : protection contre les pollutions et nuisances », in *Encyclopédie des collectivités locales*, Dalloz, 1999, folio 5450, 130–148.
[681] F.G. Trébulle, « Responsabilité sociale des entreprises : entreprise et éthique environnementale », *Répertoire des sociétés*, Dalloz, mars 2003 (actualisation novembre 2022).

des terminaux physiques. Ce coût de l'activité est porté par la société dans son ensemble (aggravation du changement climatique) et non par l'entreprise fournissant le système. Enfin, un dernier exemple peut être donné avec le risque de discrimination lié à l'automatisation de certaines prises de décision portant sur des personnes. Certains systèmes d'IA peuvent conduire à la reproduction de biais discriminatoires pour des minorités[682]. Le risque de discrimination peut ne concerner que certaines minorités ethniques, politiques ou religieuses. Dans ces circonstances, le risque lié à l'utilisation de l'IA ne pèserait pas sur les utilisateurs dans leur ensemble, mais seulement sur une minorité d'entre eux. Ce constat peut expliquer pourquoi il est si difficile de mobiliser largement sur les risques juridiques liés aux usages de l'IA, tant ils n'incombent qu'à une minorité d'individus. En termes économiques, on peut considérer que ce risque est une externalité négative pour l'entreprise ayant développé le système d'IA, puisque son coût est porté par un tiers (en l'occurrence, la minorité discriminée). À terme, si l'externalité négative n'est pas réinternalisée grâce à des mesures de régulation, c'est l'intégralité des systèmes d'IA qui pourraient être rejetés par la société, quand bien même de nombreuses applications pourraient lui être bénéfiques. En effet, la confiance dans les technologies d'IA diminue à chaque scandale éthique et, à moins de créer un cadre juridique capable de la préserver, il y a de grandes chances que cette confiance finisse par s'épuiser[683].

333. **Un objectif commun de conciliation entre innovation et prévention des risques.** Par ailleurs, les mesures de régulation, en particulier en matière environnementale et financière, ont également pour objectif (ou contrainte) de trouver un équilibre entre la promotion de l'innovation et la prévention des risques.

334. **L'objectif de conciliation dans la régulation environnementale.** En ce qui concerne la protection de l'environnement, certains considèrent que le progrès technologique peut être

[682] J. Angwin, J. Larson, S. Mattu, L. Kirchner, « Machine Bias : There's software used across the country to predict future criminals, and it's biased against blacks », *ProPublica*, 2016, disponible en ligne : <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>, consulté le 17 mars 2020.

[683] P.P. Swire, R.E. Litan, « None of your business: world data flows, electronic commerce, and the European privacy directive », *Brookings inst. Press*, 1998, p. 8 ; sur l'épuisement d'une ressource perçue comme un bien commun (en l'espèce la confiance des individus dans la technologie) : G. Hardin, « The Tragedy of the Commons », *Science*, 1968, vol. 162, 1243.

source de bénéfices pour la société[684]. En effet, il conduit, ou a conduit, à l'invention de nouveaux moyens renouvelables de produire de l'énergie tels que l'éolien, le solaire ou encore l'hydrogène, la modernisation des procédés existants pour en optimiser les coûts (financiers et environnementaux) de fonctionnement, ou la découverte de nouvelles méthodes de captation du CO2. D'autres estiment au contraire que le progrès technologique coûte plus à la société que ce qu'il ne lui apporte[685]. D'un point de vue objectif et sans adopter de position partisane, il semble nécessaire que le Droit de l'environnement poursuive un double objectif : règlementer les activités industrielles pour limiter l'impact de l'homme sur son environnement, tout en n'entravant pas sa capacité à innover pour découvrir de nouvelles solutions technologiques moins néfastes pour l'environnement[686].

335. **L'objectif de conciliation dans la régulation financière.** Dans le secteur financier, le développement technologique permet l'émergence de nouveaux services pour les consommateurs, une diversification de l'offre et, plus généralement, tous les bénéfices attendus de la libre concurrence dans une logique libérale. La dérèglementation du secteur[687] traduit la prise de conscience de la nécessité de proportionner l'intervention législative. Jusqu'alors, la priorité des autorités normatives était de prévenir le risque systémique, *via* une règlementation prudentielle mais, aujourd'hui, on retrouve bien le double objectif de prévention des risques d'une part, et de promotion de l'innovation.

336. **L'objectif de conciliation dans la régulation relative à la protection de la vie privée.** La même logique se retrouve dans le RGPD, se voulant neutre technologiquement. Le double objectif se retrouve d'ailleurs clairement dans ses considérants, présentant le droit à la protection des données personnelles comme *« un droit fondamental »*[688] tout en lui déniant tout caractère *« absolu »*[689] et en reconnaissant que l'objectif poursuivi est principalement de *« de*

---

[684] D. Maier, A. Maier, I. Aşchilean, *et al.*, « The Relationship between Innovation and Sustainability: A Bibliometric Review of the Literature », *Sustainability*, 2020, vol. 12, n°10, 4083.
[685] A. Vallée, *Economie de l'environnement*, Editions du Seuil, 2011, nouvelle édition, 299 p.
[686] O. Boiral, « Concilier environnement et compétitivité, ou la quête de l'éco-efficience », *Revue française de gestion*, 2005, vol. 5, n°158, pp. 163-186.
[687] N. Mathey, G. Bourdeaux, « Vers une régulation des FinTechs ? », *Revue de Droit bancaire & financier*, mars 2017, vol. 2, dossier 15.
[688] RGPD, Considérant 1.
[689] RGPD, Considérant 4.

susciter la confiance qui permettra à l'économie numérique de se développer dans l'ensemble du marché intérieur »[690].

337. **Conclusion sur les similitudes dans les objectifs des modèles de régulation.** La recherche d'une proportionnalité entre prévention des risques et promotion de l'innovation est donc un second point commun aux cadres juridiques relatifs à la protection de l'environnement, à la stabilité du système financier et aux traitements de données à caractère personnel. Ces régimes répondent aux mêmes objectifs que ceux visés dans la régulation de l'IA, à savoir la correction d'une défaillance de marché conduisant à la production de systèmes non souhaitables pour la société ainsi que la mise en balance entre risques et bénéfices de la technologie. Ainsi, il peut être pertinent de s'inspirer des mécanismes juridiques mobilisés dans chacun d'eux pour bâtir le cadre juridique le plus adapté.

338. **Conclusion du §2 sur l'expérience de la régulation comme source d'inspiration.** Les développements précédents ont permis de démontrer les similitudes existantes entre la nature des risques générés par l'utilisation de l'IA et des risques environnementaux, liés à la stabilité financière ou encore liés au traitement des données à caractère personnel. Aussi, les règlementations nées pour répondre à ces trois derniers risques poursuivent des objectifs communs avec ceux de la nécessaire régulation de l'IA. Ces points communs justifient notre intérêt pour ces matières, pouvant servir de source d'inspiration à notre réflexion sur le cadre juridique nécessaire pour réguler l'IA.

339. **Conclusion de la Section 2 sur le choix du mode de régulation.** Le choix du mode de régulation n'est pas aisé tant les options à disposition des autorités normatives sont nombreuses. Sur son champ, la régulation de l'IA doit être à même de couvrir l'ensemble des risques que son utilisation engendre, tout en étant adaptée à chaque secteur d'activité tant ces risques diffèrent suivant l'utilisation concernée. Des principes communs, horizontaux, pourraient guider, dans une logique de subsidiarité, les autorités sectorielles de régulation dans leur propre encadrement de la technologie. En effet, ce sont elles, en coopération avec les acteurs régulés,

[690] RGPD, Considérant 7.

qui sont le mieux placées pour identifier les cas d'usage les plus dangereux et les mesures à prendre pour en limiter les risques. De plus, l'objectif de la régulation étant de stimuler l'innovation tout en prévenant les risques, une règlementation stricte sur tous les aspects n'est pas souhaitable. Il convient de penser une combinaison d'approches contraignantes et non-contraignantes pour parvenir à un cadre le plus proportionné possible. Afin de réaliser les choix dans les instruments juridiques à utiliser et les mécanismes de régulation à mettre en place, l'étude de secteurs régulés présentant des similitudes avec la situation du développement de l'IA nous apparaît nécessaire et pertinente.

340. **Conclusion Chapitre 1 sur la nécessaire création d'un cadre juridique pour l'IA.** Plusieurs raisons peuvent pousser les autorités normatives à réguler une activité. Cela peut être, par exemple, le défaut de fonctionnement d'un marché, auquel répondront des mesures de régulation économique, ou l'existence de risques inacceptables pour la société, pouvant être gérés par une règlementation stricte.

Le développement de l'IA dans tous les secteurs d'activité peut générer des risques d'atteinte pour les droits et libertés des individus, plus ou moins graves, suivant le cas d'usage concerné. Ces atteintes peuvent avoir trait à la sécurité des personnes, avec le cas des dommages causés par des systèmes autonomes, à leur vie privée, avec des capacités de traitement de données décuplées, ou encore à leur environnement. L'existence de risques est avérée et fait consensus. En parallèle, les systèmes d'IA peuvent conduire à d'immenses progrès pour la société s'ils sont utilisés à bon escient (aux fins de réduire les émissions de gaz à effet de serre, de prévision de catastrophes naturelles, d'optimisation des ressources…).

Ce double constat nous pousse à dire que l'IA doit faire l'objet d'une nouvelle régulation. Or, lorsqu'un besoin de régulation est avéré, les autorités normatives disposent de nombreux instruments juridiques pour la mettre en œuvre, allant des mesures de droit souple non contraignantes aux règlementations les plus strictes. Au sein-même de chacune de ces catégories, de multiples possibilités s'offrent aux régulateurs : mécanismes d'incitation fiscale, encadrement de la conception d'une technologie, interdiction de l'utilisation de certains produits, standardisation… A cette multitude d'options de régulation s'ajoute également la question du champ d'une éventuelle règlementation : sectorielle ou transversale, nationale ou extraterritoriale… Bien entendu, l'ensemble de toutes ces mesures ne sont pas exclusives et peuvent, voire doivent, être combinées. Le choix des moyens juridiques pour réguler l'IA n'est donc pas aisé, si bien que le juriste, empreint d'une certaine forme de conservatisme, peut chercher à s'inspirer de l'existant.

La situation n'est pourtant pas nouvelle. Les autorités normatives ont déjà eu, par le passé, à réguler des situations présentant un certain nombre de similitudes avec le développement actuel de l'IA et les risques qu'il engendre. En particulier, nous avons identifié trois corpus juridiques qui ont été créés pour des raisons similaires à celles qui nous poussent à proposer un cadre pour l'IA. L'étude du Droit de l'environnement, en raison de la nature sociétale, nouvelle et complexe des risques qu'il entend encadrer, du Droit financier, en réponse aux risques systémiques évoluant rapidement selon les avancées technologiques, et des régimes de protection des données à caractère personnel, traitant également des risques pour la vie privée, devrait nous permettre de saisir les avantages et inconvénients des différentes options de régulation.

341. **Transition.** S'intéresser aux mesures mises en œuvre et aux options de régulation retenues dans ces précédents peut être utile pour déterminer les instruments juridiques les plus pertinents pour réguler l'IA[691], ce qui est l'objet du Chapitre suivant.

[691] N.A. Smuha, « Beyond the Individual: Governing AI's Societal Harm », *Internet Policy Review*, 30 septembre 2021, vol. 10, n°3, spec. p. 9.

Chapitre 2 : Des moyens juridiques appropriés à la création d'une régulation proportionnée

342. **Plan.** De nombreuses leçons peuvent être tirées de l'expérience de la régulation. En particulier, la régulation de l'IA pourrait s'inspirer de certaines approches ou mécanismes du droit de l'environnement, de la protection des données ou de la régulation financière, puisque les risques adressés et les objectifs poursuivis présentent des similitudes[692]. Parmi tous les mécanismes juridiques employés dans ces secteurs, plusieurs nous semblent particulièrement utiles et pertinents dans le cadre de la régulation de l'IA.

En premier lieu, dans les domaines précités prévaut une approche de co-régulation alliant règlementation contraignante et incitations à l'autorégulation. Cette approche, visant à responsabiliser les acteurs régulés par la mise en place de mécanismes de conformité, à les impliquer dans l'édiction des normes, nous paraît adaptée aux enjeux de la régulation de l'IA (**Section liminaire**). Si le principe de la co-régulation semble aujourd'hui s'imposer comme une norme dans les nouvelles régulations au niveau européen, les mécanismes juridiques concrets pour la réaliser sont nombreux. Exigences de transparence, promotion des codes de bonne conduite, ou autres mécanismes de marché incitatifs tels que des taxes ou la diffusion de normes non contraignantes sont autant d'exemples qui pourraient être répliqués dans la régulation de l'IA. Il s'agira alors de présenter nos propositions de moyens juridiques permettant la responsabilisation des acteurs (**Section 1**).

Dans un deuxième temps, la majorité des nouvelles régulations ont été accompagnées par la création d'autorités de contrôle et de supervision. Ces autorités ont pour mission d'assurer la mise en œuvre de la régulation par l'accompagnement des acteurs régulés dans leur mise en conformité, l'édiction de standards sectoriels, l'information du public ou, le cas-échéant, la sanction des comportements contraires aux principes contenus dans la régulation. Nécessairement sectorielles, leur existence garantit que la régulation réponde de façon pertinente aux contraintes et attentes des acteurs régulés, tout en s'assurant qu'elle remplit bien les objectifs poursuivis, notamment en termes de prévention des risques et de lutte contre les

[692] Voir Supra, 325-337.

comportements déviants. Au vu du haut niveau de technicité des systèmes d'IA et de la nécessaire harmonisation des bonnes pratiques dans leur conception et leur développement, la création d'une telle autorité de régulation apparaît pertinente, si tant est que sa structure, son champ de compétence et ses pouvoirs lui permettent d'accomplir les missions susmentionnées (**Section 2**).

Enfin, au vu de la nature sociétale des risques générés par les systèmes d'IA et des transformations qu'ils amènent, le public doit avoir un rôle important dans leur régulation afin de répondre à la nécessité d'une gouvernance démocratique de la technologie[693]. À cet égard, l'expérience de la régulation montre que la reconnaissance de droits aux individus est une solution efficace[694]. Qu'ils prennent la forme de « garantie d'accès à la justice » et « droit à l'information du public » en droit de l'environnement ou « droits des personnes » dans le RGPD, ces droits individuels permettent de rendre effectives des règles et d'impliquer les citoyens dans la régulation d'une activité présentant un risque sociétal. Dans le cadre de la régulation de l'IA, de tels droits garantiraient aux individus un véritable contrôle sur une technologie qui, bien souvent, leur sera imposée (**Section 3**).

[693] Sur l'importance accordée à la gouvernance démocratique de l'IA par les institutions européennes, voir GROUPE D'EXPERTS DE HAUT NIVEAU EN IA, *Lignes directrices du 8 avril 2019 pour une IA digne de confiance*, 8 avril 2019, 2 : « *it is through Trustworthy AI that we, as European citizens, will seek to reap its benefits in a way that is aligned with our foundational values of respect for human rights, **democracy** and the rule of law* » ; GROUPE D'EXPERTS DE HAUT NIVEAU EN IA, *Policy and Investment Recommendations for Trustworthy AI*, rapport, 26 juin 2019, p. 37 : « *Ensuring Trustworthy AI necessitates […] a framework that promotes socially valuable AI development and deployment, ensures and respects fundamental rights, the rule of law, and **democracy** while safeguarding individuals and society from unacceptable harm* » ; N.A. Smuha, E. Ahmed-Rengers, A. Harkens, *et al.*, « How the EU can achieve legally trustworthy AI : a response to the European Commission's proposal for an Artificial Intelligence Act », *SSRN Electronic Journal*, 5 août 2021, 64.

[694] Pour un exemple concret avec la création de « droits individuels environnementaux » au Canada, voir D.D. Babor, « Environmental Rights in Ontario: Are Participatory Mechanisms Working », *Colorado Journal of International Environmental Law and Policy*, 1999, vol. 10, n°1998, p. 121-135.

Section liminaire : La pertinence d'un modèle de co-régulation

343. **Définitions.** Dans une logique de régulation, le cadre institutionnel fournit un spectre d'alternatives, allant des règles les plus précises et contraignantes aux mesures de droit souple. On distingue classiquement trois approches de régulation : la règlementation, l'autorégulation et la co-régulation. Il convient de définir les deux premières afin de comprendre la troisième. Ces notions ont été abordées et définies notamment dans l'accord européen interinstitutionnel « Mieux légiférer » publié en 2003[695], qui constitue une source essentielle pour saisir l'approche européenne de la régulation. En premier lieu, la règlementation contraignante correspond à une approche positiviste du Droit et consiste en l'édiction de règles précises pour encadrer les comportements individuels et collectifs[696]. La conformité des comportements aux règles édictées peut être contrôlée soit par les tribunaux[697], soit par des autorités administratives[698]. Ensuite, l'autorégulation consiste en « *la possibilité pour les opérateurs économiques, les partenaires sociaux, les organisations non gouvernementales ou les associations, d'adopter entre eux et pour eux-mêmes des lignes directrices communes au niveau européen (notamment codes de conduite ou accords sectoriels)* »[699]. Il s'agit donc d'une approche « par le bas » ou « *bottom-up* », par laquelle les opérateurs économiques édictent eux-mêmes les normes, de façon unilatérale, par le biais de charte éthique individuelle par exemple, ou multilatérale avec l'adoption de normes collectives[700]. Enfin, l'approche de la co-régulation se situe entre les deux premières. Elle est définie par les institutions européennes comme « *le mécanisme par lequel un acte législatif […] confère la réalisation des objectifs définis par l'autorité législative aux*

---

[695] *Accord interinstitutionnel 2003/C 321/01 du 13 avril 2016 entre le Parlement européen, le Conseil de l'Union européenne et la Commission européenne « Mieux légiférer »*, publiée au JOUE n°L123/1 le 12 mai 2016.
[696] L'idée rejoint notamment la vision positiviste d'Hans Kelsen sur le Droit : H. Kelsen, *General Theory of the Law and the State*, Harvard University Press, 1949, 544 p.
[697] C'est le cas des lois générales sur la propriété et la responsabilité, par exemple les règles de la propriété intellectuelle contenues dans le Code de la Propriété intellectuelle ou les régimes de responsabilité civile du Code civil. Leur caractère coercitif est garanti par l'ordre juridictionnel judiciaire.
[698] C'est le cas, notamment, de la règlementation en matière de protection des données à caractère personnel, contenue dans une loi et un règlement européen et mise en œuvre par une autorité administrative indépendante, la CNIL. Le Droit de l'environnement ou la règlementation financière peuvent également être pris en exemple.
[699] *Accord interinstitutionnel 2003/C 321/01 du 13 avril 2016 entre le Parlement européen, le Conseil de l'Union européenne et la Commission européenne « Mieux légiférer »*, publiée au JOUE n°L123/1 le 12 mai 2016, pt 22.
[700] ARBORUS, ORANGE, Charte internationale pour une IA inclusive, disponible en ligne : <https://charteia.arborus.org/>, consulté le 25 janvier 2022.

parties concernées […] »[701]. Elle permet donc aux pouvoirs publics de déléguer aux opérateurs économiques une partie de la responsabilité de la mise en application de la norme. La doctrine internationale a ainsi pu la qualifier de véritable « entre-deux »[702] de régulation. Elle se traduit souvent par la notion de « responsabilisation » des acteurs régulés[703], prégnante en matière de protection des données notamment[704], et qui peut être concrétisée de différentes manières. L'expérience de la régulation laisse penser que les deux premières approches – prises séparément – ne sont pas adaptées pour créer un cadre juridique favorable à l'innovation éthique et écologique dans le domaine de l'IA.

344. **Plan.** L'approche prescriptive (**§1**) et l'autorégulation (**§2**), considérées isolément, ne nous semblent pas adaptées dans le cadre de la régulation de l'IA. Toutefois, leur combinaison, dans une logique de co-régulation, nous apparaît comme la solution la plus pertinente (**§3**).

§1 : Le manque d'adaptabilité d'une règlementation stricte

345. **Un modèle de régulation déjà expérimenté.** La question de la nécessité, ou non, d'une règlementation contraignante pour réguler des technologies présentant des risques pour la société s'est déjà posée par le passé. En particulier, l'histoire du Droit de l'environnement sur le continent américain montre que cette approche, seule, n'est pas la plus adaptée pour réguler une technologie comme l'IA. En effet, lorsqu'il a fallu, dans les années 70, adopter les premières mesures en matière environnementale, les États-Unis ont opté pour une approche descendante et prescriptive de la régulation, appelée *« Command and control »* ou *« First generation of environmental law »*[705]. Après 25 ans de pratique et à la suite d'un rapport très

[701] *Accord interinstitutionnel 2003/C 321/01 du 13 avril 2016 entre le Parlement européen, le Conseil de l'Union européenne et la Commission européenne « Mieux légiférer »*, publiée au JOUE n°L123/1 le 12 mai 2016, pt 18.
[702] U. Pagallo, P. Casanovas, R. Madelin, « The Middle-out Approach: Assessing Models of Legal Governance in Data Protection, Artificial Intelligence, and the Web of Data », *The Theory and Practice of Legislation*, 2019, vol. 7, n°1, 25.
[703] J. Sénéchal, « Responsabilisation ab initio, régulation ex ante et responsabilités a posteriori : le coeur des débats européens sur les systèmes d'intelligence artificielle, hors et dans le secteur du commerce électronique », *Dalloz IP/IT*, 2020, p. 667.
[704] L. Marignol, « Principe de responsabilité et action en responsabilité dans le Règlement général sur la protection des données », *Revue Lamy Droit de l'Immatériel*, 1er janvier 2020, n°166.
[705] D. D. Hirsch, « Protecting the Inner Environment: What Privacy Regulation Can Learn from Environmental Law », *Georgia Law Review*, 2006, vol. 41, n°1, p.10.

influent publié par le Congrès américain[706], les États-Unis ont changé d'approche et adopté des mesures de co-régulation, plus souples et impliquant les acteurs régulés dans la recherche des meilleures solutions pour diminuer l'impact environnemental des activités industrielles notamment. Cette nouvelle approche sera connue en doctrine sous le nom de « *Second generation of environmental law* ». Les conclusions du rapport du Congrès américain et diverses contributions académiques ont permis de mettre en évidence les limites de l'approche contraignante et prescriptive. Cette dernière consistait, en résumé, en la désignation par les pouvoirs publics des secteurs considérés comme trop polluants, puis en l'identification par un régulateur sectoriel des meilleures technologies disponibles pour réduire les émissions et enfin en l'obligation des acteurs à utiliser ces technologies (« *design based regulation* ») ou à ne pas dépasser certains seuils d'émissions (« *rate-based regulation* »)[707].

346. **Les enseignements de l'expérience en Droit de l'environnement.** Les enseignements qui peuvent être tirés de cette expérience au regard de l'approche contraignante et prescriptive de la régulation sont nombreux. D'abord, la règlementation contraignante présente l'avantage d'être relativement facile à mettre en œuvre et d'être la plus efficace pour atteindre des résultats précis (en l'occurrence une baisse des émissions de gaz à effet de serre dans les secteurs concernés). Toutefois, l'approche prescriptive s'avère **très coûteuse**, à la fois pour les acteurs régulés qui n'ont pas le choix dans les technologies auxquels ils doivent recourir (peu importe la taille de l'entreprise concernée) et pour les pouvoirs publics qui doivent disposer des moyens suffisants pour définir les standards et contrôler leur respect. Enfin, elle semble **inadaptée à l'innovation** en raison du temps long de la définition des standards (qui peut les rendre obsolètes au moment de leur sortie ou incapables de changer rapidement en cas d'avancée technologique) et du fait qu'elle n'incite, ni ne permet, aux entreprises de dépasser les objectifs fixés ou d'innover dans la façon de les atteindre[708]. En somme, l'approche prescriptive n'est

[706] *Ibid.* ; U.S. CONGRESS, OFFICE OF TECHNOLOGY ASSESSMENT, *Environmental Policy Tools: A User's Guide*, U.S. Government Printing Office, OTA-ENV-634, Washington DC, September 1995, 224p.
[707] Cette approche est aussi désignée dans la doctrine « Best available technologies regulation » : B.A. Ackerman, R.B. Stewart, « Reforming Environmental Law », *Stanford Law Review, 1985*, vol. 37, n°1333 ; E.W. Orts, « A Reflexive Model of Environmental Regulation », octobre 1995, *Business Ethics Quarterly*, vol. 5, n°4, pp. 779-794.
[708] D.D. Hirsch, *op. cit.*, p.32 ; V. aussi C. Rechtschaffen, « Deterrence vs. Cooperation and the Evolving Theory of Environmental Enforcement », *Southern California Law Review*, 1997, n°71, pp. 1181-1272.

adaptée que pour des dangers précisément identifiés, quantifiables et suffisamment importants pour justifier la mise en place d'une régulation coûteuse et difficilement adaptables[709]. Elle ne semble ainsi pas pertinente dans le contexte de la régulation de l'IA, une technologie encore peu mature et dont les usages évoluent très rapidement.

En effet, l'IA requiert une régulation plus souple, en mesure de s'adapter à la découverte de nouveaux risques ou au développement de nouveaux cas d'usage.

§2 : Le défaut d'effectivité de l'autorégulation

347. **L'utilité de l'autorégulation.** L'autorégulation, quant à elle, présente l'avantage de la souplesse et de l'adaptabilité. Les opérateurs économiques sont les plus fins connaisseurs de leur activité, il apparaît logique qu'ils soient les mieux placés pour identifier les solutions, méthodes et technologies les moins coûteuses et les plus adaptées pour répondre à un objectif fixé (tel que la réduction d'un risque). Les Professeurs Bensamoun et Loiseau considèrent également que l'autorégulation permet « *une efficacité et une acceptation plus large de la norme par les milieux concernés* »[710]. Un exemple d'autorégulation parfaitement efficace peut être trouvé dans le droit des fusions-acquisitions pratiqué dans la City, à Londres[711]. La pratique a été érigée en norme et n'a jamais été codifiée par les pouvoirs publics. Pourtant, toute entreprise qui ne respecterait pas les normes implicites de la pratique londonienne se verrait sanctionnée par ses pairs, notamment par la pratique du « *Cold shoulder* »[712]. Ainsi, dans cet exemple, tant la construction de la norme que sa mise en œuvre et sa sanction émanent

[709] C. Coglianese, « The Limits of Performance-Based Regulation », *University of Michigan Journal of Law Reform*, 2017, vol. 50, n°3, pp. 525-564.
[710] A. Bensamoun, G. Loiseau, « L'intégration de l'intelligence artificielle dans l'ordre juridique en droit commun : questions de temps », *Dalloz IP/IT*, 2017, 239.
[711] *The City Code on Takeovers and Mergers*, 13ème édition, 5 juillet 2021, disponible en ligne : < https://www.thetakeoverpanel.org.uk/the-code/download-code>, consulté le 1er février 2022 ; D. Kershaw, , *Principles of takeover regulation*, Oxford University Press, 2016, 418 p.
[712] Dans la pratique des fusions-acquisitions dans la City à Londres, une entreprise ne respectant pas les normes fixées par ses pairs peut se voir « blacklistée » par eux, sur décision d'une autorité d'auto-régulation telle que le *Securities and Investment Board*. La nature de la sanction, non étatique, produit alors un effet hautement dissuasif car elle prive l'opérateur de futures opportunités économiques et de partenaires essentiels à son activité s'il ne respecte pas les règles fixées par ses pairs : The City Code on Takeovers and Mergers, *op. cit.*, Introduction, 11(b)(v) ; T. Shea, « Regulation of Takeovers in the United Kingdom », *Brooklyn Journal of International Law*, 1990, vol. 16, n°1, p. 89.

directement des entreprises. L'autorégulation s'appuie sur d'autres mécanismes de régulation que la règle de droit tels que le pouvoir du marché ou le risque réputationnel pour les entreprises. En principe, sa souplesse en fait une approche adaptée à la régulation des secteurs qui évoluent rapidement, tels que le développement technologique.

348. **Les limites de l'autorégulation.** Toutefois, elle présente plusieurs limites qui nous font penser qu'elle ne peut pas assurer seule la régulation du développement de l'IA. En effet, à défaut de sanction effective du non-respect de la norme, elle favorise l'émergence de comportements de « *passagers clandestins* »[713]. On peut déjà l'observer dans le domaine de l'IA, où se multiplient les chartes éthiques non contraignantes en vue d'éviter une règlementation contraignante[714]. De plus, l'autorégulation ne peut fonctionner que s'il existe un large consensus entre les entreprises et les pouvoirs publics sur les risques à adresser et les objectifs de la régulation[715]. Or, les divergences observées dans le contenu des différentes chartes éthiques multilatérales[716] et la difficulté à parvenir à un large consensus international[717] laissent penser que l'autorégulation ne pourra jouer pleinement son rôle.

Il apparaît alors nécessaire de la combiner avec des approches plus contraignantes afin de pallier ses limites et le risque que des entreprises profitent de l'absence de sanction contraignante.

---

[713] H.J. Levin, « The Limits of Self-Regulation », *Columbia Law Review*, avril 1967, vol. 67, n° 4, pp. 603-644.
[714] R. Ochigame, « How Big Tech Manipulates Academia to Avoid Regulation », *The Intercept (blog)*, 20 décembre 2019, disponible en ligne : <https://theintercept.com/2019/12/20/mit-ethical-ai-artificial-intelligence/>, consulté le 23 janvier 2020.
[715] H.J. Levin, « The Limits of Self-Regulation », *op. cit.*
[716] A. Jobin, M. Ienca, E. Vayena, *op. cit.* ; Y. Meneceur, « Analyse des principaux cadres supranationaux de régulation de l'intelligence artificielle », 31 mai 2021, disponible en ligne : <https://lestempselectriques.net/ANALYSE_IA.pdf>, consulté le 20 janvier 2022.
[717] La divergence des approches européenne, américaine, russe et chinoise (C. Cath, S. Wachter, B. Mittelstadt, *et al.*, « Artificial intelligence and the 'Good Society' : the US, EU, and UK approach », *Science and Engineering Ethics*, 2018, vol. 24, 505-528 ; H. Roberts, J. Cowls, J. Morley, *et al.*, « The Chinese approach to artificial intelligence : an analysis of policy, ethics and regulation », *AI & Society*, 2021, vol. 36, 59-77) démontre l'impossibilité de compter sur une autorégulation des acteurs de l'IA au niveau international. Elle ne pourrait être pensée, au mieux, au niveau régional.

§3 : L'atteinte d'un équilibre grâce au modèle de la co-régulation

349. **La pertinence de la combinaison d'approches souples et contraignantes.** Les mécanismes d'autorégulation, ou de « droit souple », peuvent être favorisés par une intervention étatique. Comme le soulignent les Professeurs Bensamoun et Loiseau : *« L'idée n'est pas de substituer du droit mou au droit substantiel, mais de permettre, grâce à l'autorégulation, une efficacité et une acceptation plus large de la norme par les milieux concernés. Les mesures prudentielles peuvent encore être un préalable à un droit substantiel plus réfléchi et donc plus adapté, dans un secteur où l'influence de la technologie oblige à une flexibilité et à une réactivité importantes »*[718]. Le droit substantiel peut donc venir à l'appui de l'autorégulation, en créant des mécanismes qui encourageraient les entreprises à développer leurs propres normes (par exemple par la reconnaissance des codes de bonne conduite) ou qui favoriseraient la régulation des comportements par le marché (par exemple par l'obligation de divulguer au marché certaines informations pouvant orienter le comportement des investisseurs ou des consommateurs).

350. **L'expérience de la co-régulation.** L'idée de combiner les approches contraignantes et non-contraignantes pour réguler une activité n'est pas nouvelle et a déjà fait ses preuves dans d'autres secteurs. Elle se retrouve, d'abord, dans l'exemple du Droit de l'environnement européen[719] ou américain[720] à travers sa « régulation environnementale de seconde génération ». Le rapport du Congrès des États-Unis précité[721] concluait à cet égard que les mécanismes juridiques renversant la responsabilité de déterminer les moyens de la régulation aux entreprises ou à des groupements d'entreprises sont les plus à-mêmes d'aboutir à des solutions efficientes (au meilleur rapport coût – efficacité) aux questions climatiques[722]. Le principal argument avancé par la doctrine à cet égard est que les entreprises ont une meilleure

[718] A. Bensamoun, G. Loiseau. « L'intégration de l'intelligence artificielle dans l'ordre juridique en droit commun : questions de temps », *Dalloz IP/IT*, 2017, 239.
[719] N.A. Smuha, « Beyond the Individual: Governing AI's Societal Harm », *op. cit.*
[720] D.D. Hirsch, *op. cit.*, p. 37.
[721] U.S. CONGRESS, OFFICE OF TECHNOLOGY ASSESSMENT, *Environmental Policy Tools: A User's Guide*, *op. cit.*
[722] *Ibid.*

connaissance de leur activité et des moyens à leur disposition[723]. Après avoir fait ses preuves en Droit de l'environnement, l'idée de la responsabilisation a été reprise dans la construction du régime européen de la protection des données à caractère personnel[724]. L'article 24 du RGPD pose ainsi le principe selon lequel le responsable de traitement *« met en œuvre des mesures techniques et organisationnelles appropriées pour s'assurer et être en mesure de démontrer que le traitement est effectué conformément au présent règlement »*[725]. Il doit être en mesure de prouver la conformité des traitements dont il a la charge, tout en étant libre sur la nature des mesures mises en place. Le modèle de co-regulation construit dans le RGPD est ainsi fondé sur le principe *d'accountability* ou de responsabilisation des responsables de traitement. Ils ont la charge d'identifier eux-mêmes les risques inhérents aux traitements effectués et de trouver les moyens de les limiter, à la fois *ex-ante* en prenant en compte la protection de la vie privée dès la conception[726] et *ex-post* en mettant en place des mesures techniques et organisationnelles pour assurer leur sécurité et confidentialité[727]. Cette méthode permet de laisser suffisamment de souplesse aux acteurs régulés pour définir eux-mêmes les meilleurs moyens pour respecter les principes fixés par les textes[728]. En revanche, cela nécessite d'avoir un consensus sur les grands principes à respecter, ce qui n'est pas le cas dans le cadre de la régulation de l'IA à l'heure actuelle[729]. C'est pourquoi nous proposons de fonder la régulation juridique de l'IA non pas sur un corpus de valeurs morales mais plutôt sur le corpus des droits fondamentaux qui fait, lui, l'objet d'un consensus beaucoup plus large au niveau international[730].

---

[723] J. Salzman, « Creating Markets for Ecosystem Services: Notes from the Field », *New York University Law Review*, 2005, vol. 80, 870, spec. 887-888 ; V. aussi R.B. Stewart, « A New Generation of Environmental Regulation ? », *Capital University Law Review*, 2001, vol. 29, 21, spec. 38-151.

[724] Voir notamm. U. Pagallo, P. Casanovas, R. Madelin, « The Middle-out Approach: Assessing Models of Legal Governance in Data Protection, Artificial Intelligence, and the Web of Data », *The Theory and Practice of Legislation*, 2 janvier 2019, vol. 7, n°1, 25 ; J. Sénéchal, « Responsabilisation ab initio, régulation ex ante et responsabilités a posteriori : le coeur des débats européens sur les systèmes d'intelligence artificielle, hors et dans le secteur du commerce électronique », *Dalloz IP/IT*, 2020, p. 667.

[725] RGPD, article 24.

[726] Il s'agit des principes de *privacy by design* et de *privacy by default* (RGPD, article 25 « Protection des données dès la conception et protection des données par défaut »).

[727] RGPD, article 24.

[728] U. Pagallo, P. Casanovas, R. Madelin, *op. cit.* ; EUROPEAN PARLIAMENTARY RESEARCH SERVICE, *The impact of the GDPR on AI*, Rapport de recherche, Juin 2020, n° 641530.

[729] U. Pagallo, P. Casanovas, R. Madelin, *op. cit.*, p.12.

[730] K. Yeung, A. Howes, G. Pogrebna, « AI Governance by Human Rights-Centred Design, Deliberation and Oversight: An End to Ethics Washing », in *The Oxford Handbook of AI Ethics*, dir. M. Dubber, F. Pasquale, Oxford University Press, 2019, spec. p. 9.

351. **La nécessaire recherche d'un équilibre entre contrainte et souplesse règlementaires.** Il conviendra toutefois de toujours veiller à l'équilibre entre les approches souples et contraignantes pour ne pas parvenir à une règlementation trop exhaustive qui viendrait freiner l'innovation. Cet écueil a par exemple été observé dans le secteur financier, lequel fait désormais l'objet d'un mouvement de déréglementation[731].

352. **L'optimum de régulation.** Dans chacun des exemples cités ci-dessus, il n'est jamais question de déléguer intégralement la régulation aux opérateurs économiques. Chaque fois, le Droit cherche à trouver un équilibre entre des règles prudentielles contraignantes et des mécanismes d'incitation à l'autorégulation[732]. Il s'agit alors, pour les pouvoirs publics, de rechercher « l'optimum social » ou « l'optimum de régulation » chers aux économistes du Droit[733].

353. **Conclusion sur la pertinence du modèle de la co-régulation.** La co-régulation vise donc à responsabiliser les opérateurs économiques dans leur propre régulation par une conjonction d'approches souples et contraignantes. Comme nous avons pu le voir à travers l'expérience de la régulation, elle présente plusieurs avantages. Premièrement, elle permet d'aboutir à des règles au meilleur rapport coût-efficacité et de promouvoir l'innovation car les entreprises sont incitées à développer les meilleures solutions au meilleur coût pour optimiser leur rentabilité. Deuxièmement, les mécanismes de co-régulation s'affranchissent des lourds processus administratifs liés à la création de standards par un régulateur central, ce qui permet une plus grande souplesse et adaptabilité des normes. Cette méthode semble être la plus adaptée pour réguler un domaine aussi mouvant que celui de l'IA, qui peut évoluer très rapidement au gré des avancées scientifiques.

---

[731] Sur le phénomène de déréglementation des marchés financiers et ses justifications, V. notamm. C.-A. Dubreuil, « Règlementation des marchés financiers », *JurisClasseur Administratif*, LexisNexis, 2010, Fasc. 265, spec. 9.

[732] Il existe différents niveaux de co-régulation, allant des approches les plus souples se rapprochant de l'autorégulation, à la supervision d'une activité par une autorité indépendante composée de pairs. Sur l'échelle des différents niveaux de co-régulation, voir : C. Marsden, « Internet co-regulation and constitutionalism : Towards European judicial review », *International Review of Law Computers and Technology*, 2012, vol. 26, n°2, 211-228.

[733] W. Maxwell, *A Method to Assess Regulatory Measures Designed to Limit Access to Harmful Content on the Internet*, Thèse pour le doctorat en sciences économiques, Telecom ParisTech, 2016, spec. p. 87 et s., citant notamm. D. Helm, « Regulatory Reform, Capture, and the Regulatory Burden », *Oxford Review of Economics*, 2006, vol. 22, n°169.

354. **Transition.** Toutefois, on l'a vu, l'efficacité de la co-régulation dépend en grande partie de l'équilibre trouvé entre règles contraignantes et mécanismes incitatifs. Cet équilibre ne peut être trouvé qu'en sélectionnant intelligemment les mécanismes de responsabilisation à mettre en œuvre.

Section 1 : Les moyens de la responsabilisation des acteurs

355. **La diversité des moyens de la co-régulation ou de la responsabilisation des acteurs.** De nombreux mécanismes juridiques peuvent être utilisés pour aboutir à un cadre qui permettrait à la fois de prévenir les risques pour les droits fondamentaux et l'environnement liés aux usages de l'IA et d'inciter les entreprises à développer des systèmes d'IA vertueux. Les possibilités sont nombreuses et, encore une fois, l'expérience de la régulation permet d'avoir un aperçu des mécanismes auxquels il est possible d'avoir recours, leurs avantages et inconvénients. Ces moyens vont des plus prescriptifs tels que la règlementation précise de la conception d'une technologie[734] ou la mise en place de processus de certification obligatoire[735], aux plus souples tels que la mise en place de codes de bonne conduite non contraignants[736].

356. **Plan.** Face à la diversité de ces moyens, il convient de sélectionner ceux qui permettront de responsabiliser les acteurs sans entraver l'innovation pour aboutir à un cadre juridique proportionné. Deux moyens juridiques, expérimentés dans d'autres domaines par le passé, nous semblent particulièrement pertinents dans le cadre de la régulation de l'IA. Premièrement, la mise en place de mécanismes de conformité *ex-ante* semble être la meilleure option pour s'assurer que la technologie respecte la règle de Droit dès la conception (**§1**). Deuxièmement, l'implication des acteurs régulés dans l'édiction de ces normes permettrait de s'assurer de leur

---

[734] La « *design-based regulation* », voir Supra, 345.
[735] A. Grenard, « Normalisation, certification : quelques éléments de définition », *Revue d'économie industrielle*, 1996, vol. 75, pp. 45-60 ; CONSEIL DE L'EUROPE, *Possible introduction of a mechanism of certifying artificial intelligence tools and services in the sphere of justice and the judiciary*, Etude de faisabilité de la Commission européenne pour l'efficacité de la justice, 8 décembre 2020, CEPEJ(2020)15Rev, spec. pp. 5-10, disponible en ligne : <https://rm.coe.int/feasability-study-en-cepej-2020-15/1680a0adf4>, consulté le 8 février 2022 ; C. Galan, « The Certification as a Mechanism for Control of Artificial Intelligence in Europe », *SSRN Electronic Journal*, 11 septembre 2019, disponible en ligne : <https://ssrn.com/abstract=3451741>, consulté le 8 février 2022.
[736] P. Boddington, *Towards a Code of Ethics for Artificial Intelligence*, Springer, 2017, 124 p. ; L. Venema, « Code of conduct for using AI in healthcare », *Nature Machine Intelligence*, 2019, vol. 1, 265-266.

acceptabilité et de leur pertinence, notamment en incitant les entreprises à proposer elles-mêmes des solutions pour remplir les objectifs de la régulation (**§2**).

§1 : Une responsabilisation des acteurs par le recours à la conformité

357. **Principe de la conformité.** La mise en place de mécanismes de conformité *ex-ante* consiste en l'édiction de principes et méthodes à respecter lors de la phase de conception d'un produit ou service, avant son utilisation ou mise sur le marché[737]. En cela, elle consiste en une règlementation, pouvant être définie comme un mode d'encadrement de l'économie consistant en *« l'encadrement unilatéral des conduites par l'édiction de normes juridiques à l'objet varié »*[738]. Pour être pleinement efficace, elle est régulièrement couplée avec un examen obligatoire de la conformité avant la mise sur le marché ou utilisation[739] ou de certification[740]. Contraignante par nature, la conformité doit être limitée aux risques les plus importants et son contenu strictement proportionné à ce qui est nécessaire pour faire passer le risque sous le seuil de l'acceptabilité. On retrouve cette démarche dans de nombreux secteurs, en particulier dans la régulation prudentielle du secteur financier[741] ou dans la gestion des risques environnementaux[742].

358. **La délégation du contrôle de la conformité.** Dans une logique de co-régulation, la preuve de la conformité avec les normes et principes édictés unilatéralement par les pouvoirs publics peut être déléguée aux acteurs régulés. C'est la traduction concrète du principe de responsabilisation évoqué précédemment. Cette délégation permet d'éviter de créer de lourds

[737] L. Rapp, P. Terneyre, « Démarche de/par la conformité… diversité des expressions », *Lamy droit public des affaires*, 615-616 ; C. Boiteau, « L'entreprise régulée », *RFDA*, 2018, 469.
[738] J.-P. Colson, P. Idoux, *Droit public économique*, LGDJ, 4e éd., 2008, p. 185.
[739] Pour un exemple, voir : *Directive 2001/95/CE du Parlement européen et du Conseil du 3 décembre 2001 relative à la sécurité générale des produits*, publié au JOCE n°L11/4 du 15 janvier 2002.
[740] S. Bernatchez, « La certification en tant que droit de la gouvernance », *Ethique publique*, 2019, vol. 21, n°1.
[741] Dans le secteur financier, l'arrêté du 3 novembre 2014 a été le premier à donner une définition de la conformité : *Arrêté du 3 novembre 2014, relatif au contrôle interne des entreprises du secteur de la banque, des services de paiement et des services d'investissement soumises au contrôle de l'Autorité de contrôle prudentiel et de résolution*, publié au JORF n°0256 du 5 novembre 2015, NOR : FCPT1423259A ; C.-A., Dubreuil, *op. cit.*, spec. 12 et s.
[742] R. Vallier, *La conformité environnementale, une politique juridique au service de la performance globale de l'entreprise*, thèse pour le doctorat en Droit, Université Côte d'Azur, 357 p., spec. pp. 51-54.

mécanismes d'autorisation administrative préalable comme cela existe en matière environnementale[743] ou bancaire[744].

359. **L'encadrement de l'IA par la conformité.** Il est donc possible d'encadrer la conception des systèmes d'IA par des mécanismes de conformité, tout en prévoyant des modalités plus souples qu'une règlementation prescriptive supervisée par une autorité externe. Assujettir le développement des systèmes d'IA à certaines normes devrait permettre de les rendre compatibles à la fois avec les principes contenus dans le corpus législatif existant[745] et avec les droits fondamentaux tels que le respect de la vie privée ou l'autonomie individuelle. Il serait alors possible de fixer de grands principes de façon horizontale, relatifs, par exemple, à la prise en compte du risque de biais dès la conception, à la réalisation d'analyse d'impact pour identifier les risques pour les droits fondamentaux que le système d'IA pourrait générer, à la supervision humaine en toutes circonstances ou à l'utilisation des meilleurs moyens disponibles pour prévenir les risques de piratage. De façon subsidiaire, les modalités de la mise en conformité et les méthodes précises pour remplir ces objectifs devraient être déclinées sectoriellement par les autorités de régulation, avec la participation des entreprises qui sont les mieux placées pour identifier les meilleurs moyens pour atteindre les objectifs fixés dans la régulation horizontale.

360. **Plan.** Dans le but d'assurer la proportionnalité de la régulation de l'IA et afin de ne pas freiner l'innovation, il convient de bien circonscrire le champ d'application des exigences de conformité. En effet, au vu de la diversité des systèmes d'IA et de leurs finalités, il semble peu opportun d'imposer de lourdes obligations à tous les fournisseurs. Au contraire, il convient de concentrer les contraintes sur les applications qui posent de réelles menaces pour les droits et libertés des individus et d'alléger la charge règlementaire pour les systèmes les moins dangereux, en adoptant une approche « par les risques » (**A**). De plus, les exigences règlementaires pour les systèmes méritant la plus grande attention doivent permettre à la fois de prévenir les risques pour les droits et libertés des individus, et de ne pas imposer une contrainte insurmontable pour les entreprises. C'est pourquoi nous proposons dans notre thèse

[743] Code de l'environnement, articles R122-1 et s.
[744] Code monétaire et financier, articles L511-9 et s.
[745] Tels que les régimes de responsabilité civile en droit français, voir Supra 85 et s.

plusieurs propositions d'exigences de conformité semblant raisonnables au vu des bonnes pratiques utilisées dans les entreprises et de l'état de l'art en la matière (**B**).

A/ Des exigences de conformité fondées sur une approche par les risques

361. **La notion de risque appliquée à l'IA.** Parler des risques liés à l'utilisation des systèmes d'IA nécessite dans un premier temps de définir cette notion et de la distinguer de ses concepts voisins. Pour ce faire, l'expérience en matière de régulation des risques technologiques, dans le domaine de la sécurité et de l'environnement nous semble particulièrement pertinente[746]. Le risque peut y être défini comme la « *combinaison de la probabilité d'un dommage et de sa gravité* »[747] ou comme la « *possibilité de survenance d'un dommage résultant d'une exposition aux effets d'un phénomène dangereux* »[748]. Le danger, quant à lui, peut se définir dans le même contexte comme « *la propriété intrinsèque […] d'un système technique de nature à entraîner un dommage […]* »[749]. À ce titre, certaines utilisations de l'IA pourraient être considérées comme « dangereuses » en ce qu'elles peuvent causer des dommages, physiques ou immatériels (financiers, psychologiques…), aux individus. Ils peuvent, en l'occurrence, constituer des atteintes aux droits et libertés des individus (atteinte à la propriété, à l'intégrité physique, à la vie privée). Si l'on s'en tient à cette définition, de nombreux systèmes techniques, dont une majorité des usages de l'IA, peuvent être considérés comme « dangereux ». Toutefois, un système « dangereux » ne présente pas forcément un « risque » élevé. En effet, il est nécessaire de recourir à la notion de « risque » pour déterminer les usages et techniques qui doivent concentrer l'attention des régulateurs. Ce raisonnement constitue le fondement de ce que l'on

---

[746] Sur l'utilité de la notion de risque du point de vue du droit et de la régulation, V. notamm. M.A. Frison-Roche, « Le droit de la régulation », *Recueil Dalloz*, 2001, 610 ; L. Boy, « Réflexions sur le 'droit de la régulation' », *Recueil Dalloz*, 2001, 3031 ; J. Chevallier, « Vers un droit post-moderne ? Les transformations de la régulation juridique », *Revue du droit public et de la science politique*, Librairie Générale de Droit et de Jurisprudence, 1998, 659–714.

[747] Norme ISO/CEI 51:2014, *Aspects liés à la sécurité – Principes directeurs pour les inclure dans les normes*.

[748] *Circulaire du 10 mai 2010 récapitulant les règles méthodologiques applicables aux études de dangers, à l'appréciation de la démarche de réduction du risque à la source et aux plans de prévention des risques technologiques (PPRT) dans les installations classées en application de la loi du 30 juillet 2003*, publiée au BO du MEEDDM n° 2010/12 du 10 juillet 2010.

[749] *Ibid.*

appelle la « régulation des risques » ou « approche par les risques », méthode de régulation utilisée notamment dans le domaine de l'environnement et des risques technologiques (notamment dans les industries pétrolière, aéronautique ou de la santé publique). Cette approche nous semble particulièrement pertinente pour notre réflexion sur la régulation de l'IA. Elle impose, d'abord, de qualifier les composantes du risque (**1**) puis consiste, ensuite, en la mise en œuvre d'une méthodologie précise, dont chacune des étapes pourrait utilement être appliquée au développement de l'IA (**2**).

1. La caractérisation du « risque » lié à l'utilisation de systèmes d'IA

362. **La régulation des risques : une condition à l'acceptation d'une nouvelle technologie.** De nombreux enseignements peuvent être tirés de l'expérience de la régulation des risques industriels et technologiques. En effet, le risque est inhérent à toute activité industrielle, technologique ou humaine[750] et la question de sa régulation lors de l'émergence d'une nouvelle technologie n'est pas nouvelle. Quand bien même des mesures de prévention étaient prises pour le minimiser, il persisterait toujours une part d'incertitudes, un « risque résiduel »[751]. Notre propos n'est pas de comparer les risques liés à l'utilisation de l'IA et les risques technologiques classiques, tels que dans les industries nucléaires ou pétrolières. Evidemment, les dommages susceptibles d'être causés sont de nature et d'ampleur différentes. Toutefois, l'utilisation de l'IA peut emprunter le danger de l'activité dans laquelle elle s'insère, voire générer de nouveaux risques sociétaux notamment pour des droits fondamentaux tels que la non-discrimination ou la vie privée. Certains risques sont régulés par les comportements individuels (dois-je traverser la route malgré le risque de me faire écraser ?), d'autres par la société en raison du nombre d'individus concernés par les effets de la réalisation du risque ou parce que le risque est pris indépendamment de la volonté des individus. C'est le cas des risques technologiques et en

[750] G.B. Bruna, « Du risque et de sa perception », *Variances (blog)*, 2 septembre 2020 disponible en ligne : <http://variances.eu/?p=5246>, consulté le 14 septembre 2020.
[751] J. Couturier, G.B. Bruna, F. Tarallo, *et al.*, « Après Fukushima, quelques considérations sur le risque résiduel dans l'industrie nucléaire », in *Risques majeurs, incertitudes et décisions - Approche pluridisciplinaire et multisectorielle*, dir. M. Merad, N. Dechy, L. Dehouck, *et al.*, MA Edition, ESKA, Paris, 2016, 315 p.

particulier de celui porté par l'IA : un petit nombre d'individus (en l'occurrence, les entreprises décidant d'avoir recours à des systèmes d'IA dans leur activité) fait courir un risque à un groupe beaucoup plus large. De nombreux exemples de modèles de régulation des risques technologiques peuvent être trouvés en France, que ce soit à travers l'application de la directive européenne SEVESO[752] sur la prise en compte des accidents industriels majeurs ou la construction du droit du nucléaire[753]. L'exemple de la régulation des risques nucléaires au Royaume-Uni, explicité par la *Health and Safety Executive* dans un document de référence[754], nous semble également particulièrement pertinent.

363. **Un « danger » existant, un « risque » à démontrer.** Toutefois, ce n'est pas parce qu'un danger existe que la probabilité de sa réalisation est importante et ses conséquences graves et irréversibles. La probabilité de la réalisation du dommage ainsi que la gravité des conséquences, composantes de la notion de « risque », diffèrent suivant l'application de l'IA concernée. De plus, le « danger » vient du contexte dans lequel l'IA est utilisée. Les systèmes d'IA ne sont pas dangereux en tant que tels mais leur intégration dans des activités elles-mêmes dangereuses peuvent l'être (conduite d'une centrale nucléaire, maintenance d'un barrage hydroélectrique).

Plan. Ce constat nous conduit à devoir identifier précisément les composantes du « risque » lié à l'utilisation de l'IA (**a**) et notamment, pour l'objet de notre thèse, dans le secteur de l'électricité (**b**).

---

[752] A.H. Maynard, « Environnement : protection contre les pollutions et nuisances », *Dalloz*, 1999, n°5450, 35 et s. ; P. Savin, Y. Martinet, « Risques technologiques et réparation des dommages : points saillants de la loi du 30 juillet 2003 », *Les Petites Affiches*, 10 octobre 2003, n°203, p. 4.

[753] Code de l'Environnement, Livre V, Titre IX, « La sécurité nucléaire et les installations nucléaires de base », articles R592-1 à R596-17 ; S. Emmerechts, « Droit de l'environnement et droit nucléaire : une symbiose croissante », *Bulletin de droit nucléaire*, 2009, vol. 2008/2 ; N. Reboul-Maupin, « La prévention des risques technologiques : aspects juridiques », *Les Petites Affiches*, 16 décembre 2004, n° 251, pp. 6-13.

[754] HEALTH AND SAFETY EXECUTIVE, « The Tolerability of Risk from Nuclear Power Stations », Document officiel, 1988, disponible en ligne : <https://www.onr.org.uk/documents/tolerability.pdf>, consulté le 16 mai 2020 ; V. aussi sur le processus de régulation des risques par la HSE au Royaume-Uni : HEALTH AND SAFETY EXECUTIVE, *Reducing risks, protecting people : HSE's decision making process*, HSE books, 2001, 88 p., disponible en ligne : <https://www.hse.gov.uk/managing/theory/r2p2.pdf>, consulté le 9 septembre 2021.

a) La réalité du risque lié à l'utilisation de l'IA

364. **Un consensus sur l'existence de risques inhérents à l'utilisation de l'IA.** L'existence de dangers liés à l'utilisation de l'IA constitue un postulat de la présente thèse. Les textes institutionnels de référence, qu'ils soient européens[755] ou nationaux[756], ainsi que la doctrine académique[757] ont également déjà présenté dans le détail la façon dont l'utilisation de systèmes d'IA peut conduire à l'atteinte à de nombreux droits et libertés. Par ailleurs, des exemples de situations litigieuses impliquant l'utilisation de techniques d'IA peuvent être trouvés en abondance dans la presse internationale[758] et nationale[759]. Le présent paragraphe synthétise, pour rappel, les grandes catégories de risques connus relatifs à l'utilisation de l'IA.

365. **Le risque lié à la sécurité des individus.** En premier lieu, l'expérience a montré que l'utilisation de systèmes d'IA en remplacement de l'homme pouvait mettre en danger la sécurité des individus, avec l'exemple des accidents causés par des voitures autonomes[760]. Ce premier exemple est assez paradoxal dans la mesure où la promesse initiale de l'application en question

---

[755] COMMISSION EUROPÉENNE, *Livre blanc du 19 février 2020 sur l'Intelligence Artificielle – Une approche européenne axée sur l'excellence et la confiance*, 19 février 2020, COM(2020) 65, p. 10 ; *Résolution 2015/2103(INL) du Parlement européen du 16 février 2017 contenant des recommandations à la Commission concernant des règles de droit civil sur la robotique*, publiée au JOUE n°C252/239 le 18 juillet 2018.

[756] C. Villani, *Donner un sens à l'intelligence artificielle*, Rapport dans le cadre d'une mission parlementaire du 8 septembre 2017 au 8 mars 2018 confiée par le Premier Ministre Edouard Philippe, La Documentation Française, 8 mars 2018 ; CNIL, *Comment permettre à l'homme de garder la main ? Les enjeux éthiques des algorithmes et de l'intelligence artificielle*, Synthèse du débat public animé par la CNIL dans le cadre de la mission de réflexion éthique confiée par la loi pour une République Numérique, 15 décembre 2017.

[757] Voir notamm. C. Castets-Renard, « Comment construire une intelligence artificielle responsable et inclusive ? », *Recueil Dalloz*, 2020, p. 225 ; E. Magrani, « New Perspectives on Ethics and the Laws of Artificial Intelligence », *Journal of Internet Regulation*, 2019, vol. 8, n°3, 19.

[758] M. Zhang, « Google Photos Tags Two African-Americans As Gorillas Through Facial Recognition Software », *Forbes*, 1er juillet 2015, disponible en ligne : <https://www.forbes.com/sites/mzhang/2015/07/01/google-photos-tags-two-african-americans-as-gorillas-through-facial-recognition-software/>, consulté le 10 septembre 2019 ; D. Wakabayashi, « Self-Driving Uber Car Kills Pedestrian in Arizona, Where Robots Roam », *New York Times*, 19 mars 2018, disponible en ligne : <https://www.nytimes.com/2018/03/19/technology/uber-driverless-fatality.html>, consulté le 10 septembre 2019 ; D. Victor, « Microsoft Created a Twitter Bot to Learn From Users. It Quickly Became a Racist Jerk », *New York Times*, 24 mars 2016, disponible en ligne : <https://www.nytimes.com/2016/03/25/technology/microsoft-created-a-twitter-bot-to-learn-from-users-it-quickly-became-a-racist-jerk.html>, consulté le 10 septembre 2019.

[759] M. Viennot, « Intelligence artificielle et éthique ne font pas (encore) bon ménage », *France Culture (blog)*, 25 mai 2019, disponible en ligne : <https://www.franceculture.fr/emissions/la-bulle-economique/intelligence-artificielle-et-ethique-ne-font-pas-encore-bon-menage>, consulté le 22 août 2021 ; S. Gavois, « Une étude pointe les possibles effets pervers et dangers de l'intelligence artificielle », *Nextimpact (blog)*, 26 février 2018, disponible en ligne : <https://www.nextinpact.com/article/28064/106188-une-etude-pointe-possibles-effets-pervers-et-dangers-intelligence-artificielle>, consulté le 22 août 2021.

[760] D. Wakabayashi, « Self-Driving Uber Car Kills Pedestrian in Arizona, Where Robots Roam », *op. cit.*

(la voiture autonome) était de gagner en sécurité et d'éviter des accidents routiers. Or, dans l'exemple de l'accident causé par la voiture Tesla, on reproche au contraire à l'humain de ne pas avoir empêché l'accident[761].

366. **Le risque de reproduction de biais discriminatoires.** Un deuxième danger lié à l'utilisation de l'IA réside dans sa tendance à reproduire et perpétuer des discriminations fondées sur la race[762], le genre[763], ou l'orientation sexuelle[764]. Cette tendance, due aux données sur lesquelles le modèle a appris, peut systématiser la discrimination et en aggraver les conséquences pour les minorités, notamment dans le cadre de technologies utilisées par les forces de l'ordre ou la justice[765].

367. **Le risque pour la vie privée des individus.** Le fonctionnement des systèmes d'IA peut également présenter des dangers pour la vie privée des individus et pour l'autonomie individuelle. En effet, ils nécessitent une grande quantité de données (parfois personnelles) pour fonctionner et ils permettent une hyperpersonnalisation des services pouvant aboutir à une abolition du libre arbitre[766].

368. **Transition.** Ainsi, les systèmes d'IA peuvent effectivement être considérés comme « dangereux » dans la mesure où certains de ses usages peuvent causer des dommages. L'ensemble des dangers présentés ici ne sont que des exemples, non exhaustifs, et ne sont pas

---

[761] V. notamm. NATIONAL TRANSPORTATION SAFETY BOARD (NTSB), « 'Inadequate Safety Culture' Contributed to Uber Automated Test Vehicle Crash - NTSB Calls for Federal Review Process for Automated Vehicle Testing on Public Roads », *Communiqué de presse du NTSB Office of Safety Recommendations and Communications*, 19 novembre 2019, spec. : « *Had the vehicle operator been attentive, the operator would likely have had enough time to detect and react to the crossing pedestrian to avoid the crash or mitigate the impact* ».

[762] E. Ntoutsi, P. Fafalios, U. Gadiraju, *et al.*, « Bias in data-driven artificial intelligence systems : an introductory survey », *WIREs Data Mining Knowlede Discovery*, décembre 2020, 10:1356.

[763] F. Zuiderveen Borgesius, « Strengthening legal protection against discrimination by algorithms and artificial intelligence », *The International Journal of Human Rights*, 2020, 24:10, 1572-1593.

[764] Y. Wang, M. Kosinski, « Deep Neural Networks are more accurate that humans at detecting sexual orientation from facial images », *Journal of personality and social psychology*, février 2018, vol. 114, n°2, 246-257.

[765] J. Angwin, J. Larson, S. Mattu, L. Kirchner, « Machine Bias : There's software used across the country to predict future criminals, and it's biased against blacks », *ProPublica*, 2016, disponible en ligne : <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>, consulté le 17 mars 2020.

[766] G. Koenig, *La fin de l'individu : voyage d'un philosophe au pays de l'intelligence artificielle*, L'Observatoire, coll. De Facto, 189 p.

communs à toutes les applications de l'IA. On les retrouve cependant dans le secteur objet de notre étude : le secteur de l'électricité.

b) La réalité du risque lié à l'utilisation de l'IA dans le secteur de l'électricité

369. **Des potentielles utilisations à haut risque de systèmes d'IA dans le secteur de l'électricité.** Le secteur de l'électricité contient des activités critiques telles que la production d'électricité, d'origine nucléaire ou hydraulique. Il est donc un parfait exemple pour illustrer les risques que pourraient représenter une utilisation de l'IA non maîtrisée. Le présent paragraphe vise à illustrer des cas où des systèmes d'IA conduiraient à des risques importants pour les individus ou la société dans son ensemble.

370. **Une application risquée de l'IA dans les centrales nucléaires.** Un premier exemple est l'utilisation de l'IA dans le système de conduite d'une centrale nucléaire de production d'électricité. Elle est aujourd'hui parfaitement incompatible avec les règlementations actuelles en matière de sécurité et de sûreté nucléaires[767]. Toutefois, en faisant abstraction des contraintes règlementaires, des modèles d'IA pourraient en théorie être utilisés aux fins d'automatisation de certaines fonctions de sécurité (fermeture de valves, réalisation automatique de tâches mécaniques lors d'un rechargement de combustible, opérations de maintenance…)[768]. Une telle utilisation pourrait mettre l'algorithme en capacité d'agir sur des systèmes critiques d'une centrale, et donc le rapprocher du combustible radioactif, dont la libération dans l'environnement pourrait avoir des conséquences graves et irréversibles. Cette hypothèse placerait donc l'IA dans une position où son fonctionnement serait dangereux, dans le sens où il pourrait causer des dommages tels que le dysfonctionnement des mécanismes de sûreté, un relâchement accidentel de matière radioactive, voire un accident nucléaire avec la fusion du cœur. Pour rappel, cette hypothèse est hautement irréaliste, et ce, en raison de la régulation particulièrement contraignante (et justifiée) de l'opération d'une centrale nucléaire, dont la

[767] Voir Supra, 137-180.
[768] Pour un éventail des utilisations possibles des techniques d'apprentissage machine dans le domaine du nucléaire, voir M. Gomez-Fernandez, K. Higley, A. Tokuhiro, *et al.*, « Status of Research and Development of Learning-Based Approaches in Nuclear Science and Engineering: A Review », *Nuclear Engineering and Design*, avril 2020, 359, 110479.

sûreté est supervisée par l'ASN. Les usages les plus dangereux de l'IA dans le nucléaire sont déjà couverts par la règlementation existante. En revanche, d'autres cas d'usage de l'IA moins proches des fonctions essentielles de systèmes critiques, peuvent présenter des dangers qui ne seraient pas couverts par la régulation existante.

371. **Une application risquée dans les barrages hydroélectriques.** L'exemple de la détection de failles sur la digue de barrage hydroélectrique au moyen d'algorithmes de reconnaissance d'image peut illustrer notre propos. Afin d'assurer le bon fonctionnement d'un barrage hydroélectrique, des opérateurs ont la charge d'identifier les failles qui représentent un danger pour la solidité de la digue retenant l'eau. Les conséquences d'une digue qui cède en raison d'une faille mal entretenue peuvent être désastreuses : inondations, dégâts humains et matériels, coûts financiers liés à la perte d'un moyen de production d'électricité, et bien d'autres. Des algorithmes sont utilisés pour aider les opérateurs à identifier, sur des photos, les défauts de la digue nécessitant une opération de maintenance. Aujourd'hui, la décision de déclencher une opération de maintenance reste humaine, l'IA n'intervient que comme une aide à la décision. De plus, le plan de maintenance, notamment sous la supervision des Directions Régionales de l'Environnement, de l'Aménagement et du Logement (DREAL), nécessite toujours une observation humaine de la digue pour détecter d'éventuels défauts. La technologie ne peut venir qu'en supplément du plan de maintenance existant. Toutefois, comme pour tout système d'aide à la décision, il est possible que l'humain suive systématiquement les résultats de la machine, en oubliant ses propres responsabilités. Dès lors, si la confiance dans la machine est trop grande, l'opérateur ne remettra en cause aucune de ses recommandations, ni ne les complètera par sa propre expertise afin de vérifier la pertinence du résultat. Dans le cas qui nous intéresse, cela pourrait conduire les opérateurs à passer à côté de failles dangereuses que l'IA n'aurait pas détectées, ou de prioriser certains défauts identifiés par le système, au détriment d'autres tout aussi importants mais qui n'auront pas été correctement diagnostiqués par l'IA. Encore une fois, dans ces situations, le recours à l'IA peut conduire à des situations dangereuses, en l'occurrence liées à une mauvaise priorisation des opérations de maintenance ou au défaut d'identification de défauts majeurs.

372. **La nécessaire réflexion sur la tolérabilité et la gestion des risques liés aux usages de l'IA.** Une réflexion est nécessaire pour déterminer dans quelle mesure les risques liés aux usages de l'IA, notamment dans le secteur de l'électricité, sont tolérables. Des risques sont dits « tolérables » lorsque la société est prête à vivre avec afin de tirer les bénéfices de l'activité risquée, à la condition que celle-ci soit correctement contrôlée et que ses risques soient

minimisés[769]. La tolérabilité se distingue de l'acceptabilité dans la mesure où cette dernière consiste en l'acceptation d'un risque tel qu'il est, sans volonté de le réduire. Rendre des risques tolérables est donc possible, sous réserve qu'ils soient correctement régulés. C'est pourquoi une méthodologie rigoureuse, et éprouvée par l'expérience, doit être adoptée.

2. La méthodologie de la régulation des risques appliquée à l'IA

373. **Plan.** La méthodologie proposée ici est tirée des principes traditionnels de la régulation des risques technologiques tels qu'appliqués en France[770], au Royaume-Uni[771] ou outre-Atlantique[772]. En partie appliquée par la proposition de règlement sur l'IA par la Commission européenne en 2021[773], il convient de revenir sur les fondements de la régulation des risques avant de pouvoir mettre en lumière ses potentiels écueils. La bonne régulation des risques liés aux usages de l'IA devrait passer par plusieurs étapes : l'identification des risques (**a**), leur évaluation et leur classification en fonction de leur acceptabilité (**b**) et enfin la gestion des risques résiduels pour les rendre tolérables (**c**). Dans un souci d'indépendance de l'expertise scientifique et d'efficacité politique, les différentes étapes de cette méthodologie devraient être réalisées par des autorités publiques distinctes, avec la coopération des acteurs régulés. En effet, l'identification et l'évaluation des risques reposent exclusivement sur des données et savoirs scientifiques, tandis que la gestion des risques (c'est-à-dire le choix concret des mesures à prendre) doit prendre en considération des données socio-économiques.

---

[769] *Ibid.*, 10 ; 24.
[770] P.J. Baralle, « Maîtrise de l'urbanisation autour des installations dangereuses », *JurisClasseur Environnement et Développement durable*, LexisNexis, 15 janvier 2009, Fasc. 4035 ; *Loi n° 2003-699 du 30 juillet 2003 relative à la prévention des risques technologiques et naturels et à la réparation des dommages*, publiée au JORF du7 janv. 2004, p. 644 ; E. Gaillard, « Principe de précaution : Systèmes juridiques internationaux et européens », *JurisClasseur Environnement et Développement durable*, LexisNexis, 1er mars 2015, Fasc. 4035.
[771] HEALTH AND SAFETY EXECUTIVE, *op. cit.*, 25.
[772] C. Hood, H. Rothstein, R. Baldwin, *The government of risk : understanding risk regulation regimes*, Oxford University Press, 2001, 232 p.
[773] COMMISSION EUROPÉENNE, *Proposition de règlement du Parlement européen et du Conseil établissant des règles harmonisées concernant l'intelligence artificielle (législation sur l'intelligence artificielle) et modifiant certains actes législatifs de l'Union*, 21 avril 2021, 2021/0106 (COD), spec. article 7§2.

a) L'identification des risques

374. D'abord, il est nécessaire d'identifier clairement les risques technologiques, qu'ils soient transversaux tels que les risques de biais discriminatoires ou spécifiques à un secteur d'activité, comme le risque lié à l'opacité des systèmes utilisés dans des infrastructures critiques. Pour ces derniers, il convient de laisser les autorités sectorielles mener cette étude, avec la coopération des acteurs régulés, puisqu'elles sont les mieux placées pour identifier ces dangers spécifiques. Toutefois, il convient de retenir un standard commun et partagé afin d'éviter des divergences d'analyse des risques, c'est pourquoi nous avons proposé de nous concentrer sur les risques potentiels pour les droits de l'homme (droit à la vie privée, droit à un environnement sain, droit à la sécurité…). Les méthodes classiques de cartographie des risques ou d'analyse d'impact, déjà répandues dans plusieurs domaines tels que l'application de la loi Sapin II[774] ou du RGPD[775], pourront être utilisées. Pour la réalisation de cette étape, les pouvoirs publics pourront faire appel au débat public, à la consultation des parties prenantes ou au conseil d'experts en IA. On notera que ce travail est déjà bien avancé au niveau international, grâce au travail de certaines institutions[776], d'organisations non gouvernementales[777] ou encore de la communauté universitaire[778].

b) L'évaluation et la classification des risques

375. Une fois identifiés, les risques liés aux usages de l'IA devraient ensuite faire l'objet d'une évaluation, pilotée par les pouvoirs publics. Pour rappel, un risque est caractérisé par trois

---

[774] J. Camy, « Loi sur le devoir de vigilance et loi Sapin II : quelles obligations des entreprises ? », *La Semaine Juridique Entreprise et Affaires*, 18 mars 2021, n° 11, 1135 ; S.J. Saud Neto, N. Tollet, « Cartographie des risques : Retour d'expérience brésilien pour réussir l'exercice », *Cahiers de droit de l'entreprise*, mars 2017, n° 2, dossier 7.
[775] G. Georgiadis, G. Poels, « Towards a privacy impact assessment methodology to support the requirements of the general data protection regulation in a big data analytics context: A systematic literature review », *Computer Law & Security Review*, avril 2022, vol. 44, 105640.
[776] GROUPE D'EXPERTS DE HAUT NIVEAU EN IA, *Lignes directrices du 8 avril 2019 pour une IA digne de confiance*, 8 avril 2019 ; CONSEIL DE L'EUROPE, *Algorithmes et droits humains*, Etude menée par le comité d'experts sur les intermédiaires d'internet MSI-NET, 2017, DGI (2017)12, disponible en ligne : <https ://rm.coe.int/algorithms-and-human-rights-fr/1680795681>, consulté le 8 mars 2020.
[777] INSTITUT MONTAIGNE, *Algorithmes : contrôle des biais S.V.P.*, Rapport, mars 2020 ; AI NOW INSTITUTE, *2019 Annual Report*, rapport, décembre 2019, disponible en ligne : < https://ainowinstitute.org/AI_Now_2019_Report.pdf>, consulté le 12 février 2022.
[778] M.D. Dubber, F. Pasquale, S. Das, *The Oxford handbook of ethics of AI*, Oxford University Press, 2020, 896 p.

composantes[779] : la probabilité de sa réalisation, l'évènement indésirable attaché à la probabilité (exemple : la mise en danger des personnes en raison du dysfonctionnement d'un système critique), et enfin la gravité des conséquences de sa réalisation (Combien de personnes seraient concernées par l'évènement indésirable ? Les conséquences de l'événement seraient-elles irréversibles ?). La caractérisation du risque suivant ces composantes va permettre sa classification dans l'une des trois catégories suivantes :

- Les risques si importants ou dont les conséquences sont si inacceptables qu'ils doivent être refusés en intégralité. Ces risques peuvent couvrir par exemple des situations difficilement contrôlables, qui mettraient en danger un trop grand nombre de personnes ou dont la probabilité de réalisation est très grande et insusceptible d'être minimisée.

- Les risques qui sont, ou ont été rendus, si mineurs qu'aucune précaution supplémentaire n'est nécessaire.

- Les risques tombants entre les deux catégories précédentes, qui devraient être minimisés au niveau le plus bas possible, en prenant en considération les bénéfices potentiels que la société pourrait tirer de la technologie concernée et les coûts de tout effort supplémentaire de prévention des risques. Le principe de prévention impose en effet la minimisation des risques (technologiques[780] ou environnementaux[781]) connus. On retrouve la même idée dans le principe « ALARA »[782] ou « ALARP »[783] dans le droit

[779] O. Sutterlin, « Synthèse – Principe de précaution », *LexisNexis*, Encyclopédies, 2019, spec. 3 : « *L'évaluation des risques a pour objectif d'identifier, de comprendre et de caractériser, en termes qualitatifs et quantitatifs, chacune des trois composantes de la définition du risque (danger, transfert, cible). Il s'agit donc de déceler et de décrire les effets défavorables liés à une activité, une situation ou chose, d'identifier les individus ou les composantes de l'environnement qui sont exposés et d'analyser la façon dont cette exposition peut se produire*. » ; HEALTH AND SAFETY EXECUTIVE, *op. cit.*, 12 ; G.B. Bruna, *op. cit.*

[780] M. Burg, *Droit des risques technologiques*, Legitech, 2020, 1ère ed., 216 p. ; *Directive 2012/18/UE du Parlement européen et du Conseil du 4 juillet 2012 concernant la maîtrise des dangers liés aux accidents majeurs impliquant des substances dangereuses, modifiant puis abrogeant la directive 96/82/CE du Conseil*, publiée au JOUE n°L197 du 24 juillet 2012, dite « SEVESO 3 ».

[781] *Charte de l'environnement*, article 3 : « Article 3. Toute personne doit, dans les conditions définies par la loi, prévenir les atteintes qu'elle est susceptible de porter à l'environnement ou, à défaut, en limiter les conséquences », *Loi constitutionnelle n°2005-205 du 1er mars 2005 relative à la Charte de l'environnement*, publiée au JORF n°0051 du 2 mars 2005, 3697 ; Code de l'environnement, article L110-1, II, 2° ; M. Dreyfus, « Principe de prévention », in *Dictionnaire Collectivités territoriales et Développement Durable*, 2017, pp. 393-395.

[782] « ALARA » pour « As Low As Reasonably Achievable ».

[783] « ALARP » pour « As Low As Reasonably Praticable ».

anglais de la régulation des risques[784]. En France, ce principe est d'interprétation restrictive et n'a été concrétisé qu'en matière de risques environnementaux[785]. Son contenu semble pourtant riche d'enseignements pour la régulation des risques pour les droits humains liés aux usages de l'IA. Preuve en est, la Commission européenne a adopté une méthodologie quasiment similaire dans sa proposition de règlement européen en 2021.

376. **Transition.** Après l'identification et l'évaluation des risques, la troisième étape de la méthodologie de régulation des risques consiste en leur gestion. Autrement dit, elle consiste à répondre à la question : quels mécanismes de contrôle ou de prévention faut-il mettre en place pour rendre les risques résiduels (ceux qui ne sont ni totalement inacceptables, ni totalement acceptables en l'état actuel de la technique) tolérables et ainsi permettre à la société de tirer bénéfice de la technologie ?

c) La gestion des risques

377. **Objectif de la phase de gestion des risques.** Après la phase d'évaluation, la gestion des risques désigne « *l'ensemble des mesures effectivement mises en œuvre pour prévenir la réalisation d'un dommage ou limiter ses conséquences* »[786]. Ces mesures de prévention visent à faire passer le risque résiduel d'une activité dangereuse en dessous d'un seuil fictif de tolérabilité, correspondant à un niveau de risque que la société est prête à accepter pour jouir des bénéfices liés à l'activité concernée (en l'occurrence, de l'usage de l'IA).

378. **Contenu des mesures de prévention des risques.** Les autorités publiques ont un large choix quant aux mesures de gestion des risques. Elles peuvent par exemple fixer des standards contraignants et encadrant la conception de certains produits pour s'assurer qu'aucun acteur ne

[784] H. Pike, F. Khan, P. Amyotte, « Precautionary Principle (PP) versus As Low As Reasonably Practicable (ALARP): Which one to use and when », *Process Safety and Environmental Protection*, 2020, vol. 137, 158-168 ; O. Sutterlin, *op. cit.*, spec. 2 « Prévention et précaution » ; 8 et s.
[785] F.G. Trébulle, « Droit de l'environnement », *Recueil Dalloz*, 2010, n°2468 ; CE 2 septembre 2009, n° 318584 ; *AJDA*, 2009, 1522.
[786] O. Sutterlin, *op. cit.*, 3.

s'éloigne des normes de sécurité[787]. Elles peuvent également mettre en place des mécanismes d'autorisation préalable, conditionnée à l'analyse d'une documentation exhaustive[788]. Dans l'industrie nucléaire, la conception, la construction et la mise en service d'une Centrale Nucléaire de Production d'Electricité (CNPE) sont soumises à la supervision par l'Autorité de Sûreté nucléaire, qui requiert la réalisation d'une démonstration de sûreté des équipements[789]. Cette démonstration, à la charge des « régulés », doit permettre de prouver que la CNPE respecte toutes les normes de sécurité en vigueur et que toutes les mesures techniques, organisationnelles et humaines ont été mises en œuvre pour limiter les risques d'incident[790]. Le choix de ces mesures doit faire l'objet d'une large consultation pour garantir son effectivité et son acceptation dans le secteur régulé. Des développements sur le choix des mesures à adopter pour prévenir les risques liés aux usages de l'IA considérés comme dangereux seront présentés plus loin.

379. **L'identification du responsable de la mise en œuvre des mesures de prévention des risques.** Il est également nécessaire d'identifier le responsable de la limitation des risques, qui aura la charge de mettre en œuvre concrètement les mesures de prévention susmentionnées. Dans le cas de l'IA, les entreprises développant et commercialisant les systèmes logiciels semblent être les mieux placées pour jouer ce rôle. Cette option a d'ailleurs été utilement retenue par la Commission européenne dans sa proposition de règlement sur l'IA en avril 2021.

---

[787] Dans le secteur de l'aviation, plusieurs règlements européens organisent la certification préalable à la mise en service des aéronefs, garantissant notamment leur « navigabilité » sur des critères de sécurité et de sûreté : *Règlement (UE) n°748/2012 de la Commission du 3 aout 2012 établissant des règles d'application pour la certification de navigabilité et environnementale des aéronefs et produits, pièces et équipements associés, ainsi que pour la certification des organismes de conception et de production*, publié au JOUE n°L224/1 du 21 août 2012 ; *Règlement (UE) n°1321/2014 de la Commission du 26 novembre 2014 relatif au maintien de la navigabilité des aéronefs et des produits, pièces et équipements aéronautiques, et relatif à l'agrément des organismes et des personnels participant à ces tâches*, publié au JOUE n°L363/2 du 17 décembre 2014.

[788] Dans le secteur de la santé, la mise sur le marché des médicaments est soumise à la délivrance d'une autorisation préalable par une autorité de surveillance : M. Baumevieille, « L'industrie du médicament – Autorisation de mise sur le marché – Évolution du cadre juridique – Principes généraux », *Litec Droit pharmaceutique*, version mise à jour le 13 juillet 2015, Fasc. 33.

[789] M. Moliner-Dubost, « Nucléaire », *JurisClasseur Administratif*, version mise à jour le 14 avril 2021, Fasc. 378 ; W. Gremaud, « Le droit administratif de la production électronucléaire », *RFDA*, 2021, n°4, 711.

[790] *Ibid.*, spec. 192 et s. ; *Arrêté du 11 janvier 2016 portant homologation de la décision n°2015-DC-0532 de l'Autorité de sûreté nucléaire du 17 novembre 2015 relative au rapport de sûreté des installations nucléaires de base*, publié au JORF n°0012 du 15 janvier 2016, texte n° 7.

380. **La gestion des risques par les autorités de régulation.** Enfin, les autorités de régulation bénéficient généralement de moyens d'audit, de contrôle, de supervision du marché (en imposant par exemple aux acteurs de remonter un certain nombre d'informations précises), et parfois de sanction afin de s'assurer que les risques sont correctement gérés. C'est le rôle de la CNIL en matière de données à caractère personnel, de l'ASN pour la sûreté nucléaire, ou encore l'ANSSI sur la sécurité des systèmes d'information. À leur image, une autorité de régulation dédiée à l'IA aurait donc toute sa place dans la mise en œuvre des exigences de conformité et dans la prévention des risques[791].

381. **Transition.** L'identification et l'évaluation des risques liés aux usages de l'IA devraient permettre de mettre en évidence les systèmes d'IA pour lesquels les risques ne sont pas acceptables. Pour ces derniers, des mesures préventives doivent être mises en œuvre afin de prévenir le risque. La suite de notre propos contient des propositions d'exigences de conformité (puisque nous avons retenu le modèle de la co-régulation fondé sur la conformité) poursuivant cet objectif.

B/ Proposition d'exigences de conformité pour les systèmes d'IA risqués

382. **La nécessaire normalisation de la conception et de l'utilisation des systèmes d'IA.** La doctrine scientifique et les professionnels de l'IA se rejoignent sur la possibilité de mettre en place des mesures techniques et organisationnelles afin de prévenir les risques générés par certains systèmes d'IA[792]. Les principes éthiques à appliquer aux systèmes d'IA divergent selon

[791] Voir Infra, Section 2.

[792] Voir par exemple l'ensemble des travaux universitaires sur « l'IA responsable » (V. Dignum, *Responsible Artificial Intelligence : how to develop and use AI in a responsible way*, Springer, coll. Artificial Intelligence, 2019, 1st ed., 127p. ; J. Fjeld, H. Hilligoss, N. Achten, *et al.*, « Principled Artificial Intelligence : Mapping Consensus in Ethical and Rights-based Approaches to Principles for AI», *Berkman Klein Center for Internet & Society (blog)*, 2020, disponible en ligne : <https://dash.harvard.edu/bitstream/handle/1/42160420/HLS%20White%20Paper%20Final_v3.pdf?sequence=1&isAllowed=y>, consulté le 28 février 2022) ou des cadres de normalisation volontaire publiés par des entreprises (Rolls Royce, *The Aletheia Framework*, 14 décembre 2020, disponible en ligne : <https://www.rolls-royce.com/~/media/Files/R/Rolls-Royce/documents/stand-alone-pages/aletheia-framework-booklet-2021.pdf>, consulté le 28 février 2022 ; IBM, *IBM's Principles for trust and transparency,* 30 mai 2018, disponible en

leur source[793], tout comme les moyens pour les concrétiser dans la pratique. Si certaines entreprises tentent de prendre les devants en adoptant, volontairement, un cadre organisationnel pour le développement de systèmes d'IA, d'autres patientent, dans l'attente de normes uniformes. Les principaux risques générés par les systèmes d'IA, bien qu'ils diffèrent suivant l'application et le secteur concernés, commencent aujourd'hui à faire consensus : protection de la vie privée, remplacement de l'humain dans la prise de décision, biais algorithmiques… Afin de limiter la fragmentation du marché et de guider au mieux les entreprises vers un développement responsable de systèmes d'IA, les mesures techniques et organisationnelles pour y parvenir gagneraient à être uniformisées. Le cadre juridique de l'IA proposé dans notre thèse repose sur un double principe de proportionnalité et de subsidiarité. Ainsi, la règlementation de l'IA doit fixer les grands principes guidant le développement de l'IA, tandis que leur déclinaison sectorielle se fera par les autorités de régulation compétentes et les acteurs régulés et ce, afin que les normes ainsi créées collent au mieux aux contraintes de la pratique.

383. **Plan.** Dans la suite de cette Section, seront présentées dix exigences de conformité qui, sans aller jusqu'aux mesures techniques à mettre en place, visent à identifier précisément les problématiques auxquelles les entreprises devront prêter attention lorsqu'elles développeront, vendront ou utiliseront des systèmes d'IA. Ces propositions de « normes », au sens où elles constituent des règles *« fixant les conditions de la réalisation d'une opération, de l'exécution d'un objet ou de l'élaboration d'un produit dont on veut unifier l'emploi »* [794], peuvent concerner la phase de conception du système (« amont ») (**1**) ou son maintien en conditions opérationnelles et son utilisation (« aval ») (**2**). En effet, si la plupart des initiatives de régulation se concentrent sur l'encadrement de la phase de conception, la pratique en entreprise révèle que de nombreux risques sont mieux adressés par des mesures de maintenance ou de supervision après la mise en service du système. Enfin, ces normes doivent être construites comme des grands principes. Les entreprises, accompagnées par leurs autorités de régulation sectorielles, auront la charge de déterminer, en pratique, les mesures précises à mettre en œuvre pour les

ligne : <https://www.ibm.com/blogs/policy/wp-content/uploads/2018/06/IBM_Principles_SHORT.V4.3.pdf>, consulté le 20 octobre 2019).

[793] A. Jobin, M. Ienca, E. Vayena, « The Global Landscape of AI Ethics Guidelines », *Nature Machine Intelligence*, septembre 2019, vol. 1, n°9, 389-399

[794] *Dictionnaire Larousse*, 2022 ; V. aussi G. Cornu, *Vocabulaire juridique*, PUF, 10ème ed., 2014, p. 689.

respecter. L'environnement numérique et l'état de la technique, surtout dans le domaine de l'IA, étant particulièrement évolutifs, ces principes doivent prendre la forme d'obligations de moyens, contraignant les entreprises à prendre toutes les mesures techniques et organisationnelles afin de les atteindre. Elles devront être en mesure, en cas de contrôle par l'autorité de régulation compétente[795], de produire les preuves des mesures mises en place *via* de la documentation technique, des audits des systèmes concernés ou encore des rapports de certification par des tiers indépendants. Des standards sectoriels et mécanismes de certification pourront être créés pour faciliter la mise en conformité.

1. Une normalisation « en amont » de la mise en service

384. **Le processus de conception d'un système d'IA.** Le processus de conception d'un système d'IA est composé de nombreuses étapes, lesquelles dépendent de la finalité recherchée et des techniques utilisées. Il commence généralement par l'identification du problème à résoudre : modélisation d'un phénomène physique, traitement automatique de documents, aide à la prise de décision, automatisation, analyse de données… Cette étape initiale conditionne le reste du processus. Une fois que la tâche à exécuter par le système est identifiée, les porteurs du projet vont identifier les solutions techniques pour y parvenir. À cette fin, des développeurs et des experts « métiers »[796] vont collaborer pour déterminer les contraintes et choisir les techniques pertinentes pour réaliser le système. Les briques technologiques (vision par ordinateur[797], détection de texte[798] ou de voix[799], jumeau numérique[800], analyse automatique de données[801]…) et les techniques précises à mobiliser (réseau de neurones profonds[802], ontologies[803], traitement automatique du langage[804], approches logiques ou symboliques…)

[795] Voir Infra, Section 2.
[796] Des experts de l'activité concernée et de la tâche à réaliser par la machine.
[797] Dans la littérature et la pratique : *Computer Vision.*
[798] Dans la littérature et la pratique : « OCR » pour *Optical Character Recognition.*
[799] Dans la littérature et la pratique : *Speech Recognition*.
[800] Dans la littérature et la pratique : *Digital Twin.*
[801] Dans la littérature et la pratique : *Automated data analysis*.
[802] Dans la littérature et la pratique : *Deep neural networks*.
[803] Dans la littérature et la pratique : *Ontology*.
[804] Dans la littérature et la pratique : « TALN » en français, « NLP » pour *Natural Language Processing* en anglais.

sont choisies à ce stade. En parallèle, si l'application le justifie, devront être collectées les données pertinentes au regard de la finalité poursuivie (documents à analyser, mesures physiques à simuler…). Ces données doivent être collectées, sélectionnées, nettoyées et labellisées, souvent par des experts « métiers » avant de pouvoir être utilisées pour construire le système d'IA. De plus, dans la pratique, une partie de ces données prétraitées sont mises de côté aux fins de validation en fin de processus. Les parties prenantes peuvent ensuite concevoir l'algorithme qui traitera les données, souvent directement à partir des données préparées si des techniques d'apprentissage automatique sont utilisées. L'algorithme sera ensuite testé dans un environnement dédié sur des données fictives, réelles ou issues de la préparation des données. Des modifications, ou un « réentrainement », peuvent être nécessaires à ce stade en fonction des résultats obtenus. Si la performance de l'algorithme est satisfaisante[805] alors il pourra être validé et passer « en production », en conditions réelles. D'autres choix peuvent être effectués en fonction du système concerné, tels que le choix du matériel à utiliser ou de l'architecture informatique. Enfin, lorsque le système est « en production », il est supervisé et mis à jour par les personnes en charge de son « maintien en conditions opérationnelles », qui ne sont généralement pas les mêmes qui ont conçu le système.

385. **La nécessaire uniformisation des bonnes pratiques de conception des systèmes d'IA.** Prévenir les risques de défaut des systèmes d'IA ou les risques que son utilisation peut engendrer lors de son utilisation implique d'uniformiser les bonnes pratiques dans l'intégralité de ce processus. En effet, la diffusion de bonnes pratiques permettrait de limiter, par exemple, le risque de biais discriminatoire, les failles de sécurité informatique ou l'utilisation abusive de données personnelles. Puisque le système n'évolue qu'à la marge après sa mise en service, au gré des mises à jour, les erreurs encapsulées dans l'algorithme lors de la conception ont vocation à se répéter au cours de son cycle de vie. C'est donc un moment crucial de la vie de l'algorithme, qu'il nous faut encadrer pour anticiper les risques.

386. **Plan.** Certaines bonnes pratiques nécessitent d'être partagées par toutes les entreprises développant des systèmes d'IA, en particulier lorsque ces derniers sont utilisés de telle manière

[805] Généralement, la performance de l'algorithme est évaluée selon des « indices de confiance » et en fonction d'un pourcentage de succès. Le « seuil de confiance » diffère suivant les secteurs et les entreprises (85%, 90%, 95%, 100%).

qu'ils font peser un risque sur les individus ou leurs biens. Tout en respectant l'approche par les risques conditionnant la contrainte de la norme à la sévérité du risque, les développements qui suivent contiennent une série de sept grands principes qui, au vu de la pratique, gagneraient à être partagés plus largement. À réaliser en amont de la mise en service du système, ils consistent d'abord en la réalisation d'une analyse d'impact préalable (**a**) pour certains systèmes d'IA susceptibles de générer un risque élevé pour les droits et libertés des individus. Ils relèvent ensuite de la mise en œuvre de mesure visant à garantir la qualité des données utilisées (**b**) et la sécurité informatique, conformément à l'état de l'art en la matière (**c**). Ils recommandent enfin la création d'exigences de conformité uniformisées relatives à l'explicabilité (**d**), aux processus de vérification et de validation (**e**) et à la sobriété numérique (**f**).

a) La réalisation d'une analyse d'impact préalable

387. **Les moyens pour définir le niveau de dangerosité d'un système.** Afin de définir le niveau de risque d'un système d'IA, plusieurs moyens sont envisageables : soit l'on préqualifie certaines applications comme dangereuses, par exemple *via* une liste établie par les autorités normatives que ce soit au niveau horizontal ou sectoriel, soit l'on charge les entreprises de définir elles-mêmes le niveau de risque de chacun de leurs systèmes. Le premier moyen est celui choisi par les autorités européennes dans la proposition de règlement européen sur l'IA d'avril 2021[806], comprenant quatre catégories de risques : risque minimum, faible risque, haut risque et risque inacceptable[807]. Il présente l'avantage de la sécurité juridique à la condition que les définitions, et le cas-échéant les listes, soient suffisamment précises. Toutefois, ce modèle ne permet que peu de souplesse. Toutes modifications des définitions, ou ajouts, devront passer

[806] COMMISSION EUROPÉENNE, *Proposition de règlement du Parlement européen et du Conseil établissant des règles harmonisées concernant l'intelligence artificielle (législation sur l'intelligence artificielle) et modifiant certains actes législatifs de l'Union*, 21 avril 2021, 2021/0106 (COD)..

[807] Voir la « pyramide des risques de l'IA » dans A. Bensamoun, « Artificial Intelligence Act : l'Union européenne invente la pyramide des risques de l'intelligence artificielle », *Le Club des juristes (blog)*, 21 mai 2021, disponible en ligne : <https://blog.leclubdesjuristes.com/artificial-intelligence-act-lunion-europeenne-invente-la-pyramide-des-risques-de-lintelligence-artificielle/>, consulté le 28 février 2022 ; Y. Meneceur, « Proposition de règlement de l'IA de la Commission européenne : entre le trop et le trop peu ? », *Les Temps électriques (blog)*, 22 avril 2021, disponible en ligne : <https://lestempselectriques.net/index.php/2021/04/22/proposition-de-reglement-de-lia-de-la-commission-europeenne-entre-le-trop-et-le-trop-peu/#more-1762>, consulté le 28 février 2022.

par une modification des textes, ce qui peut prendre beaucoup de temps alors que le champ de l'IA évolue très rapidement. Le second moyen, l'autodéfinition des niveaux de risques, présente l'avantage d'éviter cette lourdeur administrative mais pourrait générer des abus de la part d'entreprises peu scrupuleuses. La combinaison des deux approches nous paraît être la solution optimale : une prescription des autorités normatives sur les systèmes d'IA devant être considérés en toutes circonstances comme inacceptables ou à haut risque, d'une part, et une obligation pour les entreprises à réaliser une étude d'impact des autres systèmes susceptibles de générer des risques pour les droits et libertés des individus, d'autre part. Si cette dernière venait révéler l'existence de dangers relatifs, par exemple, à la sécurité physique des personnes ou à la discrimination de minorités, alors le système devrait être traité comme « à haut risque ». Construite sur l'exemple de l'analyse d'impact relative à la protection des données (AIPD) prescrite dans le RGPD pour les traitements « présentant un risque élevé pour la vie privée des personnes »[808], cette proposition apparaît proportionnée dans la mesure où les systèmes vraiment peu risqués n'auraient pas à se soumettre à cette analyse préalable.

388. **L'analyse d'impact préalable appliquée aux systèmes d'IA.** Le modèle de l'AIPD du RGPD, conditionné à l'existence d'un « risque élevé », peut ici être repris. Sa complétion par des listes non-exhaustives des traitements pour lesquels l'analyse est requise ou non requise[809], ainsi que la publication des critères à prendre en compte pour définir si un traitement présente un risque élevé[810] peuvent être reproduit dans le cadre de la régulation de l'IA. Concernant son contenu, il peut également s'inspirer de l'AIPD[811] : description des systèmes d'IA et de leurs finalités, évaluation des risques pour les droits et libertés des individus susceptibles d'être affectés par son fonctionnement, mesures envisagées les limiter et garantir le respect des autres

[808] RGPD, article 35.
[809] CNIL, *Délibération n°2019-118 du 12 septembre 2019 portant adoption de la liste des types d'opérations de traitement pour lesquelles une analyse d'impact relative à la protection des données n'est pas requise*, publiée au JORF n°0246 du 22 octobre 2019, texte n°90 ; CNIL, *Délibération n°2018-327 du 11 octobre 2018 portant adoption de la liste des types d'opérations de traitement pour lesquelles une analyse d'impact relative à la protection des données est requise*, publiée au JORF n°0256 du 6 novembre 2018, texte n°82.
[810] GROUPE DE TRAVAIL « ARTICLE 29 » SUR LA PROTECTION DES DONNÉES, *Lignes directrices concernant l'analyse d'impact relative à la protection des données (AIPD) et la manière de déterminer si le traitement est « susceptible d'engendrer un risque élevé » aux fins du règlement (UE) 2016/679*, adoptées le 4 avril 2017, WP 248 rév. 01 ; CNIL, *Délibération n°2018-326 du 11 octobre 2018 portant adoption de lignes directrices sur les analyses d'impact relatives à la protection des données (AIPD) prévues par le règlement général sur la protection des données (RGPD)*, publiée au JORF n°0256 du 6 novembre 2018, texte n° 81.
[811] RGPD, article 35, 7.

exigences de conformité applicables. La différence entre l'AIPD et l'analyse d'impact préalable pour les systèmes d'IA proposée ici concerne le champ des risques à analyser. Dans le cas de l'IA, seront également recherchés tous les risques relatifs à la sécurité physique des personnes, à un défaut de fonctionnement ou à l'impact du système sur son environnement. De nombreux travaux de recherche universitaire, en particulier à l'échelle internationale, promeuvent l'idée de la création d'une « analyse d'impact algorithmique »[812], si bien que le gouvernement canadien l'a déjà adopté en 2019[813] et que les États-Unis en prennent la voie[814].

389. **Proposition d'obligation de réalisation d'une analyse d'impact algorithmique :**

Les entreprises seraient ainsi tenues de réaliser une analyse d'impact algorithmique dans le cas où le système d'IA qu'elles projettent de mettre en service serait susceptible de générer un risque élevé pour les droits et libertés des individus.

Trois listes accompagneraient cette obligation, suivant le même schéma que l'AIPD. La première serait une liste non exhaustive des systèmes d'IA présentant, d'emblée, des risques élevés, pour lesquels l'analyse d'impact est obligatoire. La deuxième serait une liste non exhaustive contenant des systèmes pour lesquels une analyse d'impact n'est pas requise. Enfin, la troisième contiendrait les critères à prendre en considération pour définir si un système est susceptible de générer un risque élevé et est donc soumis à analyse d'impact. Ces critères pourraient avoir trait, par exemple, au degré d'autonomie du système, à sa faculté d'interaction directe avec des individus, ou aux conséquences probables de son dysfonctionnement.

390. **Conséquences de l'identification de risques graves.** Un système ayant fait l'objet d'une analyse d'impact algorithmique confirmant la gravité des risques qu'il fait peser sur les droits et individus devra se plier aux autres exigences de conformité afin de les prévenir. À ce titre,

[812] K. Yeung, A. Howes, G. Pogrebna, « AI Governance by Human Rights-Centred Design, Deliberation and Oversight: An End to Ethics Washing », in *The Oxford Handbook of AI Ethics*, dir. M. Dubber, F. Pasquale, Oxford University Press, 2019 ; M.E. Kaminski, G. Malgieri, « Algorithmic Impact Assessments under the GDPR: Producing Multi-Layered Explanations », *SSRN Electronic Journal*, 2019.

[813] GOVERNMENT OF CANADA, *Directive on Automated Decision-Making*, 1er avril 2019, disponible en ligne : <https://www.tbs-sct.gc.ca/pol/doc-eng.aspx?id=32592>, consulté le 16 octobre 2019 ; GOVERNMENT OF CANADA, « Algorithmic Impact Assessment », *Site gouvernemental (blog)*, disponible en ligne : <https://open.canada.ca/aia-eia-js/?lang=en.>, consulté le 16 octobre 2019.

[814] *Algorithmic Accountability Act of 2019*, proposition de loi par les sénateurs C. Booker et R. Wyden devant le Congrès américain, 4 octobre 2019, H.R.2231, 116th U.S. Congress (2019-2020).

les mesures techniques et organisationnelles prises pour garantir la qualité des données utilisées, la sécurité informatique, la traçabilité, la vérification et la sobriété numérique devront figurer dans l'analyse d'impact.

b) La garantie de la complétude, de la représentativité et de la qualité des données

391. **L'importance de la qualité des données dans la prévention des risques.** La qualité et la pertinence des données utilisées pour concevoir un système d'IA sont les principaux critères de son bon fonctionnement. Par ailleurs, la plupart des dysfonctionnements de systèmes d'IA, dont les plus retentissants médiatiquement, sont dus à des données incomplètes, non représentatives, ou contenant des biais discriminatoires[815]. Prévenir les erreurs dans les données, c'est prévenir d'éventuels dommages causés par le fonctionnement du système d'IA, qui reproduira les erreurs ou aboutira à de mauvais résultats. Pour y parvenir, les entreprises concevant des systèmes d'IA susceptibles de générer un risque élevé pourraient donc se voir imposer une obligation de moyens, les contraignant à mettre en œuvre des mesures techniques et organisationnelles destinées à garantir la complétude, la représentativité et la qualité des données utilisées pour la conception et le fonctionnement du système concerné, conformément à l'état de la technique.

392. **Les moyens pour prévenir les risques liés aux données utilisées lors de la conception et du fonctionnement du système d'IA.** Bien que les normes précises doivent être érigées par des autorités de régulation sectorielles ou par un consensus d'entreprises afin qu'elles correspondent au mieux à la réalité pratique, il convient de préciser les moyens envisageables pour limiter les risques liés aux données. Tout d'abord, les données utilisées pour la conception du système doivent être complètes. Des études de disponibilité des données peuvent être menées

[815] P. Bertail, D. Bounie, S. Clémençon, P. Waelbroeck, « Algorithmes : biais, discrimination et équité », *Telecom Paris Tech*, février 2019, disponible en ligne : <https://www.telecom-paris.fr/wp-content-EvDsK19/uploads/2019/02/Algorithmes-Biais-discrimination-equite.pdf>, consulté le 16 mars 2020 ; NATIONAL TRANSPORTATION SAFETY BOARD (NTSB), « 'Inadequate Safety Culture' Contributed to Uber Automated Test Vehicle Crash - NTSB Calls for Federal Review Process for Automated Vehicle Testing on Public Roads », *Communiqué de presse du NTSB Office of Safety Recommendations and Communications*, 19 novembre 2019 ; E. Ntoutsi, P. Fafalios, U. Gadiraju, *et al.*, « Bias in data-driven artificial intelligence systems : an introductory survey », *WIREs Data Mining Knowlede Discovery*, décembre 2020, 10:1356.

par les développeurs afin d'identifier les données à leur disposition et de vérifier si elles sont suffisantes au vu du système à concevoir. Si des données essentielles sont manquantes ou partielles, la conception du modèle pourrait être fortement compromise et le projet abandonné. Ensuite, les données doivent être représentatives. L'appréciation de la représentativité dépend de la finalité envisagée pour le système d'IA. S'il vise à simuler un phénomène physique comme l'évolution de la pression ou de la radioactivité dans un environnement fermé, alors il faudra s'assurer de disposer de suffisamment de mesures de capteurs pour le représenter numériquement. Le défaut de prise en compte d'un paramètre pourrait conduire à des résultats erronés et à des dommages si la modélisation est utilisée aux fins de prise de décision. Si le système d'IA a vocation à traiter des données personnelles, il convient de s'assurer de la représentation des diverses populations et minorités dans les données utilisées pour prévenir la reproduction de biais discriminatoires. La complétude et la représentativité ne doivent pas être entendues comme une incitation à utiliser la plus grande quantité de données. Outre le non-respect du principe de minimisation dans le cas de données à caractère personnel, l'utilisation d'un trop grand nombre de données pourrait conduire à un « surentraînement » du modèle d'IA et desservir sa performance[816]. Enfin, les données doivent être de qualité et, autant que possible, exemptes d'erreurs. Une méthodologie rigoureuse doit être suivie lors de la phase de sélection, préparation, labellisation et vérification des données. Les processus et méthodes suivies lors de cette phase devraient être documentés et justifiés dans l'analyse d'impact, pour prouver que toutes les dispositions ont été prises pour détecter d'éventuelles erreurs dans les données d'entrée.

[816] I. Bilbao, J. Bilbao, « Overfitting problem and the over-training in the era of data: Particularly for Artificial Neural Networks », *2017 Eighth International Conference on Intelligent Computing and Information Systems (ICICIS)*, 2017, pp. 173-177 ; M.L. Astion, M.H. Wener, R.G. Thomas, *et al.*, « Overtraining in neural networks that interpret clinical data », *Clinical Chemistry*, Volume 39, Issue 9, 1 September 1993, pp. 1998-2004 ; J. Sjöberg, L. Ljung, « Overtraining, regularization and searching for a minimum, with application to neural networks », *International Journal of Control*, 1995, vol. 62, n°6, pp. 1391-1407.

393. **Proposition d'obligation relative à la qualité des données utilisées dans la conception de l'IA :**

Les entreprises développant un système d'IA susceptible de générer un risque élevé seraient ainsi tenues de mettre en place toutes les mesures techniques et organisationnelles pertinentes au regard de l'état de la technique afin de garantir la complétude, la représentativité et la qualité des données utilisées dans la conception dudit système.

Ces mesures devraient être documentées et jointes à l'analyse d'impact algorithmique préalable.

394. Outre les données, les choix des techniques employées et de l'architecture informatique peuvent également contribuer aux risques générés par les systèmes d'IA, d'où l'importance d'y prêter également une attention particulière.

c) La garantie de la conformité à l'état de l'art en matière de sécurité

395. **Les risques liés à la sécurité de l'information et inhérents aux spécificités des systèmes d'IA.** La sécurité de l'information peut être définie comme la *« protection de la confidentialité, de l'intégrité et de la disponibilité de l'information »*[817] et fait l'objet de plusieurs normes ISO[818] ou autres référentiels publiés par l'ANSSI[819]. La confidentialité est la propriété selon laquelle l'information n'est pas diffusée ni divulguée à des personnes, des entités ou des processus non autorisés ; l'intégrité, la propriété d'exactitude et de complétude ; et la disponibilité, la propriété d'accessibilité et d'utilisabilité à la demande par une entité autorisée. Les risques liés à la sécurité informatique correspondent à des *« événements liés à la sécurité de l'information, indésirables ou inattendus, présentant une probabilité forte de compromettre*

[817] ISO, *Norme ISO/IEC 27000:2018*, Technologies de l'information – Techniques de sécurité – Systèmes de management de la sécurité de l'information – Vue d'ensemble et vocabulaire, février 2018, 5ème ed., 29 p. ; ANSSI, *Glossaire*, Site institutionnel, disponible en ligne : < https://www.ssi.gouv.fr/particulier/glossaire/i/>, consulté le 1er mars 2022.
[818] V. notamm. : ISO, *Norme ISO/IEC 27000:2018*, *op. cit.* ; ISO, *Norme ISO/IEC 27001*, Management de la sécurité de l'information, octobre 2013, 2ème ed., 23 p.
[819] ANSSI, *Référentiel général de sécurité v2.0*, Référentiel d'exigences techniques, 13 juin 2014, 25 p. ; ANSSI, *Référentiel « PRIS v2 » - Prestataires de réponse aux incidents de sécurité*, Référentiel d'exigences techniques, 2 août 2017, 53 p.

les opérations liées à l'activité de l'organisme et de menacer la sécurité de l'information »[820]. Ces risques d'incidents liés à la sécurité de l'information sont inhérents à tous systèmes logiciels ou informatiques. Ils peuvent, dans certains cas, constituer des délits punis par le Code pénal depuis la Loi Godfrain[821] : accès non autorisé à un système automatisé de données, entrave à son fonctionnement, introduction, modification ou extraction frauduleuse de données. La probabilité de survenance de ces risques est accrue lorsque le système logiciel concerné est un système d'IA. En effet, plusieurs de leurs spécificités multiplient les vulnérabilités. Premièrement, leur dépendance à la donnée augmente le risque d'empoisonnement des données dès la conception[822] ou de manipulation lors du fonctionnement. Deuxièmement, leur complexité structurelle multiplie les voies d'entrée pour d'éventuelles attaques informatiques. Leur architecture peut être composée de serveurs principaux, de nombreux capteurs collectant et transmettant des données en temps réel, lesquelles peuvent être traitées par des services d'informatique en nuage, potentiellement de différents éditeurs, avant de déclencher une action sur un autre système matériel… La démultiplication des maillons de la chaîne crée autant de nouveaux risques. Troisièmement, la potentielle autonomie des systèmes d'IA accroît la gravité des potentielles conséquences d'une attaque informatique. Plus un procédé critique est automatisé, plus un défaut dans le système le pilotant aura d'importants impacts sur son environnement. Enfin, quatrièmement, l'opacité du fonctionnement des systèmes d'IA pourrait faire naître de nouveaux risques liés à la sécurité de l'information si les développeurs ne sont pas en mesure d'anticiper complètement le fonctionnement dudit système. Ainsi, les spécificités de l'IA accroissent la gravité des risques liés à la sécurité informatique ou en multiplient la probabilité de survenance. C'est la raison pour laquelle il est essentiel que des bonnes pratiques, uniformisées, soient suivies afin de limiter ce phénomène.

396. **Assurer la sécurité des systèmes d'IA par la publication de standards techniques.** Il est essentiel que les systèmes d'IA susceptibles de générer un risque élevé pour les droits et

[820] ISO, *Norme ISO/IEC 27000:2018*, *op. cit.*, item « Incident lié à la sécurité de l'information ».
[821] Code pénal, articles 323-1 et suivants ; *Loi n°88-19 du 5 janvier 1988 relative à la fraude informatique*, publiée au JORF du 6 janvier 1988.
[822] Expression en référence au terme anglais « *Data poisoning attacks* » utilisé dans la pratique : A. Schwarzschild, M. Goldblum, A. Gupta, *et al.*, « Just How Toxic is Data Poisoning? A Unified Benchmark for Backdoor and Data Poisoning Attacks », *Proceedings of the 38th International Conference on Machine Learning*, PMLR, 2021, vol. 139, 9389-9398.

libertés des individus présentent le plus haut niveau de sécurité informatique. Pour ce faire, la meilleure solution semble être l'adoption de standards que pourraient suivre les entreprises. En matière de sécurité informatique, la solution est déjà expérimentée et est plébiscitée par les entreprises qui, malgré son coût de mise en œuvre, présente l'avantage de la sécurité juridique[823]. En plus de présenter les bonnes pratiques et références techniques à utiliser pour correspondre à l'état de l'art en matière de sécurité informatique, l'existence de standards permet la mise en place de processus de certification. La certification peut être réalisée de façon autonome par l'entreprise qui, en documentant son processus de certification, sera à même de prouver qu'elle a mis en place l'ensemble des mesures envisageables pour que le système d'IA développé présente le niveau le plus élevé de sécurité. Le processus de certification peut également être externalisé et réalisé par un tiers indépendant. Le cas-échéant, la conformité aux standards sera présumée. Pour une problématique aussi technique que celle de la sécurité informatique d'un système logiciel, et encore plus d'IA, la combinaison entre référentiels techniques et processus de certification nous semble la plus prometteuse[824]. Les travaux engagés par l'ENISA et par l'AFNOR en ce sens sont prometteurs[825] et pourraient constituer des références majeures sur l'état de l'art en matière de sécurité informatique des systèmes d'IA.

[823] Nous pouvons ici citer, par exemple, les travaux de normalisation de l'agence européenne chargée de la sécurité des réseaux et de l'information (ENISA), de l'AFNOR ou de l'ANSSI.

[824] Sur l'utilité de la certification dans une logique de régulation, voir : J.M. Pontier, « La certification, outil de modernité normative », *Recueil Dalloz*, 1996, n°41, 355-360 ; E. Guiraud, « Le rôle de l'éthique dans la mise en place d'une certification pour l'utilisation d'algorithmes dans le système juridique », *Ethique publique*, 2019, vol. 21, n°1.
Sur la pertinence de la certification en matière de sécurité informatique dans des secteurs éminemment techniques, voir : V. Louis, C. Baron, « Vers une certification continue des logiciels critiques en aéronautique », *Techniques de l'ingénieur*, coll. Technologies logicielles Architectures des systèmes, novembre 2019, n°h8060 ; N. Leveson, *Safeware : System Safety and Computers*, Addison-Wesley Professional, 1995, 704 p.

[825] ENISA, *Securing Machine Learning Algorithms*, rapport, 14 décembre 2021, 70 p., disponible en ligne : <https://www.enisa.europa.eu/publications/securing-machine-learning-algorithms>, consulté le 12 février 2022 ; AFNOR, *Normes en cours de conception pour le terme « intelligence artificielle »*, disponible en ligne : <https://norminfo.afnor.org/search?term=intelligence+artificielle>, consulté le 12 février 2022.

397. **Proposition d'obligation relative à la sécurité de l'IA dès la conception :**

Les entreprises développant un système d'IA susceptible de générer un risque élevé seraient ainsi tenues de mettre en place toutes les mesures techniques et organisationnelles pertinentes au regard de l'état de la technique et des standards en vigueur afin de garantir, dès la conception, le plus haut niveau de sécurité.

Ces mesures devraient être documentées et jointes à l'analyse d'impact algorithmique préalable. La conformité aux standards européens ou nationaux en matière de sécurité des systèmes d'IA pourra être prouvée *via* la documentation des étapes suivies si la certification s'effectue en interne, ou *via* l'obtention d'une certification par un tiers indépendant et agréé par les autorités nationales compétentes.

398. **Transition.** Toutefois, même en suivant les standards les plus exigeants, il est impossible de supprimer complètement le risque d'incident de sécurité. C'est pourquoi il faut que les entreprises mettent en place, dès la conception, des mécanismes afin d'assurer la traçabilité du fonctionnement du système d'IA, la reproductibilité d'éventuelles erreurs afin d'en identifier l'origine et l'explicabilité des résultats.

d) La garantie de la traçabilité, de la reproductibilité et de l'explicabilité

399. **Les enjeux de la traçabilité des résultats d'un système d'IA.** Plusieurs raisons justifient la nécessité de pouvoir tracer, dans une certaine mesure, les résultats produits par un système d'IA. Selon le système concerné, ces résultats peuvent être des diagnostics, des recommandations ou encore le déclenchement d'actions pouvant influencer son environnement. Dans ce dernier cas, si l'action entreprise par le système de façon autonome cause un dommage à un individu ou un bien, il est important de pouvoir identifier la cause du dommage : dysfonctionnement du système, utilisation non conforme aux instructions d'usage, manipulation du système ou simple défaut dans la collecte de données de capteurs… Ce besoin est justifié, en partie, pour garantir la bonne application des régimes de responsabilité civile[826].

[826] Voir Supra, 100.

Dans les cas où le système d'IA ne produirait qu'un diagnostic d'une situation ou des recommandations, la décision resterait humaine donc l'attribution de la responsabilité devrait être plus aisée en cas de préjudice. Toutefois, il pourrait s'avérer utile de pouvoir rejouer le fonctionnement du système, ou du moins pouvoir tracer le résultat, afin de pouvoir déterminer si l'humain ayant pris une décision sur ce fondement n'aurait pas dû s'en détacher en raison d'une erreur manifeste. Somme toute, cela reviendrait à venir contrôler, *a posteriori*, si l'humain a exercé un contrôle effectif sur le traitement automatisé de données qui lui sert de fondement pour prendre des décisions[827]. Dans l'ensemble de ces situations, et pour assurer la bonne application des règles de droit en vigueur, il serait utile de pouvoir tracer le fonctionnement d'un système d'IA susceptible de causer des dommages.

400. **Traçabilité, reproductibilité et explicabilité dès la conception.** La traçabilité doit permettre aux entreprises de pouvoir fournir, en cas de contentieux, de contrôle par une autorité normative ou de demande justifiée des usagers, des informations sur le fonctionnement et les résultats produits par un système d'IA susceptible de générer un risque élevé. Les concepteurs devraient à ce titre le concevoir de telle manière que des données relatives au fonctionnement et aux résultats soient générées, stockées et accessibles pendant une durée proportionnée au regard du système concerné. La nature et la durée de conservation de ces données doivent être appréciées au cas par cas et ne doivent pas excéder ce qui est nécessaire aux fins de reproductibilité en cas de dommage causé par le système d'IA. La reproductibilité consiste à pouvoir « rejouer un scénario » *a posteriori* afin de pouvoir l'analyser. Si les résultats d'un système ne peuvent générer des actions que dans la semaine qui suit, alors il n'est pas utile de conserver les données de traçabilité plus longtemps que cela, sauf si les dommages résultant des actions peuvent elles-mêmes mettre du temps à se manifester. Enfin, pour que ces données de traçabilité permettent d'identifier l'origine d'un défaut par exemple, elles doivent permettre à des experts d'expliquer les résultats. De nombreuses méthodes visant à expliquer le fonctionnement, parfois opaque, des systèmes d'IA sont aujourd'hui développées[828]. Elles

[827] Lorsque le système constitue un traitement de données à caractère personnel, la situation est déjà encadrée par le RGPD à travers l'application de son article 22 sur les « décisions prises exclusivement sur le fondement d'un algorithme », voir Supra, 111 et s.

[828] D. Gunning, M. Stefik, J. Choi, *et al.*, « XAI : Explainable artificial intelligence », *Science Robotics*, 18 décembre 2019, vol. 4, n°37 ; F.K. Došilović, M. Brčić, N. Hlupić, « Explainable artificial intelligence: A survey »,

doivent être sélectionnées en fonction du destinataire de l'explication et des techniques d'IA utilisées[829].

401. **Proposition d'obligation relative à la traçabilité du fonctionnement de l'IA.**

Les entreprises développant un système d'IA susceptible de générer un risque élevé seraient ainsi tenues de mettre en place toutes les mesures techniques et organisationnelles pertinentes au regard de l'état de la technique afin de garantir, dès la conception, la traçabilité du fonctionnement et des résultats du système.

Les données générées et leur durée de stockage doivent être proportionnées à la nature du système et peuvent être supprimées dès que le besoin de traçabilité n'est plus justifié. Ces données doivent permettre la reproductibilité d'un résultat litigieux ou ayant causé un préjudice et l'explication de la façon dont le système a produit ce résultat (notamment en analysant les données d'entrée et la logique sous-jacente de l'algorithme). Ces obligations permettraient de garantir la bonne application des règles de droit existantes, en particulier des régimes de responsabilité civile. Les mesures adoptées devraient figurer dans l'analyse d'impact du système d'IA présentant un risque élevé, afin que les autorités normatives puissent juger de leur conformité à l'état de l'art. Ces mesures visent pour la majorité à faciliter l'attribution de la responsabilité *a posteriori* de la réalisation du dommage, *via* la prise en compte dès la conception de la traçabilité des résultats du système.

402. **Transition.** Néanmoins, afin de minimiser le risque de survenance des dommages, les entreprises devraient mettre en place des processus de vérification et de validation des systèmes d'IA présentant un risque élevé, avant leur mise en service.

41st International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO), 2018, pp. 0210-0215 ; W. Samek, K.R. Müller, « Towards Explainable Artificial Intelligence », in *Explainable AI: Interpreting, Explaining and Visualizing Deep Learning. Lecture Notes in Computer Science*, dir. W. Samek, G. Montavon, A. Vedaldi, *et al.*, Springer, 2019, vol. 11700.

[829] ACPR, *Gouvernance des algorithmes d'intelligence artificielle dans le secteur financier*, Document de réflexion, spec. p. 12 et s., disponible en ligne : <https://acpr.banque-france.fr/sites/default/files/medias/documents/20200612_gouvernance_evaluation_ia.pdf>, consulté le 24 mai 2021 ; V. Beaudouin, I. Bloch, D. Bounie, *et al.*, « Flexible and Context-Specific AI Explainability: A Multidisciplinary Approach », *SSRN Electronic Journal*, 2020, disponible en ligne : < https://www.ssrn.com/abstract=3559477>, consulté le 29 avril 2020.

e) Le suivi d'une procédure de vérification et de validation

403. **L'expérience des logiciels critiques pour la sûreté.** Les systèmes d'IA ne sont pas les premiers systèmes logiciels faisant peser un risque important pour la sécurité des biens et des personnes. Depuis longtemps, des logiciels sont utilisés dans l'aéronautique, le nucléaire ou encore l'automobile, et ce, y compris pour des fonctions critiques telles que le pilotage automatique ou les systèmes de sécurité. Or, ces logiciels, constitués d'algorithmes déterministes, sont aujourd'hui parfaitement maîtrisés et leurs utilisations régulées. Bien que les techniques employées diffèrent, leur nature ne diffère pas tant que cela avec les systèmes d'IA actuels, si bien que nous pouvons nous inspirer des méthodes de validation des premiers pour encadrer efficacement les seconds.

404. **Pour l'application de la méthode « VVQI » aux systèmes d'IA avant leur mise en service.** L'étude des règlementations sectorielles de l'énergie applicables aux systèmes d'IA réalisée dans la Première Partie a mis en évidence l'applicabilité de la méthode dite « VVQI » pour « vérification, validation et quantification des incertitudes »[830]. Cette méthode est composée de trois étapes et a pour objectif de garantir la pertinence des choix effectués et la performance du système. La vérification vise à constater l'absence d'anomalies dans le code informatique et sa performance pour accomplir la tâche voulue. Lors de la validation, les évaluateurs vont questionner la pertinence des choix réalisés pour atteindre le résultat attendu et, le cas-échéant, comparer le fonctionnement du modèle avec des données de validation « du monde réel ». Enfin, la dernière étape consiste en la quantification des incertitudes. Une étude de sensibilité peut être réalisée pour identifier les données d'entrée ayant une influence sur le résultat si nécessaire[831]. Seules les entreprises évoluant dans un environnement critique pour la sûreté (aéronautique, nucléaire, automobile…) sont familières avec ces méthodes de qualification des logiciels, qui sont pourtant aujourd'hui bien maîtrisées. De nombreux travaux sont en cours pour les adapter pleinement aux différentes techniques utilisées pour construire

[830] Voir Supra, 190-200.
[831] Pour une présentation plus détaillée de la méthode « VVQI » et sur les adaptations à lui apporter pour être pleinement applicables aux systèmes d'IA, voir Supra, 211-241.

des systèmes d'IA, y compris les plus opaques. Cette expérience doit être mobilisée pour construire le cadre des systèmes d'IA susceptibles de générer un risque élevé pour les droits et libertés des individus.

405. **Proposition d'obligation relative à la certification des systèmes les plus risqués :**

Les entreprises développant un système d'IA susceptible de générer un risque élevé seraient ainsi tenues de mettre en place une procédure de vérification et de validation du système en amont de sa mise en service.

Cette procédure pourra faire l'objet de standards, lesquels pourront utilement s'inspirer de ceux existants dans les procédures de qualification des logiciels critiques pour la sûreté dans les industries aéronautique ou nucléaire par exemple.

406. **Transition.** La dernière proposition d'exigence de conformité à imposer aux systèmes d'IA est sans doute la plus novatrice, bien qu'elle soit certainement la plus difficile à mettre en œuvre dans un horizon de temps court. Qu'importe leur niveau de risque, l'ensemble des systèmes d'IA devraient être conçus de telle manière qu'ils minimisent leur empreinte sur l'environnement.

f) La mise en œuvre de mesures pour garantir la sobriété numérique

407. **L'impératif de la sobriété numérique.** Du rapport Villani[832] aux lignes directrices du groupe d'experts de haut niveau en IA établi par la Commission européenne[833], en passant par nos institutions[834] et le milieu scientifique[835], il existe aujourd'hui un consensus global sur la nécessité de réduire l'empreinte environnementale du numérique et en particulier de l'IA, caractérisée par sa voracité en énergie. Les systèmes d'IA dont le fonctionnement nécessite des quantités déraisonnables d'énergie ou dont l'infrastructure a généré de grandes quantités d'émissions de gaz à effet de serre, devraient être considérés comme présentant un risque élevé pour les individus. Le seuil de l'acceptabilité en matière d'émissions directes ou indirectes de CO2 liées à un système d'IA doit faire l'objet d'un consensus le plus large possible. Quoi qu'il en soit, sans égard au risque et à l'application concernée, la conception de n'importe quel système d'IA devrait être fondée sur une démarche de sobriété numérique.

408. **La concrétisation de la sobriété numérique dans la conception d'un système d'IA : un besoin de normalisation.** Les entreprises ont tendance à reconnaître l'importance de la sobriété numérique et lancent elles-mêmes des programmes de « green IT » ou de numérique responsable[836]. Toutefois, elles se heurtent à un obstacle bloquant : le défaut de norme pour

---

[832] C. Villani, *Donner un sens à l'intelligence artificielle*, Rapport dans le cadre d'une mission parlementaire du 8 septembre 2017 au 8 mars 2018 confiée par le Premier Ministre Edouard Philippe, *La Documentation Française*, 8 mars 2018, p. 20.

[833] GROUPE D'EXPERTS DE HAUT NIVEAU EN IA, *Lignes directrices du 8 avril 2019 pour une IA digne de confiance*, 8 avril 2019, p. 16.

[834] CONSEIL NATIONAL DU NUMÉRIQUE, *Feuille de route sur l'environnement et le numérique - 50 mesures pour un agenda national et européen sur un numérique responsable c'est-à-dire sobre et au service de la transition écologique et solidaire et des objectifs de développement durable*, Rapport remis à la ministre de la Transition écologique et solidaire et au secrétaire d'État chargé du Numérique, juillet 2020, disponible en ligne : <https://cnnumerique.fr/environnement_numerique>, consulté le 01/04/2021 ; M. Vivant, « Publication des travaux du CNNum sur l'environnement et le numérique », *Revue Lamy Droit de l'Immatériel*, 1er août 2020, n°173.

[835] P. Dhar, « The Carbon Impact of Artificial Intelligence », *Nature Machine Intelligence*, 2020, vol. 2, n°8, 423 ; E. Strubel, « Energy and Policy Considerations for Deep Learning in NLP », *57th Annual meeting of the Association for Computational Linguistics*, 5 juin 2019 ; ou M. Whittaker, R. Dobbe, « AI and Climate Change: How they're connected, and what we can do about it », *Medium (blog)*, 17 octobre 2019, disponible en ligne : <https://medium.com/@AINowInstitute/ai-and-climate-change-how-theyre-connected-and-what-we-can-do-about-it-6aa8d0f5b32c>, consulté le 7 avril 2021.

[836] H. D'Agrain, « Du Green IT au Green by IT : exemples d'applications dans les grandes entreprises », *CIGREF*, janvier 2017, disponible en ligne : <https://www.cigref.fr/wp/wp-content/uploads/2017/01/CIGREF-Du-Green-IT-au-Green-by-IT-2017.pdf>, consulté le 11 mars 2020 ; Collectif Numérique Responsable, site officiel, disponible en ligne : <https://collectif.greenit.fr/>, consulté le 7 mai 2021.

mesurer l'empreinte environnementale d'un système d'IA. Des travaux de normalisation doivent être menés pour créer un référentiel commun et proposer des exemples de méthodes ou outils pouvant être utilisés par les entreprises pour réaliser un diagnostic de leur empreinte environnementale numérique. Nous proposons, à ce sujet, des pistes de réflexion, inspirées de l'expérience des bilans de gaz à effet de serre, dans le Titre 2 de cette Partie[837]. Outre la quantification, les entreprises devraient en tout état de cause adopter des mesures pour minimiser l'empreinte environnementale liée à la conception et au fonctionnement du système d'IA. Cela implique de réaliser des choix quant aux techniques et langages utilisés, puisqu'il est prouvé que chacun ne présente pas la même efficacité énergétique[838]. La quantité de données utilisées, à collecter ou à stocker, ainsi que le lieu de stockage auront également des incidences sur l'empreinte environnementale et devront donc être justifiés[839].

409. **Proposition d'obligation relative à la sobriété écologique de l'IA.**

Les entreprises développant un système d'IA, quel que soit son niveau de risque, devraient être en mesure de démontrer qu'elles ont mis en œuvre, dès la conception, des mesures techniques et organisationnelles afin de quantifier et minimiser l'empreinte environnementale du système concerné.

Ces mesures devront être documentées et pourront concerner les méthodes de développement, les techniques employées, les matériels utilisés, la gestion des données ou tout autre aspect pouvant contribuer directement ou indirectement à la dégradation de l'environnement.

[837] Voir Infra, 648-652.

[838] R. Pereira, M. Couto, F. Ribeiro, *et al.*, « Energy efficiency across programming languages: how do energy, time, and memory relate ? », *Proceedings of the 10th ACM SIGPLAN International Conference on Software Language Engineering (SLE 2017)*, Association for Computing Machinery, New York, NY, USA, 256–267 ; G. Pinto, F. Castor, « Energy efficiency: A new concern for application software developers », *Communications of the ACM*, Novembre 2017.

[839] V. Moreau, *Méthodologie de représentation des impacts environnementaux locaux et planétaires, directs et indirects - Application aux technologies de l'information*, thèse pour le doctorat en sciences et génie de l'environnement, Ecole Nationale Supérieure des Mines de Saint-Etienne, 2012, 361 p.

410. **Conclusion du 1. sur la nécessaire normalisation de la phase de conception des systèmes d'IA.** Afin de lutter contre les risques générés par l'utilisation des systèmes d'IA, des bonnes pratiques peuvent être mises en place en amont de la mise en service. Dans une logique de proportionnalité, ces bonnes pratiques ne devraient être contraignantes que pour les systèmes d'IA susceptibles de présenter un risque élevé pour les droits et libertés des individus. À cette fin, nous proposons la réalisation d'une étude d'impact algorithmique préalable à la mise en service des systèmes d'IA considérés comme présentant un risque élevé par les autorités normatives ou que les entreprises jugent susceptibles de générer un risque élevé. Dans le cas où le système présenterait effectivement des risques élevés, des exigences de conformité devront être respectées, sous la forme d'une obligation de moyens, par les entreprises le concevant. Les mesures techniques et organisationnelles mises en œuvre pour se mettre en conformité devraient être documentées et jointes à l'étude d'impact algorithmique. Le recours à des standards et à des certifications par des tiers indépendants pourraient être facilités afin d'assouplir la contrainte administrative imposée aux entreprises. Les exigences de conformité lors de la phase de conception visent à garantir la qualité des données utilisées, la sécurité informatique, la traçabilité des résultats, la validation technique du système, ainsi que la minimisation de son empreinte environnementale. Toutefois, la phase de conception n'est pas le seul moment de la vie d'un système d'IA lors duquel les risques sont créés. C'est pourquoi, ces exigences doivent être accompagnées d'obligations « en aval » de la mise en service.

2. Une normalisation « en aval » de la mise en service

411. **La nécessaire uniformisation des bonnes pratiques de maintien en conditions opérationnelles des systèmes d'IA.** La finalité et les modalités de fonctionnement d'un système d'IA sont en grande partie déterminées et fixées lors de la phase de conception. Il est donc important de l'encadrer, afin d'éviter d'encapsuler des erreurs ou défauts dans l'algorithme, susceptibles de générer des dommages une fois mis en service. Toutefois, il ne serait pas raisonnable de ne compter que sur la mise en place de bonnes pratiques lors de la conception pour prévenir les risques que peut générer un système d'IA lors de son fonctionnement. En effet, leur complexité, l'avancée rapide des techniques et de la connaissance dans ce domaine, ainsi que la potentielle évolution des systèmes au cours de leur utilisation justifient leur surveillance et leur suivi tout au long de leur cycle de vie. C'est la raison pour laquelle nous recommandons également la mise en place obligatoire de bonnes

pratiques encadrant le maintien en conditions opérationnelles des systèmes d'IA les plus risqués.

412. **Plan.** Ces bonnes pratiques visent à garantir le maintien du plus haut niveau de qualité, de sécurité et de performance tout au long du cycle de vie du système d'IA (**a**), par des mesures techniques et organisationnelles « en aval » de la mise en service. Puisqu'en pratique les risques apparaissent majoritairement lorsque les systèmes agissent de façon autonome ou que leurs recommandations, potentiellement défectueuses, sont suivies par l'humain sans contrôle effectif, les entreprises développant les systèmes d'IA les plus risqués devraient garantir leur supervision humaine (**b**). Enfin, pour pallier l'opacité de ces technologies et permettre son contrôle effectif, tant par les autorités normatives à travers des audits que par la société à travers l'exercice des droits individuels, les systèmes d'IA risqués devraient n'être utilisés que dans des conditions de transparence vis-à-vis des parties prenantes et d'information complète des utilisateurs (**c**).

a) Le maintien de la qualité tout au long du cycle de vie

413. **Le nécessaire maintien de la qualité au cours du cycle de vie.** La qualité d'un système d'IA peut décroître au fil du temps. Les raisons sont nombreuses et peuvent avoir trait, par exemple, à la découverte de techniques plus précises, à l'utilisation de nouveaux matériels informatiques plus performants, à la découverte de nouvelles failles informatique jusque là inconnues, à un auto-apprentissage sur des données de faible qualité, à la reproduction d'erreurs commises lors de la conception ou encore à la manipulation frauduleuse du système. Lorsque ce dernier est susceptible de générer un risque élevé pour les droits et libertés des individus, cette dégradation de la qualité dans le temps n'est pas acceptable. Des mesures devraient être prises pour s'assurer que le haut niveau de sécurité et de performance atteint à l'issue de la phase de conception soit maintenu tout au long du cycle de vie. Un système d'IA dont la sécurité et la précision n'est ni surveillée ni maintenue au cours du temps ne devrait pas être utilisé dans des situations susceptibles de générer un risque pour les droits et libertés des individus.

414. **Les moyens du maintien de la qualité au cours du cycle de vie.** De nombreuses mesures peuvent être mises en œuvre dans le maintien en conditions opérationnelles d'un système d'IA risqué. D'abord, un plan de suivi peut être créé afin de surveiller qu'aucune anomalie n'apparaît dans le fonctionnement du système, les données ou les résultats produits. Ce suivi est conditionné par la mise en place, dès la conception, de mesures techniques de

traçabilité du système[840]. Ensuite, tout changement significatif dans l'utilisation du système, son environnement ou les données d'entrée devrait conduire à l'actualisation de l'analyse d'impact algorithmique pour vérifier que les mesures mises en place sont toujours pertinentes et suffisantes. Enfin, un plan de mises à jour, visant à maintenir dans le temps la conformité du système d'IA risqué avec l'état de l'art en matière de sécurité informatique et de performance devrait être construit et rigoureusement appliqué.

415. **Proposition d'obligation relative au maintien de la qualité tout au long du cycle de vie de l'IA :**

Les personnes physiques ou morales opérant un système d'IA susceptible de générer un risque élevé pour les droits et libertés des individus devraient être tenues de mettre en œuvre un plan de maintien de la qualité, de la sécurité et de la performance du système tout au long de son cycle de vie.

Ce plan devrait couvrir le suivi de l'utilisation aux fins de détection d'éventuelles anomalies, l'actualisation de l'analyse d'impact algorithmique en cas d'évolution significative du système ou de son environnement, ainsi que la réalisation des mises à jour nécessaires au maintien dans la durée de la sécurité et de la performance du système au niveau de l'état de l'art. Tout système d'IA non maintenu dans ces conditions ne devrait pas être utilisé dans des conditions susceptibles de générer des risques pour les individus.

416. **Transition.** Ledit plan de maintien de la qualité devrait inclure les mesures prises pour assurer la supervision humaine du système.

b) La supervision humaine

417. **Les justifications de la supervision humaine du fonctionnement des systèmes d'IA présentant un risque élevé.** L'exemple des systèmes logiciels utilisés dans des infrastructures critiques telles que les avions, les centrales de production d'électricité ou de gaz, par exemple, peut être utilisé pour illustrer la nécessité d'une supervision humaine des systèmes d'IA risqués.

[840] Voir Supra, 399-401.

Dans ces environnements, ces systèmes logiciels classiques ne peuvent opérer des fonctions critiques de façon automatique que parce qu'ils ont été dûment qualifiés en amont et surtout parce qu'il est établi que leur comportement est strictement déterministe[841]. Dans le cas de l'IA, le fonctionnement précis peut être opaque, partiellement autonome et, dans certains cas marginaux, évoluer par un apprentissage au cours du cycle de vie. Ces caractéristiques rendent difficile la qualification *a priori*[842] qui est pourtant une condition à la confiance que l'on peut accorder au système pour fonctionner correctement après sa mise en service. Ces incertitudes structurelles, c'est-à-dire contre lesquelles on ne peut lutter, puisqu'elles sont inhérentes aux spécificités des systèmes d'IA, constituent une première justification de la nécessaire supervision humaine des systèmes d'IA présentant un risque élevé pour les droits et libertés des individus. Une seconde justification tient en la bonne application de la règle de droit. Pensée pour la subjectivité humaine[843], la simple présence d'une personne physique au contrôle ou à la garde d'un objet permet de limiter grandement les problématiques juridiques liées, par exemple, à la causalité ou à l'attribution des responsabilités[844].

418. **Les moyens de la supervision humaine.** La supervision humaine sur le fonctionnement d'un système d'IA présentant des risques élevés doit être effective. À ce titre, sa concrétisation peut s'inspirer du « contrôle effectif » à exercer sur les algorithmes dans le cadre de l'application du RGPD, en particulier dans le cadre du profilage et des décisions automatisées[845]. Elle doit garantir qu'un ou plusieurs opérateurs humains surveillent le fonctionnement du système et vérifient l'absence d'anomalies, qu'ils puissent à tout moment reprendre la main, et que tout dysfonctionnement technique conduise, aussi rapidement que possible, à une intervention humaine. Si, en cas de dysfonctionnement, le risque pour les

[841] Pour une illustration dans le secteur aéronautique, voir : L. Wang, « Issues on software testing for safety-critical real-time automation systems », *The 23rd Digital Avionics Systems Conference*, IEEE, 2004, Cat. n°04CH37576, pp. 530-101.
[842] Voir Supra, 190-200.
[843] S. Merabet, *Vers un droit de l'intelligence artificielle*, thèse de doctorat en droit privé, Dalloz, 2020, pt 360.
[844] Voir par exemple : M. Lamoureux, « La causalité juridique à l'épreuve des algorithmes », *JCPG*, 2016, p.731.
[845] Voir Supra, 113 ; GROUPE DE TRAVAIL « ARTICLE 29 » SUR LA PROTECTION DES DONNÉES, *Guidelines on Automated individual decision-making and Profiling for the purpose of Regulation 2016/679*, 3 octobre 2017, n°WP251 ; N. Martial-Braz, « Le profilage », *Communication Commerce Electronique*, avril 2018, p. 70 et s., dossier spécial « Entrée en vigueur du Règlement général sur la protection des données » ; CNIL, *Délibération n°2017-299 du 30 novembre 2017 portant avis sur un projet de loi d'adaptation au droit de l'Union européenne de la loi n°78-17 du 6 janvier 1978*, p. 28.

individus est élevé, alors il doit être conçu d'une telle manière qu'un autre système prenne le relais ou qu'une intervention humaine soit sollicitée avant la réalisation de l'action dommageable. Les opérateurs humains en charge de la supervision doivent disposer des moyens et des connaissances requises pour analyser les données de fonctionnement, détecter une éventuelle anomalie et intervenir sur le système pour empêcher la réalisation d'une action dommageable, le cas-échéant. Le terme « supervision » nous semble plus adapté que « contrôle », utilisé notamment dans la proposition de règlement européen sur l'IA[846], car ce dernier exclue de fait tous les systèmes autonomes. Or, certains systèmes d'IA, comme certains logiciels traditionnels, peuvent fonctionner de façon autonome sans présenter de risques s'ils ont été correctement qualifiés et que leur comportement n'est pas amené à évoluer ou fluctuer. Si la seule « supervision » du système permet de minimiser de manière effective la survenance du risque, alors il serait disproportionné d'imposer aux entreprises de mettre en place des mesures de « contrôle » permanent sur le système. La supervision proposée implique la surveillance, suivi et capacité d'intervention. Tandis que le contrôle, impliquant l'opérateur humain dans chacun des actions du système, ne serait requis que si ces dernières ne suffisent pas à maîtriser le risque.

419. **Proposition d'obligation relative à la supervision humaine de l'IA :**

Les personnes physiques ou morales opérant un système d'IA susceptible de générer un risque élevé pour les droits et libertés des individus devraient être tenues de garantir sa supervision humaine tout au long du cycle de vie.

La supervision humaine a pour objet de minimiser le risque de survenance d'éventuels dommages causés par le système. Elle doit être effective et non artificielle : les mesures prises devraient garantir qu'un ou plusieurs opérateurs humains surveillent le fonctionnement du système et vérifient l'absence d'anomalies, qu'ils puissent à tout moment reprendre la main, et que tout dysfonctionnement technique conduise, aussi rapidement que possible, à une intervention humaine.

[846] COMMISSION EUROPÉENNE, *Proposition de règlement du Parlement européen et du Conseil établissant des règles harmonisées concernant l'intelligence artificielle (législation sur l'intelligence artificielle) et modifiant certains actes législatifs de l'Union*, 21 avril 2021, 2021/0106 (COD), article 14.

420. **Transition.** L'organisation de la supervision humaine doit être transparente, de manière à permettre aux autorités normatives ou à la société de contrôler la façon dont sont utilisés les systèmes d'IA les plus risqués. Pour que ce contrôle soit effectif, il convient de contraindre les entreprises à la transparence et à la diffusion de l'information nécessaire à l'exercice des droits individuels.

c) La transparence et l'information des utilisateurs

421. **L'utilité de la diffusion d'informations sur la conception et le fonctionnement des systèmes d'IA risqués.** Des mécanismes de transparence, donc relatifs à la disponibilité, l'accessibilité ou la fiabilité de l'information, peuvent être utilisés comme instruments de régulation[847]. La transparence peut être utilisée afin de créer de la confiance entre les acteurs d'un marché ou envers des institutions[848], ou encore pour assurer la responsabilité des parties prenantes[849] . Elle apparaît également pertinente dans le domaine des technologies émergentes et de l'IA puisque l'on peut constater, dans l'actualité, les effets négatifs que peuvent avoir des scandales publics sur des failles de sécurité ou des algorithmes discriminants. Whatsapp a par exemple vu ses utilisateurs fuir par millions sur des applications concurrentes lorsque l'entreprise a publié ses nouvelles conditions générales d'utilisation, permettant une plus grande exploitation des données des utilisateurs[850]. Dans un autre domaine, en 2019, la publication d'un rapport établissant un lien de causalité entre le logiciel de pilotage automatique de Tesla dans un accident mortel a généré une chute de 8% du cours boursier de l'entreprise[851]. Ces exemples montrent que la seule divulgation de certaines informations relatives à l'utilisation

[847] A.W. Buijze, *The principle of transparency in EU law*, thèse pour le doctorat en droit, Utrecht University, 338 p.
[848] Pour un exemple dans le secteur financier : C. Kaufmann, R.H. Weber, « The Role of Transparency in Financial Regulation », *Journal of International Economic Law*, septembre 2010, vol. 13, n°3, pp. 779-797).
[849] Pour un exemple en matière de protection des données à caractère personnel : D. Spagnuelo, A. Ferreira, G. Lenzini, « Accomplishing Transparency within the General Data Protection Regulation », *Proceedings of the 5th International Conference on Information Systems Security and Privacy* (ICISSP 2019), pp. 114-125.
[850] A. Hern, « WhatsApp loses millions of users after terms update », *The Guardian*, 24 janvier 2021, disponible en ligne : <https://www.theguardian.com/technology/2021/jan/24/whatsapp-loses-millions-of-users-after-terms-update>, consulté le 25 novembre 2021.
[851] A. Levy, L. Kolodny, « Tesla shares drop after report says its Autopilot system was engaged during a fatal crash », *CNBC (blog)*, 17 mai 2019, disponible en ligne : <https://www.cnbc.com/2019/05/17/tesla-shares-fall-on-report-autopilot-system-was-engaged-during-crash.html>, consulté le 25 novembre 2021.

d'algorithmes ou à l'exploitation des données des individus peut avoir de lourdes conséquences pour les entreprises. Imposer la transparence de certaines informations pourrait donc inciter les entreprises à adopter un comportement conforme aux attentes de la société, par crainte de se voir sanctionner économiquement par la perte de clients ou de valeur financière.

422. **La minimisation naturelle des indicateurs publiés au titre de la transparence.** De plus, la transparence présente le double intérêt d'inciter les entreprises à agir pour minimiser les indicateurs, données ou événements à divulguer mais aussi de permettre aux pouvoirs publics de disposer d'un plus grand nombre d'informations dans l'optique, le cas-échéant, de futures mesures plus précises. L'intégration d'indicateurs environnementaux tels que les émissions de gaz à effet de serre dans les rapports financiers pour les sociétés cotées[852] en est un bon exemple. En effet, cette obligation de transparence impose aux entreprises de mesurer des indicateurs, fixés par des standards[853], qu'elles n'avaient pas forcément l'habitude de quantifier par le passé. Leur publication permet à la fois au grand public et aux investisseurs de se faire une idée du sérieux de l'entreprise dans son engagement contre le réchauffement climatique mais aussi aux pouvoirs publics de pouvoir identifier les secteurs les plus émetteurs et, le cas-échéant, déterminer des actions pour lutter plus efficacement contre les émissions de gaz à effet de serre. Ce constat est confirmé dans d'autres domaines, tels que la protection des données à caractère personnel avec l'obligation de notifier les violations de données personnelles présentant un risque pour la vie privée des personnes concernées[854] ou la sûreté nucléaire avec l'information continue du public sur les incidents de sûreté, supervisée par l'Autorité de sûreté nucléaire[855].

[852] Code de l'environnement, article L229-25.
[853] Dans l'exemple donné, il s'agit des normes comptables IFRS/IAS : E.M. Barbu, N. Feleaga, L. Feleaga, « Quelles normes IAS/IFRS utiliser pour le reporting environnemental ? », *Revue Française de Comptabilité*, février 2011, n°440.
[854] RGPD, articles 33 et 34.
[855] Code de l'environnement, article L591-5 sur l'obligation de l'exploitant nucléaire de notifier à ASN tout incident ou accident affectant les installations qu'il exploite ; Code de l'environnement, article L592-1 et s. sur les missions de l'ASN en matière d'information du public dans le domaine de la sûreté nucléaire ; Code de l'environnement, article L592-32 sur l'information du public en cas d'urgence radiologique.

423. **Proposition d'obligations relatives à la transparence et à l'information du public.** Face à tous ces constats, il nous semble qu'imposer un certain degré de transparence aux acteurs du cycle de vie des systèmes d'IA permettrait à la société d'en contrôler le développement :

Les entreprises développant un système d'IA susceptible de générer un risque élevé seraient tenues de publier les analyses de risques effectuées, l'explication de la logique sous-jacente aux systèmes susceptibles de générer des risques pour les droits fondamentaux[856], les moyens techniques et tests réalisés pour lutter contre les biais discriminatoires, et les indicateurs relatifs à l'empreinte environnementale des systèmes déployés[857].

L'information du public sur ces éléments devrait être obligatoire pour éviter l'apparition de comportements de passagers clandestins mais surtout pour lutter contre l'incitation implicite des entreprises, due au fonctionnement du marché, à garder certaines informations secrètes par peur de perdre des parts de marché[858].

424. **Les objectifs poursuivis par la transparence.** Si la portée de l'obligation de transparence devra faire l'objet d'une consultation avec les acteurs concernés pour converger sur des indicateurs réalistes, les pouvoirs publics doivent garder à l'esprit que la transparence en tant qu'instrument de régulation doit poursuivre quatre objectifs[859] : mettre en place des mécanismes pour mettre en lumière des risques (pour leurs droits à la vie privée, à la sécurité, à un traitement non-discriminatoire, à un environnement sain,…), garantir la fiabilité de l'information à divulguer (par la mise en place de standards coconstruits et acceptés par les destinataires de la norme), publier et partager l'information avec toutes les parties prenantes

[856] GROUPE DE TRAVAIL « ARTICLE 29 » SUR LA PROTECTION DES DONNÉES, *Guidelines on Automated individual decision-making and Profiling for the purpose of Regulation 2016/679*, 3 octobre 2017, n°WP251 ; P.M. Schwartz, E.J. Janger, « Notification of Data Security Breaches », *Michigan Law Review*, 2006, n°105, 913-984.

[857] Sur des propositions précises sur le contenu de la transparence environnementale des systèmes d'IA, voir Infra, 662.

[858] S. Schott, C. Wing, « Information Disclosure as an Environmental Policy Instrument and a Self-Regulatory Tool », in *Innovation, Science, Environment: Canadian Policies and Performance*, dir. G.B. Doern, McGill-Queens University Press, 2006, pp. 213-232.

[859] B. Jarvis, « Accounting for the Unaccountable: Valuing the Environment in Energy Policy », in *Canadian Energy Policy and the Struggle for Sustainable Development*, dir. G.B. Doern, University of Toronto Press, 2005, pp. 105-127.

(personnes concernées par l'utilisation des systèmes d'IA, investisseurs, pouvoirs publics, chercheurs) et enfin agir sur l'information divulguée (s'il s'agit d'indicateurs, il convient d'inciter à leur minimisation ; s'il s'agit d'événements de sécurité, il convient de s'assurer que le problème soit identifié et résorbé dans les plus brefs délais,…).

425. **Conclusion du 2. sur la nécessaire normalisation du maintien en conditions opérationnelles des systèmes d'IA.** Le seul encadrement de la phase de conception d'un système d'IA ne peut suffire à la prévention de l'ensemble des risques que son utilisation peut générer. C'est pourquoi des mesures doivent être prises pour encadrer, également, son maintien en conditions opérationnelles tout au long du cycle de vie. À l'image des bonnes pratiques devant encadrer la phase de conception des systèmes d'IA susceptibles de générer un risque élevé pour les droits et libertés des individus, de nombreuses mesures sont envisageables. Trois séries de propositions, obligations incombant aux personnes physiques ou morales opérant le système, ont été faites. Premièrement, un plan de maintien de la qualité, de la sécurité et de la performance du système tout au long du cycle de vie du système d'IA devrait être mis en œuvre, visant à assurer le suivi de l'utilisation aux fins de détection d'éventuelles anomalies, l'actualisation de l'analyse d'impact algorithmique en cas d'évolution significative du système ou de son environnement, ainsi que la réalisation des mises à jour nécessaires au maintien dans la durée de la sécurité et de la performance du système au niveau de l'état de l'art. Deuxièmement, les systèmes risqués devraient faire l'objet d'une supervision humaine tout au long du cycle de vie, garantissant qu'un ou plusieurs opérateurs humains surveillent le fonctionnement du système et vérifient l'absence d'anomalies, qu'ils puissent à tout moment reprendre la main, et que tout dysfonctionnement technique conduise, aussi rapidement que possible, à une intervention humaine. Troisièmement, les personnes physiques ou morales opérant le système d'IA risqué devraient se voir imposer une double obligation de transparence et d'information du public tout au long du cycle de vie, afin de permettre aux autorités de régulation d'exercer un réel contrôle et au public de disposer des informations nécessaires pour exercer leurs droits en justice en cas de préjudice.

426. **Conclusion du §1 relatif à la responsabilisation des acteurs par la création de mécanismes de conformité.** Dans une démarche de régulation proportionnée, il semble déraisonnable d'imposer le même degré de contraintes à tous les systèmes d'IA, qu'importe

leur niveau de risque. Une approche graduelle, par les risques, réservant les normes contraignantes pour les systèmes d'IA susceptibles de générer un risque élevé pour les droits et libertés des individus, apparaît plus adaptée et à-même de remplir le double objectif de prévention des risques et de promotion de l'innovation. L'expérience et la méthodologie en matière de régulation des risques pourra être utilisée afin d'aboutir à un cadre de régulation proportionné : identification des risques, évaluation et classification, gestion. Au titre de la gestion des risques, de nombreuses propositions d'exigences de conformité, visant à la normalisation et à la diffusion de bonnes pratiques déjà ancrées dans le domaine de l'informatique, ont été présentées. Afin de prévenir au mieux les risques graves, ces bonnes pratiques devraient encadrer à la fois la conception des systèmes d'IA risqués, « en amont » de la mise en service, et le maintien en conditions opérationnelles, « en aval ». Les propositions présentées sont issues de la littérature scientifique, de rapports d'institutions publiques et, pour la majeure partie, directement de la pratique en entreprise. Elles ne visent pas l'exhaustivité mais constituent une base de réflexion à comparer et confronter avec les initiatives actuelles de régulation de l'IA, qu'elles soient européenne[860], américaine[861] ou internationale[862].

427. **Transition.** La définition des normes à imposer par la voie de la conformité devrait faire l'objet du plus large consensus entre les parties prenantes afin d'en garantir la pertinence et l'efficacité : entreprises développant les systèmes d'IA, associations représentant les consommateurs, organisations gouvernementales et non gouvernementales… A cet effet, la responsabilisation des acteurs régulés peut aussi passer par leur implication dans l'édiction de ces normes.

---

[860] COMMISSION EUROPÉENNE, *Proposition de règlement du Parlement européen et du Conseil établissant des règles harmonisées concernant l'intelligence artificielle (législation sur l'intelligence artificielle) et modifiant certains actes législatifs de l'Union*, 21 avril 2021, 2021/0106 (COD).

[861] *Algorithmic Accountability Act of 2019*, proposition de loi par les sénateurs C. Booker et R. Wyden devant le Congrès américain, 4 octobre 2019, H.R.2231, 116th U.S. Congress (2019-2020).

[862] UNESCO, *Avant-projet de Recommandation sur l'éthique de l'intelligence artificielle*, *Bibliothèque numérique de l'UNESCO*, 7 septembre 2020, p. 2, disponible en ligne : <https://unesdoc.unesco.org/ark:/48223/pf0000373434_fre>, consulté le 01/04/2021.

§2 : Une responsabilisation des acteurs par leur implication dans la construction des normes contraignantes

428. **La nécessaire co-construction de la norme.** Les destinataires de la norme, à savoir les entreprises qui conçoivent, produisent et vendent des systèmes d'IA, doivent être associés à l'identification des moyens de la régulation. Cela permettrait de garantir que les règles adoptées sont réalistes et adaptées aux contraintes des entreprises, d'une part, et que la régulation soit plus facilement acceptée et mise en œuvre par les acteurs régulés, d'autre part. Cette implication des acteurs régulés dans l'édiction des règles peut se concrétiser de plusieurs façons.

429. **La consultation des parties prenantes.** Premièrement, le choix des principes et règles contraignantes pour encadrer la conception des systèmes d'IA doit faire l'objet d'une large consultation des parties prenantes. Cette consultation, entreprise notamment dans le cadre de la construction du règlement européen sur l'IA[863], doit permettre aux pouvoirs publics de prendre en compte les attentes et contraintes des destinataires de la norme. Il serait également utile de consulter chercheurs et associations de consommateurs de façon à bien identifier les risques que soulèvent les usages de l'IA. Il est primordial que cette consultation se fasse à la fois en amont de la rédaction d'une éventuelle règlementation mais également en aval, puisque les premières propositions de textes législatifs peuvent contenir des dispositions irréalisables en pratique. Or, seules les entreprises qui conçoivent et déploient les systèmes visés par la régulation sont capables d'identifier précisément les mesures disproportionnées ou techniquement irréalisables. Leur implication dans la définition des normes, le cas-échéant par leur intégration dans les instances de normalisation européennes et internationales telles que le Comité européen de normalisation[864], l'ENISA[865] ou l'AFNOR[866] par exemple, est essentielle.

---

[863] COMMISSION EUROPÉENNE, *Public consultation on the AI White Paper : final report*, rapport final présentant les résultats de la consultation publique sur le livre blanc sur l'IA, Novembre 2020, disponible en ligne : <https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=68462>, consulté le 24 janvier 2022.

[864] CEN et CENELEC, *Position Paper – Proposal for a Regulation laying down harmonized rules on artificial intelligence (Artificial Intelligence Act-COM 2021/206)*, papier de position, octobre 2021, 5 p., disponible en ligne : <https://www.cencenelec.eu/media/CEN-CENELEC/AreasOfWork/Position%20Paper/2021/positionpaper_aia_2021.pdf>, consulté le 12 février 2022.

[865] ENISA, *Securing Machine Learning Algorithms*, rapport, 14 décembre 2021, 70 p., disponible en ligne : <https://www.enisa.europa.eu/publications/securing-machine-learning-algorithms>, consulté le 12 février 2022.

[866] Voir l'ensemble des normes sur l'intelligence artificielle en cours de conception par l'AFNOR (27 à la date du 12 février 2022 : <https://norminfo.afnor.org/search?term=intelligence+artificielle>, consulté le 12 février 2022.

430. **L'incitation à l'autorégulation.** Deuxièmement, les entreprises développant des systèmes d'IA peuvent également être incitées à proposer leurs propres solutions pour garantir que les systèmes soient vertueux à la fois pour les droits fondamentaux et pour la protection de l'environnement. Cette incitation peut passer par différents mécanismes juridiques. Par exemple, il est possible de promouvoir la rédaction de codes de bonne conduite par les opérateurs économiques en leur reconnaissant une valeur juridique une fois validés par une autorité de supervision. Se conformer aux dispositions du code de bonne conduite (fixées par les acteurs eux-mêmes) pourrait alors valoir conformité avec la règlementation. Déjà expérimenté sur d'autres sujets tels que la cybersécurité[867], ce mécanisme a prouvé sa popularité auprès des acteurs régulés qui y voient un moyen de définir des règles prenant en compte les contraintes liées à leur activité. Par ailleurs, les États-Unis ont également expérimenté un mécanisme de régulation appelé « *regulatory covenant* », que l'on pourrait traduire par « pacte de régulation ». Ce mécanisme, largement étudié dans la doctrine internationale[868], consiste à donner le choix aux acteurs du secteur à réguler : soit ils s'accordent sur un code de bonne conduite non contraignant qui devra être validé par les pouvoirs publics ; soit, à défaut, les pouvoirs publics soumettent le secteur concerné à une règlementation contraignante[869]. Cette pratique a montré, par le passé, que la perspective d'une règlementation contraignante poussait les entreprises à adopter des démarches d'autorégulation beaucoup plus fortes que lorsque cette menace n'existait pas[870]. Son utilisation dans le cadre de la régulation de l'IA permettrait

[867] ENISA, « Cybersecurity Standards and Certification », *Site officiel de l'ENISA*, disponible en ligne : < https://www.enisa.europa.eu/topics/standards?tab=details>, consulté le 12 février 2022 ; *Règlement (UE) n°2019/881 du Parlement européen et du Conseil du 17 avril 2019 relatif à l'ENISA (Agence de l'Union européenne pour la cybersécurité) et à la certification de cybersécurité des technologies de l'information et des communications, et abrogeant le Règlement (UE) n°526/201*, publié au JOUE n°L 151/15 le 7 juin 2019.

[868] D.D. Hirsch, « Protecting the Inner Environment: What Privacy Regulation Can Learn from Environmental Law », *Georgia Law Review*, 2006, vol. 41, n°1, spec. p. 50-57 ; R.B. Stewart, « A New Generation of Environmental Regulation ? », *Capital University Law Review*, 2001, vol. 29, n°1, spec. 82 ; D.J. Fiorino, « Toward a New System of Environmental Regulation: The Case for an Industry Sector Approach », *Environmental Law*, 1996, vol. 26, n°2, pp. 457-486 ; K. Perine, « The Persuader », *The industry standard*, 13 novembre 2000, 169.

[869] D.D. Hirsch, « Protecting the Inner Environment: What Privacy Regulation Can Learn from Environmental Law », *Georgia Law Review*, 2006, vol. 41, n°1, spec. p. 52.

[870] *Ibid*, spec. p. 53 ; pour un exemple européen de covenant règlementaire, voir en matière environnementale : COMMISSION EUROPÉENNE, *Communication from the Commission to the Council and the European Parliament on Environmental Agreements*, 1996, COM (96)561, spec. p. 22 : « *environmental agreements with industry have an important role to play within the mix of policy instruments sought by the Commission. They can offer cost-effective solutions when implementing environmental objectives and can bring about effective measures*

notamment de lutter contre le phénomène d'*« ethics whasing »* mené par les grandes entreprises du numérique pour éviter une règlementation[871]. Le « pacte de régulation » présente ainsi l'avantage de responsabiliser les acteurs en leur permettant de choisir entre une autorégulation supervisée ou une règlementation imposée.

431. **Le bac à sable règlementaire.** Troisièmement, d'autres mécanismes expérimentaux peuvent être mis en œuvre pour garantir la proportionnalité des règles à imposer. En particulier, le procédé du « bac à sable règlementaire », consistant en la dispense, pour des entreprises souhaitant proposer des produits et des services innovants, de certaines exigences légales et réglementaires qui pourraient être de nature à freiner leur innovation[872], nous semble pertinent dans le contexte de la régulation de l'IA. Déjà expérimenté dans différents secteurs tels que l'énergie[873], la finance[874] ou les voitures autonomes[875], le mécanisme du bac à sable règlementaire répond à deux objectifs. D'une part, il permet de stimuler l'innovation en permettant à des entreprises sélectionnées par des autorités de régulation de développer leurs produits et/ou services en bénéficiant d'un régime juridique favorable. À titre d'exemple, dans le secteur de l'énergie, la loi énergie climat du 9 novembre 2019 a autorisé l'autorité de régulation sectorielle (la CRE) à accorder des dérogations aux conditions d'accès et à l'utilisation des réseaux et installations pour déployer à titre expérimental des technologies ou des services innovants en faveur de la transition énergétique et des réseaux et infrastructures intelligents[876]. Pour son premier guichet, ce sont 9 projets qui ont été sélectionnés pour

in advance of and in supplement to legislation » ; analysé notamm. par S.M. Johnson, « Economics v. Equity II: The European Expérience », *Washington & Lee Law Review*, 2001, vol. 58, 417, spec. 442-444.

[871] R. Ochigame, « How Big Tech Manipulates Academia to Avoid Regulation », *The Intercept (blog)*, 20 décembre 2019, disponible en ligne : <https://theintercept.com/2019/12/20/mit-ethical-ai-artificial-intelligence/>, consulté le 23 janvier 2020.

[872] S. Smatt-Pinelli, « Focus : des regulatory sandbox pour l'innovation juridique », *Revue pratique de la prospective et de l'innovation*, LexisNexis, juillet 2021, n°1, 3.

[873] S. Andrieu, J. Gourdou, « Le bac à sable règlementaire dans le secteur de l'énergie », *Energie – Environnement – Infrastructures*, LexisNexis, novembre 2020, n°11, étude 16.

[874] W.G. Ringe, C. Ruof, « Regulating fintech in the EU: the Case for a Guided Sandbox », *European Journal of Risk Regulation*, septembre 2020, vol. 11, n°3, pp. 604-629 ; S. Rousseau, « La réglementation des cryptomonnaies (ICOs) au Canada : la protection des investisseurs et le bon fonctionnement du marché dans le bac à sable réglementaire », *Revue internationale des services financiers*, 2018, n°1, p. 15.

[875] *Ordonnance n° 2016-1057 du 3 août 2016 relative à l'expérimentation de véhicules à délégation de conduite sur les voies publiques*, publiée au JORF n°0181 du 5 août 2016 ; *Loi n° 2019-486 du 22 mai 2019 relative à la croissance et la transformation des entreprises*, publiée au JORF n°0119 du 23 mai 2019, article 125.

[876] *Loi n° 2019-1147 du 8 novembre 2019 relative à l'énergie et au climat*, publiée au JORF n°0261 du 9 novembre 2019, dite « Loi énergie climat », article 61.

expérimenter particulièrement « les flexibilités locales, le stockage d'électricité et l'injection de méthane de synthèse dans les réseaux »[877]. Ce mécanisme permet donc, par la sélection des projets dérogatoires, à un fléchage de l'innovation vers les technologies d'avenir. Aux termes de l'article 61 de la loi énergie climat, les expérimentations font tout de même l'objet d'obligations particulières : périmètre de la dérogation limitée aux conditions d'accès au réseau, compatibilité des projets avec les objectifs de la politique énergétique, durée limitée de l'expérimentation[878], information des utilisateurs de la nature expérimentale des services[879], publication et évaluation annuelle des projets sélectionnés[880]. D'autre part, le mécanisme du bac à sable règlementaire permet aux pouvoirs publics de tester des modèles de régulation ou de prendre le temps de définir les règles les plus pertinentes pour réguler une activité. En effet, les projets bénéficiant de la souplesse règlementaire font généralement l'objet d'une supervision rapprochée par l'autorité de régulation, qui peut ainsi identifier plus précisément les risques générés par les technologies innovantes. Dans le secteur bancaire et financier, pour favoriser le développement des « Fintechs », certains régulateurs dispensent les entreprises innovantes de règles relatives à l'agrément bancaire ou autres règles prudentielles[881]. L'approche a fait ses preuves partout dans le monde, comme au Royaume-Uni avec la Financial Conduct Authority ou à Singapour et Hong Kong[882]. D'autres autorités de régulation sectorielles se sont inspirées de cette pratique même en l'absence de possibilité légale de fournir des dispenses règlementaires, à l'instar de la CNIL en matière de données à caractère personnel[883]. Lorsque le bac à sable règlementaire n'est pas assorti d'une dispense de conformité à certaines règles contraignantes, l'intérêt pour les entreprises s'en trouve limité.

---

[877] COMMISSION DE RÉGULATION DE L'ÉNERGIE (CRE), *Délibération n° 2021-59 du 11 mars 2021 portant décision sur l'octroi des dérogations des dossiers soumis à la CRE dans le cadre du premier guichet du dispositif d'expérimentation réglementaire prévu par la loi relative à l'énergie et au climat* ; COMMISSION DE RÉGULATION DE L'ÉNERGIE (CRE), « Bac à sable réglementaire : la CRE accorde des dérogations à 9 projets innovants », *Site officiel de la CRE*, 24 mars 2021, disponible en ligne : <https://www.cre.fr/Actualites/bac-a-sable-reglementaire-la-cre-accorde-des-derogations-a-9-projets-innovants>, consulté le 4 janvier 2022.

[878] Loi énergie climat, *op. cit.*, article 61, I.

[879] Loi énergie climat, *op. cit.*, article 61, III.

[880] Loi énergie climat, *op. cit.*, article 61, V.

[881] N. Mathey, G. Bourdeaux, « Vers une régulation des FinTechs ? », *Revue de Droit bancaire & financier*, LexisNexis, 2017, n°2, dossier 15, pt. 15.

[882] L. Ashley, « Why regulators must adapt to fintech », *International Financial Law Review*, 2015, vol. 34, n°6.

[883] CNIL, « « Bac à sable » données personnelles de la CNIL : appel à projets 2021 », *Site officiel de la CNIL*, 15 février 2021, disponible en ligne : <https://www.cnil.fr/fr/bac-a-sable-2021>, consulté le 5 janvier 2022.

L'approche présente alors uniquement l'avantage de concevoir un produit dans une logique partenariale avec l'autorité de régulation, ce qui risque de priver le bac à sable règlementaire de son objectif premier : stimuler l'innovation en permettant aux entreprises volontaires d'évoluer dans un cadre juridique allégé. C'est la raison pour laquelle nous considérons que l'instauration d'un bac à sable règlementaire n'aurait d'utilité dans le cadre de la régulation de l'IA que s'il permettait aux entreprises de s'affranchir, en partie[884], des règles contraignantes qui s'appliqueraient par défaut. En l'échange de souplesses règlementaires relatives aux obligations de conformité susceptibles d'encadrer la conception des systèmes d'IA ou aux régimes de protection des données à caractère personnel, les bénéficiaires pourraient se voir imposer des obligations de transparence envers le public ou de collaboration avec les autorités de régulation. Enfin, l'octroi des dérogations par une autorité administrative compétente ne devrait pas priver les individus de leur droit à un recours effectif (et en particulier à la réparation) en cas de dommage causé par le système d'IA conçu dans un cadre juridique allégé. Cela pourrait se traduire par une responsabilité de l'entreprise en cas de dommage survenu entre le moment de l'octroi de la dérogation et la mise sur le marché, et une responsabilité de l'autorité de supervision à compter de ce moment, puisqu'elle aurait la charge de valider les étapes de mise en conformité de l'entreprise à l'issue de la phase de conception « allégée ». Bien qu'il ne soit pas exempt de toute critique[885], le mécanisme du bac à sable règlementaire nous semble être une option pertinente pour stimuler l'innovation dans l'IA, tout en accompagnant la montée en compétences des autorités publiques.

[884] En effet, le dispositif du bac à sable règlementaire peut consister en la dispense de certaines règles uniquement. Dans le secteur financier, il ne dispense pas les entreprises innovantes du respect des principes clés de la règlementation : lutte contre le blanchiment et le financement du terrorisme (LBT) ou le respect du secret bancaire. V. N. Mathey, G. Bourdeaux, *op. cit.*

[885] Les principales critiques du procédé de bac à sable règlementaire soulèvent qu'il bénéficie souvent à des entreprises qui ne nécessitaient pas une attention accrue du régulateur (i.e. des entreprises déjà agréés dans le secteur financier), que ses modalités sont parfois floues (quant aux règles auxquelles il est possible de déroger ou sa durée) ou encore qu'il contrevient au principe d'égalité en favorisant des entreprises innovantes sur le plan technologique uniquement et en excluant, de fait, toutes les entreprises non technologiques : N. Mathey, G. Bourdeaux, *op. cit.*, spec. pt. 16.

432. **Conclusion du §2 relatif à la responsabilisation des acteurs par leur implication dans la construction de la norme.** La consultation publique, l'incitation à l'autorégulation et la mise en place de bacs à sable règlementaires sont trois exemples de mécanismes permettant d'impliquer les acteurs dans la construction de la norme. Toutefois, il est important de veiller à ce qu'elles ne déresponsabilisent pas les pouvoirs publics qui conservent un rôle primordial dans la prévention des risques liés aux usages de l'IA. Loin d'être antinomiques, les approches incitatives et prescriptives sont complémentaires et vont de paires.

433. **Conclusion de la Section 1 sur les mécanismes de responsabilisation des acteurs du cycle de vie des systèmes d'IA.** Les développements précédents ont permis de démontrer que la co-régulation, consistant à déléguer la réalisation des objectifs fixés par les autorités publiques aux acteurs concernés, apparaît comme l'approche la plus adaptée pour réguler l'IA. En effet, elle se traduit en pratique par la combinaison de mesures prescriptives et incitatives, permettant de trouver un équilibre entre la prévention des risques par l'encadrement de la conception des systèmes d'IA et l'incitation à l'innovation par la souplesse accordée aux acteurs dans leur mise en conformité. Les mécanismes juridiques à mettre en place pour parvenir à cet équilibre sont nombreux. Trois d'entre eux nous paraissent pertinents dans le cadre de la régulation de l'IA : l'encadrement de la conception des systèmes d'IA par des mécanismes de conformité, l'implication des acteurs dans la définition des normes et principes de conception à respecter et l'instauration d'une obligation de transparence pour inciter les acteurs à s'autoréguler et à adopter eux-mêmes un comportement vertueux pour les droits fondamentaux.

Toutefois, les entreprises auront besoin d'être accompagnées dans la définition de ces normes et dans leur mise en œuvre. Pour garantir l'efficacité des règles, il est nécessaire que pèse sur elles la menace de contrôles et sanctions coercitives[886]. Si l'entière responsabilité de la conformité était déléguée aux acteurs régulés, il y aurait de fortes chances que certains d'entre eux profitent de l'absence de supervision externe pour se soustraire aux normes. L'ensemble de ces raisons nous font croire que la co-régulation de l'IA n'est possible qu'à condition que sa mise en œuvre soit supervisée par une autorité de régulation indépendante.

[886] E. Kant, *Philosophy of Law*, in *General Introduction to the Metaphysics of Law*, Section 1, reprinted in *The Great Legal Philosophers*, dir. C. Morris, University of Pennsylvania Press, 1979, p. 240.

Section 2 : Une mise en œuvre de la régulation par une autorité indépendante

434. **Plan.** Sur un sujet aussi technique que l'IA, la mise en œuvre de la régulation proposée dans la thèse risque d'être difficile. En effet, elle contient un nombre important d'exigences de conformité, plusieurs niveaux de risque et le tout s'appliquera à des personnes physiques ou morales qui ne disposent pas forcément de compétences techniques ou juridiques suffisantes pour s'y conformer. Dès lors, la création d'une autorité indépendante compétente dans le domaine de l'IA nous paraît opportune (**§1**) afin de faciliter la mise en œuvre de la régulation. Les compétences, rôles et pouvoirs de cette autorité seront également détaillés (**§2**).

§1 : L'opportunité de la création d'une autorité de régulation compétente dans le domaine de l'IA

435. **L'utilité d'une autorité de régulation.** Dans un domaine évoluant aussi rapidement que celui de l'IA, il est primordial que les entreprises soient accompagnées dans leur mise en conformité. De plus, depuis la fin des années 90, l'absence de supervision publique de l'IA a conduit au défaut de prise en compte des nombreux enjeux juridiques, éthiques et écologiques liés au développement de cette technologie[887]. La création d'une autorité indépendante chargée de la mise en œuvre de la régulation et de l'information du public permettrait de répondre à ces problématiques.

436. **Une utilité démontrée en Droit de l'environnement.** À cet égard, un parallèle peut être dressé avec la construction du Droit de l'environnement dans l'Union européenne et en France. Au niveau européen, l'Agence européenne pour l'environnement (AEE) a été créée par un règlement en 1990[888] avec pour missions de fournir aux institutions européennes et aux États

[887] C. Cath, S. Wachter, B. Mittelstadt, *et al.*, « Artificial Intelligence and the 'Good Society': the US, EU, and UK approach », *Science and Engineering Ethics*, 2018, vol. 24, 505–528.
[888] *Règlement (CEE)* 1210/90 *du Conseil du 7 mai 1990 relatif à la création de l'agence européenne pour l'environnement et du réseau européen d'information et d'observation pour l'environnement*, publié au JOCE n°L120/33 du 11 mai 1990.

membres *« des informations objectives, fiables et comparables au niveau européen qui leur permettent de prendre les mesures nécessaires pour protéger l'environnement, d'évaluer leur mise en œuvre et d'assurer la bonne information du public sur l'état de l'environnement »*[889], ainsi que de leur fournir, à cette fin, *« le support technique et scientifique nécessaire »*[890]. L'AEE bénéficie d'un budget propre constitué d'une subvention inscrite au budget général de l'Union et des rémunérations des services rendus dans le cadre de l'exercice de ses fonctions[891]. Son action vise donc à éclairer les institutions de l'UE et les États membres en assurant la collecte, la standardisation et la fiabilisation de données techniques ou scientifiques nécessaires à la prise de décision politique. Force est de constater que le partage, au niveau européen, de référentiels et de données consolidés pour mesurer à la fois comment les systèmes d'IA sont utilisés, les risques qu'ils génèrent, ou encore leur empreinte environnementale, serait un atout considérable pour parvenir à établir un cadre harmonisé et proportionné de régulation de l'IA.

437. **Une utilité démontrée en France.** On retrouve également ce type d'agences gouvernementales au niveau national avec, par exemple, l'ADEME, en charge de la mise en œuvre des politiques publiques en matière environnementale, ou encore l'Observatoire des impacts environnementaux du numérique[892] et placé auprès de l'ADEME et de l'ARCEP. Ces agences ou missions permettent également de réduire le déficit de connaissances autour de sujets précis et hautement techniques, notamment par la collecte et la diffusion de données[893]. À ce titre, elles rendent possibles la sensibilisation du grand public sur des thématiques nouvelles, ce qui est d'autant plus pertinent au regard des enjeux sociétaux du développement de l'IA et ne fait que confirmer la nécessité de créer une autorité de régulation compétente pour assurer la mise en œuvre de la régulation de l'IA.

---

[889] *Ibid*, article 1, maintenant codifié dans le *Règlement (CE) 401/2009 du Parlement et du Conseil du 23 avril 2009 relatif à l'Agence européenne pour l'environnement et au réseau européen d'information et d'observation pour l'environnement,* publié au JOUE n°L126/13 du 21 mai 2009.
[890] *Ibid.*
[891] *Ibid*, article 11.
[892] Instauré par l'article 4 de la *Loi n°2021-1485 du 15 novembre 2021 visant à réduire l'empreinte environnementale du numérique en France*, publiée au JORF n°0266 du 16 novembre 2021.
[893] Voir à ce propos la *Loi n° 2021-1755 du 23 décembre 2021 visant à renforcer la régulation environnementale du numérique par l'Autorité de régulation des communications électroniques, des postes et de la distribution de la presse*, publiée au JORF n°0299 du 24 décembre 2021.

§2 : Les modalités de la mise en œuvre de la régulation par une autorité indépendante

438. **Le périmètre d'une autorité de régulation dans le domaine de l'IA.** L'autorité de régulation en matière d'IA devrait logiquement avoir pour périmètre d'action un domaine correspondant au champ d'application des exigences de conformité proposées dans la présente thèse. Sa compétence devrait donc exclure tous les systèmes logiciels ne répondant pas à la définition des systèmes d'IA retenue dans la régulation. À cet égard, il est important de circonscrire son champ de compétence afin de prévenir d'éventuels risques de concurrence entre autorités de régulation. En effet, elle pourrait disposer d'un périmètre proche de deux autres types d'autorités. D'une part, son périmètre pourrait entrer en conflit avec les autorités compétentes en matière de protection des données (la CNIL en France) et de sécurité des systèmes d'information (l'ANSSI en France). Dans ce cas, l'autorité compétente sur l'IA devrait intervenir uniquement en soutien à ces autorités dont le fonctionnement est éprouvé par le temps. Concrètement, dans le cas de la CNIL, si une plainte concerne un système d'IA traitant des données personnelles, les deux autorités seront amenées à travailler ensemble afin de la traiter et, le cas-échéant, aboutit à une sanction commune. D'autre part, l'autorité proposée pourrait également entrer en conflit avec certaines autorités sectorielles. Dans le secteur de l'énergie, il s'agirait par exemple de la CRE ou de l'ASN. Dans ce cas, l'autorité compétente sur l'IA devrait agir en qualité de coordonnatrice et accompagnatrice des actions sectorielles (par exemple en organisant la rédaction de standards sectoriels). Il sera important d'organiser la coopération entre l'ensemble de ces autorités. Typiquement, dans le cadre d'une procédure de sanction, l'autorité compétente sur l'IA pourra avoir besoin de l'appui d'une autorité sectorielle pour mesurer la gravité des faits. À l'inverse, les autorités sectorielles ne disposeront pas forcément de compétences techniques suffisantes pour assurer la déclinaison sectorielle de la régulation sur l'IA : elles devront donc être elles aussi accompagnées.

439. **Les missions d'une autorité de régulation dans le domaine de l'IA.** Le rôle de l'autorité que nous appelons de nos vœux serait à la fois d'alimenter le débat public sur les risques et enjeux liés aux usages de l'IA par des informations fiables (risques pour les droits fondamentaux, quantification de l'empreinte environnementale…), de veiller à assurer la participation du public et des parties prenantes à la construction de la régulation de l'IA mais aussi, le cas-échéant, d'exercer un contrôle *ex-post* sur la mise en œuvre des mécanismes de conformité.

440. **Les pouvoirs d'une autorité de régulation dans le domaine de l'IA.** Pour accomplir ses missions, il est essentiel que l'autorité proposée soit dotée de pouvoirs d'enquête et de sanctions, ainsi que des moyens techniques et humains nécessaires à l'exercice de ses fonctions, sur le modèle de la CNIL ou des autorités de régulation financière telles que l'ACPR[894]. Ces moyens techniques, humains et financiers doivent notamment permettre à l'autorité d'instruire les plaintes déposées par les individus, ce qui est une condition à l'effectivité des droits qui leur seraient reconnus. De plus, l'effectivité de la régulation de l'IA dans son ensemble ne sera garantie que si elle est accompagnée de la menace de sanctions financières importantes. Cette proposition, inspirée de l'expérience dans d'autres secteurs tels que la protection de l'environnement ou la régulation financière, fait également consensus dans la doctrine internationale[895]. Au niveau français, le Conseil d'État a proposé dans une étude publiée le 30 août 2022 que la CNIL soit désignée comme autorité compétente dans le cadre de la règlementation européenne de l'IA[896]. Cette proposition est pertinente puisqu'elle capitalise sur les expertises à la fois technique et pédagogique développées par la CNIL.

441. **Transition.** Toutefois, si la finalité de la régulation de l'IA est bien la protection des individus contre les atteintes à leurs droits fondamentaux, la seule règlementation de la conception des systèmes d'IA ne peut suffire. En effet, elle ne vise qu'à prévenir les risques et non à garantir aux individus des voies de recours effectives pour faire valoir leurs droits, signaler des préjudices et, le cas-échéant, obtenir réparation. C'est la raison pour laquelle notre troisième proposition vise à la reconnaissance, dans la régulation de l'IA, de droits individuels spécifiques, à l'instar de ceux créés par le RGPD.

---

[894] Sur le rôle des autorités de régulation dans le secteur financier, V. notamm. A.D. Merville, « Autorités européennes de supervision financière », *Répertoire des sociétés*, Dalloz, 45 p. ; A. Boujeka, « Vers un modèle de régulation des marchés financiers dans l'Union européenne », *Recueil Dalloz*, 2012, p. 1355.

[895] K. Yeung, A. Howes, G. Pogrebna, « AI Governance by Human Rights-Centred Design, Deliberation and Oversight: An End to Ethics Washing », in *The Oxford Handbook of AI Ethics*, M. Dubber, F. Pasquale (dir.), Oxford University Press, 2019, spec. p. 9 ; N.A. Smuha, « Beyond the Individual: Governing AI's Societal Harm », *Internet Policy Review*, 30 septembre 2021, vol. 10, n°3, spec. p. 23 ; F. Zuiderveen Borgesius, *Discrimination, Artificial Intelligence and Algorithmic Decision-Making*, Rapport à la demande du Conseil de l'Europe, 2018, spec. p. 66, disponible en ligne : <https://rm.coe.int/discrimination-artificial-intelligence-and-algorithmic-decision-making/1680925d73>, consulté le 26 novembre 2021.

[896] CONSEIL D'ÉTAT, *Intelligence artificielle et action publique : construire la confiance, servir la performance*, Etude à la demande du premier Ministre, 30 août 2022, 360 p.

Section 3 : La création de droits individuels effectifs

442. **La pertinence de la création de droits individuels dans les modèles de régulation.** Face à la multiplication des prises de décision fondées sur des systèmes d'IA[897], de nombreux chercheurs recommandent la reconnaissance de droits aux individus pour leur permettre d'exercer un contrôle effectif sur les décisions algorithmiques prises à leur égard[898]. L'idée n'est pas nouvelle et a déjà été concrétisée dans d'autres domaines. Ainsi la construction du Droit européen de l'environnement a conduit à la reconnaissance des droits à l'information[899], à la participation du public[900] et d'accès la justice[901], permettant aux administrés de contester les décisions publiques qui viendraient contrevenir à leurs droits ou aux principes du droit de l'environnement d'une manière générale[902]. On retrouve également l'utilisation des droits individuels comme instrument de régulation dans la Loi pour une République numérique qui, dès 2016, précise les droits des administrés au regard des décisions individuelles les concernant et prises sur le fondement d'un traitement automatisé de données[903]. Peu après, le RGPD est

---

[897] Le constat se vérifie dans tous les domaines, de la décision publique à la recommandation de contenu en ligne (V. notamm. H. Surden, « Artificial Intelligence and Law: An Overview », *Georgia State University Law Review*, 28 juin 2019, vol. 35, 1305). Le secteur de l'électricité n'est pas épargné, que ce soit dans la prise de décisions relatives à la maintenance d'ouvrages de production d'électricité, à la gestion des infrastructures des réseaux de transport et de distribution d'électricité ou à la relation client.

[898] Autant au niveau européen (N.A. Smuha, *op. cit.*, spec. p. 16) qu'aux États-Unis (M.E. Kaminski, J.M. Urban, « The Right to Contest AI », *Columbia Law Review*, 16 novembre 2021, vol. 121, n°7, 92 p.).

[899] L. Krämer, « Transnational Access to Environmental Information », *Transnational Environmental Law*, 2012, vol. 1, n°, 95–104.

[900] Le principe de participation du public n'est pas reconnu expressément dans les textes fondateurs du Droit européen, mais la doctrine s'accorde pour dire qu'il contribue « à la réalisation concrète » des autres principes qui y sont affirmés : B. Jadot, « La participation du public en droit communautaire de l'environnement, à l'heure de la convention d'Aarhus », in *La participation du public aux décisions de l'Administration en matière d'aménagement et d'environnement*, R. Hostiou, J.F. Struillou, Les cahiers du GRIDAUH, 2007, n° 17, p. 40 ; cité notamm. par F. Jamay, « Principe de participation », *JurisClasseur Environnement et Développement durable*, version mise à jour du 3 août 2021, Fasc. 2440.

[901] C. Poncelet, « Access to Justice in Environmental Matters : Does the European Union Comply with its Obligations ? », *Journal of Environmental Law*, 2012, vol. 24, n°2, 287–309.

[902] Ces trois droits reconnus aux individus reflètent les trois grands principes contenus dans la Convention d'Aarhus, ayant inspiré la construction d'un Droit de l'environnement à l'échelle internationale. V. *Convention d'Aarhus du 25 juin 1998, ratifiée par la France le 8 juillet 2002 par la Loi n° 2002-285 du 28 février 2002 autorisant l'approbation de la Convention d'Aarhus*, publiée au JORF le 1er mars 2002, p. 3904 ; B. Drobenko, « La Convention d'Aarhus et le droit français », *Revue Juridique de l'Environnement*, numéro spécial, 1999, p. 37.

[903] L. Cluzel-Métayer, « La loi pour une République numérique : l'écosystème de la donnée saisi par le droit », *AJDA*, 2017 p.340 ; N. Martial-Braz, « Les nouveaux droits des individus consacrés par la loi pour une République numérique. Quelles innovations ? Quelle articulation avec le Règlement européen ? », *Dalloz IP/IT*, 2016, p. 525.

venu concrétiser les droits des individus sur les données les concernant : information[904], accès[905], rectification[906], effacement[907], limitation des données[908], portabilité[909], opposition[910], ainsi que le droit à ne pas faire l'objet d'une décision entièrement automatisée[911]. Nombre d'entre eux s'appliquent déjà aux systèmes d'IA, mais uniquement lorsqu'ils sont utilisés aux fins de traitement de données à caractère personnel[912]. Or, les risques inhérents aux utilisations de l'IA ne concernent pas seulement les données personnelles, mais également la sécurité et l'égalité de traitement des individus, ou tout autre effet néfaste sur leur environnement. En cela, le RGPD, en son état actuel ou amendé, ne peut suffire à couvrir tous les champs de la régulation de l'IA[913]. En particulier, les droits contenus dans le RGPD ne s'appliquent qu'aux traitements de données à caractère personnel et ne peuvent traiter des risques générés par des systèmes d'IA qui opèreraient sur des données anonymisées ou non personnelles. Toutefois, le RGPD peut être utilisé comme modèle[914]. En effet, ce texte présente une approche équilibrée entre supervision publique (avec le rôle attribué aux autorités nationales de protection des données), responsabilisation des entreprises (avec les principes relatifs aux traitements contenus dans le texte) et création de droits individuels[915]. Loin d'être antagonistes, la reconnaissance de droits

[904] RGPD, articles 12 à 14 ; C. De Terwangne, K. Rosier, *Le Règlement général sur la protection des données (RGPD / GDPR), Analyse approfondie*, Coll. du CRIDS, Larcier, 1re édition, 2018, pp. 4-34 ; M.E. Kaminski, « The Right to Explanation, Explained », *Berkeley Technology Law Journal*, 2018, vol. 34, n°1.
[905] RGPD, article 15 ; C. De Terwangne, K. Rosier, *op. cit.*, pp. 169-177.
[906] RGPD, article 16 ; C. De Terwangne, K. Rosier, *op. cit.*, pp. 51-54.
[907] RGPD, article 17 ; J. Groffe, « Du droit à l'oubli jurisprudentiel au droit à l'effacement (« droit à l'oubli ») dans le RGPD et la loi française. Retour sur l'histoire du droit à l'oubli » in *Le règlement général sur la protection des données, aspects institutionnels et matériels*, dir. A. Bensamoun, B. Bertrand, Mare et Martin, 2020.
[908] RGPD, article 18 ; C. De Terwangne, K. Rosier, *op. cit.*, pp. 100-104.
[909] RGPD, article 20 ; J. Bellesort, « Le droit à la portabilité des données à caractère personnel », in *Le règlement général sur la protection des données, aspects institutionnels et matériels*, dir. A. Bensamoun, B. Bertrand, Mare et Martin, 2020.
[910] RGPD, article 21 ; C. De Terwangne, K. Rosier, *op. cit.*, pp. 169-177.
[911] RGPD, article 22 ; GROUPE DE TRAVAIL « ARTICLE 29 » SUR LA PROTECTION DES DONNÉES, *Guidelines on Automated individual decision-making and Profiling for the purpose of Regulation 2016/679*, 3 octobre 2017, n°WP251.
[912] U. Pagallo, P. Casanovas, R. Madelin, « The Middle-out Approach: Assessing Models of Legal Governance in Data Protection, Artificial Intelligence, and the Web of Data », *The Theory and Practice of Legislation*, 2 janvier 2019, vol. 7, n°1, 25, spec. p. 13.
[913] *Ibid* ; sur les lacunes du droit à l'information du RGPD lorsqu'il est appliqué aux decisions automatisées : S. Wachter, B. Mittelstadt, L. Floridi, « Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation », *International Data Privacy Law*, 7, mai 2017, nº 2, p. 76-99.
[914] S. Wrigley, « Taming Artificial Intelligence: "Bots," the GDPR and Regulatory Approaches » in *Robotics, AI and the Future of Law*, dir. M. Corrales, M. Fenwick, N. Forgó, Springer, 2018, pp. 183-208.
[915] N.A. Smuha, *op. cit.*, spec. p. 23, note 14.

individuels apparaît comme complémentaire à l'approche prescriptive édictant des principes relatifs aux traitements de données personnelles[916]. Ainsi, il s'agit là d'un puissant instrument juridique pour donner aux individus un pouvoir de contrôle effectif sur une technologie présentant des risques.

443. **Plan.** Plusieurs droits pourraient être reconnus aux individus pour compléter ceux présents dans le RGPD et garantir un contrôle effectif sur les systèmes d'IA pouvant causer des atteintes aux droits fondamentaux. Nos propositions s'inspirent à la fois des droits individuels reconnus par le Droit de l'environnement, des droits des personnes concernées créés par le RGPD, et d'autres propositions de la doctrine[917]. En particulier, elles reprennent les trois grandes catégories de droits « aarhusiens »[918], sur lesquels nous bénéficions d'un recul suffisant[919] : la participation (**§1**) et l'information (**§2**) du public, ainsi que l'accès à la justice (**§3**), appliqués au contexte de la régulation de l'IA.

---

[916] M.E. Kaminski, « Binary Governance: Lessons from the GDPR's Approach to Algorithmic Accountability », *Southern California Law Review*, 2019, vol. 92, n°6, 1529–1616.

[917] N.A. Smuha, E. Ahmed-Rengers, A. Harkens, *et al.*, « How the EU can achieve legally trustworthy AI: A response to the European Commission's proposal for an Artificial Intelligence Act », *SSRN Electronic Journal*, 5 août 2021, spec. p. 52 et s. ; M.E. Kaminski, J.M. Urban, « The Right to Contest AI », *Columbia Law Review*, 16 novembre 2021, vol. 121, n°7, 92 p. ; N.A. Smuha, « Beyond the Individual: Governing AI's Societal Harm », *Internet Policy Review*, 30 septembre 2021, vol. 10, n°3, spec. p. 22.

[918] En référence à la Convention d'Aarhus : Convention d'Aarhus du 25 juin 1998 ratifiée par la France le 8 juillet 2002 par la *Loi n° 2002-285 du 28 février 2002 autorisant l'approbation de la Convention d'Aarhus*, publiée au JORF le 1er mars 2002, p. 3904 ; les droits « aarhusiens » ont eux-mêmes été repris du Principe 10 de la Déclaration de Rio de 1992 : *« La meilleure façon de traiter les questions d'environnement est d'assurer <u>la participation de tous les citoyens</u> concernés, au niveau qui convient. Au niveau national, chaque individu doit avoir dûment <u>accès aux informations relatives à l'environnement</u> que détiennent les autorités publiques […] et avoir la possibilité de <u>participer aux processus de prise de décision</u>. […] <u>Un accès effectif à des actions judiciaires et administratives</u>, notamment des réparations et des recours, doit être assuré »* (*Déclaration de Rio sur l'environnement et le développement adoptée par la conférence des Nations unies sur l'environnement et le développement*, Rio de Janeiro, 1992, Principe 10. Disponible en ligne : <https://www.un.org/french/events/rio92/rio-fp.htm>, consulté le 30 avril 2021).

[919] D.D.M. Babor, « Environmental Rights in Ontario: Are Participatory Mechanisms Working », *Colorado Journal of International Environmental Law and Policy*, 1999, vol. 10, n°1998, p. 121-135.

§1 : Le droit à la participation du public

444. **La participation du public à des arbitrages sociétaux : l'exemple du Droit de l'environnement.** L'utilisation de l'IA soulève, en dépit de ses potentiels bénéfices pour la société, d'importantes questions juridiques, éthiques et écologiques. Comme en matière environnementale, la majorité de ces problématiques sont sociétales, dans le sens où elles touchent à la société dans son ensemble et non pas seulement les individus pris isolément, et même transgénérationnelles[920]. Elles appellent donc, dès aujourd'hui, des arbitrages politiques entre prévention des risques et promotion de l'innovation. Puisque ces choix reviennent finalement à définir la société dans laquelle nous souhaitons vivre, la participation du public apparaît comme une condition indispensable si l'on se trouve dans une société démocratique[921]. Le procédé n'est pas nouveau et les modalités de sa mise en œuvre en Droit de l'environnement peuvent servir d'exemple. Consacrée en tant que principe à l'article 7 de la Charte constitutionnelle de l'environnement[922], la participation du public fait l'objet d'un chapitre dédié dans le Code de l'environnement[923]. Ce dernier comporte plusieurs procédures de participation visant à la participation du public aux décisions publiques ayant une incidence sur l'environnement, placées sous la responsabilité du Ministère de la transition écologique[924]. On distingue les procédures « en amont » de la décision ayant une incidence sur l'environnement, des procédures « en aval », consultant le public sur la base d'études d'impact obligatoires pour certains projets d'aménagement du territoire ayant une incidence sur l'environnement. Chacune est encadrée par des modalités précises, fixées par la loi. Elles sont la plupart du temps

[920] Sur l'empreinte environnementale de l'IA, la question n'est pas de prévenir des préjudices directs et actuels liés aux émissions de gaz à effet de serre, mais plutôt de lutter contre une dégradation future de l'environnement, telle que les générations futures verraient leurs conditions de vie nettement diminuée. La question se pose exactement de la même manière en matière environnementale (O. Sutterlin, *op. cit.*, 6).

[921] N.A. Smuha, E. Ahmed-Rengers, A. Harkens, *et al.*, « How the EU can achieve legally trustworthy AI: A response to the European Commission's proposal for an Artificial Intelligence Act », *SSRN Electronic Journal*, 5 août 2021.

[922] Charte de l'environnement, article 7 : *« Toute personne a le droit, dans les conditions et les limites définies par la loi, d'accéder aux informations relatives à l'environnement détenues par les autorités publiques et de participer à l'élaboration des décisions publiques ayant une incidence sur l'environnement ».*

[923] Code de l'environnement, Chapitre III « *Participation du public aux décisions ayant une incidence sur l'environnement* », Articles L123-1-A à L123-19-11.

[924] MINISTÈRE DE LA TRANSITION ÉCOLOGIQUE ET SOLIDAIRE, « Le cadre de la participation du public au titre du code de l'environnement », *Site officiel du Ministère de la transition écologique*, 7 février 2019, disponible en ligne : <https://www.ecologie.gouv.fr/cadre-participation-du-public-au-titre-du-code-lenvironnement>, consulté le 29 novembre 2021.

obligatoires, ce qui garantit aux individus un véritable droit à la participation[925]. Plusieurs d'entre elles nous sembleraient intéressantes à répliquer dans le cadre de la régulation de l'IA.

445. **La participation du public « en amont » de l'utilisation de l'IA.** Des procédures de participation du public pourraient d'abord être rendues obligatoires « en amont » de la mise en service ou sur le marché des systèmes d'IA susceptibles de présenter un risque élevé pour les droits fondamentaux.

446. **Le débat public et la concertation préalable transposés au domaine de l'IA.** Premièrement, le « débat public » et la « concertation préalable » sont les procédures qui interviennent le plus amont, dès le début du processus décisionnel, lorsque toutes les options sont ouvertes y compris celles de renoncer au projet ou de le faire autrement[926]. Le débat public est un dispositif participatif précis dont les principes sont fixés par le Code de l'environnement[927]. Ses modalités sont fixées au cas par cas par une autorité administrative indépendante, la Commission Nationale du Débat Public (CNDP), qui supervise également son déroulement. La « concertation préalable » est un dispositif participatif plus souple et non supervisé, visant au recueil de l'ensemble des avis des parties prenantes et/ou du grand public sur un projet, sans lier l'autorité à son origine[928]. Les méthodes permettant le recueil de l'avis du public sur un projet sont variées et la CNDP en a l'expérience : réunion publique, atelier participatif, questionnaire en ligne, conférence de citoyens tirés au sort, ou autres formes de consultations[929]. L'autorité en charge de la mise en œuvre de la régulation de l'IA pourrait utilement prévoir une coopération étroite avec la CNDP qui dispose de l'expérience nécessaire au déploiement d'une campagne de consultation et à l'organisation des débats publics. Le modèle des enquêtes publiques menées par la CNIL peut également être pris en exemple[930]. Ces

[925] *Ibid.*
[926] Code de l'environnement, articles L121-1-A à L121-24 ; COMMISSION NATIONALE DU DÉBAT PUBLIC (CNDP), « La participation du public et la CNDP », *Site officiel de la CNDP*, 31 mars 2021, disponible en ligne : <https://www.debatpublic.fr/participation-et-environnement-692>, consulté le 29 novembre 2021.
[927] Code de l'environnement, articles L121-8 et s.
[928] Code de l'environnement, articles L121-16 à L121-16-2.
[929] COMMISSION NATIONALE DU DÉBAT PUBLIC (CNDP), « Méthodes et outils », *Site officiel de la CNDP*, 31 mars 2021, disponible en ligne : <https://www.debatpublic.fr/methodes-et-outils-665>, consulté le 30 novembre 2021.
[930] Voir le rapport de synthèse du débat public sur les algorithmes : CNIL, *Comment permettre à l'homme de garder la main ? Les enjeux éthiques des algorithmes et de l'intelligence artificielle*, Synthèse du débat public

dispositifs participatifs doivent permettre de recueillir l'avis du grand public sur la régulation de l'IA. À condition de s'accompagner d'une information complète et accessible, ces campagnes pourraient viser à prendre l'avis des individus sur des usages de l'IA susceptibles de générer des risques importants pour les droits fondamentaux, tels que la reconnaissance faciale dans l'espace public ou, dans le secteur de l'énergie, l'analyse détaillée des données de consommation énergétique à des fins écologiques.

447. **La conciliation transposée au domaine de l'IA.** Deuxièmement, la procédure de « conciliation » prévue à l'article L121-2 du Code de l'environnement prévoit que, « *dès lors que le maître d'ouvrage d'un projet et une association agréée de protection de l'environnement en font la demande commune, une conciliation peut être mise en œuvre par la commission nationale du débat public. Cette procédure est non-suspensive et a notamment vocation à rétablir le dialogue entre les parties à une procédure de participation* »[931]. Un dispositif similaire permettrait de garantir le droit à la participation du public dans la régulation de l'IA dans la mesure où il agit comme un mode alternatif de règlement des conflits ou de précontentieux. En effet, une association de défense des droits fondamentaux pourrait s'opposer au déploiement de systèmes d'IA qu'elle considèrerait trop risqués, sur la base des informations rendues publiques en vertu de l'obligation de transparence incombant aux fournisseurs. La « conciliation » permettrait d'instaurer une phase de dialogue entre les parties prenantes en amont du dépôt d'une plainte auprès de l'autorité de régulation compétente et d'une éventuelle procédure contentieuse. Le dialogue ainsi créé pourrait conduire à une autorégulation des fournisseurs et à la prise en compte, en amont de la mise en service, des droits fondamentaux.

448. **La participation du public « en aval » de l'utilisation de l'IA.** D'autres procédures pourraient être utilisées « en aval » pour s'assurer que la mise en œuvre de la régulation de l'IA produit ses effets et est toujours proportionnée, ou si elle nécessite une adaptation. Par « aval », nous entendons le moment où le projet est bien défini. On ne discute plus de l'opportunité de

---

animé par la CNIL dans le cadre de la mission de réflexion éthique confiée par la loi pour une République Numérique, 15 décembre 2017.

[931] MINISTÈRE DE LA TRANSITION ÉCOLOGIQUE ET SOLIDAIRE, « Le cadre de la participation du public au titre du code de l'environnement », *Site officiel du Ministère de la transition écologique*, 7 février 2019, disponible en ligne : <https://www.ecologie.gouv.fr/cadre-participation-du-public-au-titre-du-code-lenvironnement>, consulté le 29 novembre 2021.

mener le projet, mais uniquement de ses modalités. Le public pourrait donc être consulté sur les caractéristiques principales d'un système d'IA susceptible d'être déployé et dont l'analyse d'impact a mis en lumière un risque important pour les droits fondamentaux. À la manière de l'enquête publique en matière environnementale[932], le public serait consulté sur la base d'un dossier de présentation du projet, contenant notamment l'analyse des risques pour les droits fondamentaux. L'enquête pourrait être conduite par un commissaire enquêteur impartial ou placée sous la responsabilité de l'autorité de régulation indépendante. Le rapport issu de l'enquête devrait permettre à l'autorité de régulation de l'IA de prendre en compte les intérêts des tiers et de déterminer si le risque pour les droits fondamentaux généré par le système d'IA concerné est, ou non, disproportionné par rapport aux bénéfices attendus. Cette décision pourrait même conduire, à terme, à l'amendement du cadre règlementaire, en intégrant, ou non, le système dans une catégorie « à haut risque » ou « prohibée ». Au-delà de cette enquête publique sur la base d'un dossier étayé, des mécanismes de concertation continus serait utiles pour permettre aux individus d'exprimer leur point de vue sur la règlementation d'une manière générale, sur des inquiétudes relatives à des applications de l'IA, ou plus globalement sur des interrogations. Ces mécanismes peuvent prendre la forme de consultation par voie électronique sur le modèle de l'article L123-19 du Code de l'environnement si la consultation porte sur une thématique en particulier, ou de plateformes participatives en ligne comme le site « projets-environnement.gouv.fr », répertoriant les projets soumis à étude d'impact environnemental et fournissant des informations détaillées sur chacun d'eux, y compris le contenu de l'analyse d'impact[933]. On pourrait imaginer le recensement de tous les systèmes d'IA « à haut risque » sur un site internet géré par l'autorité de régulation en charge de la mise en œuvre du cadre règlementaire et de l'information du public. Ce dernier permettrait à tout un chacun d'avoir accès à des informations générales sur les logiciels concernés, la façon dont ils sont utilisés et dont ils ont été conçus, ainsi que, le cas-échéant, de l'analyse d'impact pour les droits fondamentaux réalisés par le fournisseur. Il servirait de base à l'exercice du droit à la participation du public, en ce qu'il permettrait à chacun à la fois d'avoir une information fiable

[932] Code de l'environnement, articles L123-1 et L123-2.
[933] MINISTÈRE DE LA TRANSITION ÉCOLOGIQUE ET SOLIDAIRE, « Consultation des projets soumis à étude d'impacts », *Site gouvernemental*, disponible en ligne : <https://www.projets-environnement.gouv.fr/pages/home/>, consulté le 4 décembre 2021.

et complète et de pouvoir remonter ses préoccupations dans un « forum d'échange » consultable par les individus, industriels, universitaires et autorités de régulation. Bien entendu, cette proposition n'est pas neutre technologiquement puisque ne pourraient être répertoriés que les systèmes d'IA existants. Toutefois, l'objectif de cette plateforme est de fournir une information concrète et transparente aux individus sur les systèmes susceptibles de les affecter. Il est donc normal que n'y figurent que des systèmes déjà mis en service.

449. **Conclusion sur la forme du droit à la participation du public appliqué à l'IA.** Ainsi, les possibles modalités de l'exercice du droit à la participation du public dans la régulation de l'IA sont nombreuses. D'abord, il peut être général et porter le développement de la technologie et la régulation, ou spécifique et ne viser qu'une application précise de l'IA. Ensuite, le droit à la participation du public devrait également pouvoir s'exercer tant en amont du déploiement d'une application présentant des risques pour les droits fondamentaux, qu'en aval puisque l'usage de l'IA peut évoluer au cours de son utilisation ou les risques peuvent être découverts *a posteriori*. Enfin, il peut être concrètement mis en œuvre *via* des mécanismes ayant déjà été expérimentés en Droit de l'environnement : consultation publique et sondage, plateformes participatives type « forum » ou encore débat publique supervisé et organisé par une autorité indépendante telle que la CNDP.

450. **Le contenu de la consultation du public.** Si les modalités pratiques de mise en œuvre d'un droit à la participation du public dans la construction d'une régulation proportionnée de l'IA peuvent prendre différentes formes, il convient également de réfléchir au contenu de la consultation. Concrètement, il s'agit de répondre à la question « sur quoi le public doit-il être consulté ? ». La participation du public apparaît pertinente tant dans la décision d'autoriser le déploiement un système d'IA présentant des risques pour les droits fondamentaux que dans le choix des modalités de la régulation. L'objectif poursuivi est la prise en compte des intérêts du public dans la mise en balance des risques des applications d'IA avec les bénéfices attendus. Pour y parvenir, les individus doivent être consultés sur de nombreux aspects de la régulation. Premièrement, le public doit pouvoir se prononcer sur le champ d'application de la régulation de l'IA, à savoir sur la définition des systèmes d'IA présentant un risque suffisamment important et justifiant ainsi une attention particulière du régulateur. La nature sociétale du risque généré par de nombreuses applications de l'IA rend indispensable son appréciation par le public. Deuxièmement, le public peut également être consulté au cours du processus d'élaboration des normes devant encadrer les systèmes considérés comme risqués. Troisièmement, le public doit pouvoir faire valoir ses intérêts de façon continue au cours de la

mise en œuvre de la régulation. Il doit être associé à toute évolution significative du cadre règlementaire et devrait pouvoir en suggérer, le cas-échéant *via* la plateforme participative évoquée précédemment[934]. La perception sociétale d'un risque pouvant évoluer[935], les catégories de systèmes d'IA suivant leur niveau de risque doivent en effet pouvoir être remises en question à l'initiative des individus. En somme, on peut retenir que le public doit être impliqué dès que la décision publique porte sur le contenu de la régulation, sa mise en œuvre ou sur l'autorisation du déploiement d'un système présentant des risques importants pour les droits fondamentaux. Chacune de ces décisions réalise un arbitrage entre les risques générés par certains systèmes d'IA et les bénéfices attendus, lequel relève, selon nous, d'une dimension sociétale, d'où l'importance du droit à la participation du public.

451. **La finalité de la participation du public : la prise en compte des intérêts des individus dans la régulation.** Ce processus a tout de même des limites. On peut déjà le constater en matière environnementale avec, par exemple, un activisme fort contre des technologies telles que la fission nucléaire[936] ou l'éolien en mer[937] malgré leur nécessité pour atteindre les objectifs de neutralité carbone. Les attentes du public sur une technologie et son acceptation des risques[938] doivent être relativisés au regard de leur défaut d'information sur certains sujets. La consultation d'un public ne disposant pas d'une information complète ne pourra aboutir qu'à un résultat biaisé. Il est donc important que les pouvoirs publics soient également éclairés par l'avis expert des autorités de régulation, du secteur académique, des entreprises privées, ou de toute autre personne ou institution jouissant d'une expertise technique de la technologie à

[934] Voir Supra, 448.
[935] P. Slovic, E. Peters, « Risk Perception and Affect », *Current Directions in Psychological Science*, 2006, vol. 15, 6, 322-325.
[936] M. Chambru, « La publicisation du risque nucléaire par les usages protestataires du droit », *Sciences de la société*, 2017, 100, pp. 79-91 ; M. Chambru, « La protestation antinucléaire par-delà les frontières : mutations et temporalités de l'enjeu européen », *Communication & Organisation*, Presses Universitaires de Bordeaux, 2020, vol. 57, pp. 121-133.
[937] TA Limoges E20000066/87 EP EOL BEAULIEU, 19 avril 2021, *Conclusions et avis de la commission d'enquête publique, sur la demande d'autorisation unique présentée par la société d'exploitation éolienne de Beaulieu en vue d'obtenir l'autorisation d'exploiter un parc éolien de quatre aérogénérateurs et d'un poste de livraison sur le territoire de la commune de Beaulieu (Indre)*, spec. p. 2.
[938] En matière de vie privée, nombreux sont les individus qui ne s'intéressent pas à sa protection tant « ils n'ont rien à cacher ». Pourtant, ils n'en restent pas moins sensibles au sujet de la cybersécurité, du risque d'usurpation d'identité et autres risques pour leur vie privée (ce constant contradictoire est connu dans le monde académique sous le nom de « privacy paradox » ou « paradoxe de la vie privée » : S.B. Barnes, « A privacy paradox: Social networking in the United States », *First Monday*, 2006, vol. 11, n°9).

réguler. La participation du public, si elle doit être un droit pour les individus, ne peut constituer qu'une des pierres à l'édifice d'une régulation proportionnée de l'IA. Son objectif ne doit pas être la définition des modalités de la régulation par le grand public mais plutôt la prise en compte des intérêts des individus dans sa construction, par un procédé résolument démocratique.

452. **Un exemple prospectif d'application dans le secteur de l'électricité.** La mise en œuvre d'un droit à la participation du public aurait toute sa place dans le secteur de l'électricité, dans lequel de nombreux cas d'usage de l'IA peuvent générer des risques importants. C'est le cas, notamment, lorsque des systèmes d'IA sont utilisés pour automatiser certaines actions dans la gestion du réseau électrique. En effet, lorsque les actions automatisées concernent le rétablissement du réseau en cas de coupure ou la décision sur les zones géographiques à couper du réseau en cas de demande trop importante par rapport à la production d'électricité, le défaut de fonctionnement du système d'IA peut avoir de lourdes conséquences. En effet, toute erreur ou dysfonctionnement pourrait dans ces cas aboutir à une coupure d'électricité de longue durée (le temps qu'un humain intervienne pour débuguer ou reprendre la main sur le système) ou à une coupure injustifiée[939]. Ces coupures sont des atteintes au principe de sécurité d'approvisionnement[940] et peuvent causer d'importants préjudices aux individus, entreprises ou organismes privés d'électricité : pertes financières du fait d'une usine à l'arrêt, impossibilité de soins dans les hôpitaux… Dans l'hypothèse où un système d'IA serait utilisé pour automatiser ces fonctions essentielles de la gestion du réseau, il est naturel qu'il fasse l'objet d'une attention particulière. Une analyse des risques pourrait être réalisée, ou imposée par la règlementation, à la charge de l'opérateur de réseau. À première vue les risques sont nombreux et de différentes natures : risque pour la sécurité physique des personnes, risque financier pour les opérateurs se voyant privés d'un bien essentiel à leur activité, risque pour la vie privée des individus si le système doit collecter et traiter des données fines de consommation en temps réel, entre autres. Une cartographie complète et documentée des risques pourrait être exigée de l'opérateur de

[939] En cas de demande de consommation trop importante par rapport à la production disponible d'électricité, des règles guident le choix des opérateurs de réseau sur les clients ou zones géographiques à déconnecter en priorité.
[940] Sur le principe général, voir Code de l'énergie, article L121-1 ; sur la responsabilité incombant aux opérateurs des réseaux de transport et de distribution voir Code de l'énergie, articles L322-9 (distribution) et L321-10 (transport) : *« Le gestionnaire du réseau public de transport assure à tout instant l'équilibre des flux d'électricité sur le réseau ainsi que la sécurité, la sûreté et l'efficacité de ce réseau, en tenant compte des contraintes techniques pesant sur celui-ci »*.

réseau. Il y a de fortes chances que l'analyse d'impact ainsi réalisée mette en évidence des « risques importants pour les droits fondamentaux des individus »[941]. Dans ce cas, il serait pertinent d'en imposer la publication aux fins de consultation du public sur le site gouvernemental dédié, avant de soumettre le projet à l'autorité indépendante en charge de la mise en œuvre de la régulation. L'autorité devra alors prendre en considération à la fois les éléments techniques décrivant le système d'IA, l'analyse d'impact réalisée par son fournisseur et les conclusions de la consultation du public. Des lignes directrices pourraient être rédigées par l'autorité pour guider les fournisseurs dans la conduite de la consultation du public, ou la procédure pourrait être supervisée par un service dédié en son sein. Un tel mécanisme nécessiterait de rendre certains processus obligatoires (notamment l'étude d'impact obligatoire, la publicité et l'autorisation par une autorité indépendante en cas de « risque important pour les droits fondamentaux »), mais permettrait de garantir l'effectivité du droit à la participation du public dans la régulation de l'IA. Si le système venait à être autorisé, le public devrait toujours disposer d'un moyen de s'exprimer, que ce soit sur l'apparition éventuelle de nouveaux risques (par exemple liés à une évolution des techniques utilisées par des hackeurs, remettant en cause la sécurité initiale du système) ou sur la pertinence de la régulation. Concrètement, le même site gouvernemental d'information du public pourrait être utilisé pour sa consultation continue sur l'utilisation des systèmes d'IA à risque.

[941] Voir Supra, 387-390.

§2 : Le droit à l'information du public

453. **L'objectif du droit à l'information du public sur l'utilisation de systèmes d'IA.** La doctrine et les régulateurs européens et internationaux ont déjà mis en évidence l'importance de l'information des individus. On retrouve cette idée tant dans les premières réflexions du Parlement européen en 2017[942], que dans des articles académiques[943] et ouvrages de référence[944], ou encore dans la première mouture de règlement européen sur l'IA[945]. Pourtant, les propositions portaient principalement sur l'information des individus de la nature artificielle du système avec lequel ils échangent : assistants virtuels, réponse automatique aux mails,[946]… De plus, cette information des individus est souvent posée comme une obligation pour les fournisseurs des systèmes d'IA, et non comme un droit pour les individus. Ce choix est regrettable puisqu'il exclut l'individu et le prive de tout moyen de demander ou réclamer l'information si elle est inexistante ou incomplète. Dans ce schéma, les individus n'ont d'autres choix que de faire confiance au fournisseur, en l'absence de moyens de recours ou de droits spécifiques qui leur seraient reconnus. Cette situation est, selon nous, inacceptable et inadaptée pour garantir une protection efficace des droits fondamentaux des individus. Reconnaître un droit à l'information des individus relatif aux systèmes d'IA qu'ils utilisent ou dont le fonctionnement peut avoir des conséquences sur eux est indispensable afin de leur permettre d'exercer leurs droits existants et de réaliser des choix éclairés. Ce droit doit venir complémenter l'obligation de transparence pesant sur les fournisseurs d'IA. Ce dernier a vocation à informer les individus de l'existence d'un système d'IA, de la façon (globale) dont il fonctionne et des risques associés, tandis que le droit à l'information des individus vise à leur donner à la fois un moyen de contrôle pour s'assurer que l'information qui leur est donnée est

---

[942] *Résolution 2015/2103(INL) du Parlement européen du 16 février 2017 contenant des recommandations à la Commission concernant des règles de droit civil sur la robotique*, publiée au JOUE n°C252/239 le 18 juillet 2018.
[943] C. Castets-Renard, « Comment construire une intelligence artificielle responsable et inclusive ? », *Recueil Dalloz*, 2020, p. 225.
[944] Y. Meneceur, *L'intelligence artificielle en procès : plaidoyer pour une règlementation internationale et européenne*, Bruylant, coll. Macro Droit – Micro Droit, 2020, spec. p. 391.
[945] COMMISSION EUROPÉENNE, *Proposition de règlement du Parlement européen et du Conseil établissant des règles harmonisées concernant l'intelligence artificielle (législation sur l'intelligence artificielle) et modifiant certains actes législatifs de l'Union*, 21 avril 2021, 2021/0106 (COD), articles 1 (principe général de transparence), 13 (transparence de l'information auprès des utilisateurs des systems d'IA) et 52 (transparence renforcée pour certains systèmes d'IA tels que les chatbots).
[946] *Ibid.*, article 52.

fiable et complète, et un moyen d'obtenir les données nécessaires à l'exercice de leurs autres droits. Ces « autres droits » recouvrent par exemple les droits des personnes contenus dans le RGPD, le droit d'ester en justice pour obtenir réparation des préjudices causés sur le fondement des régimes de responsabilité civile ou encore l'ensemble des droits fondamentaux contenus, par exemple, dans la Convention européenne des droits de l'Homme. Créer un droit à l'information pour les individus revient finalement à contraindre les fournisseurs d'IA à « donner les armes » juridiques aux citoyens pour contrôler la technologie. Toutefois, le contenu de l'information que l'individu pourrait réclamer du fournisseur au titre de ce nouveau droit doit être bien délimité.

454. **Le contenu du droit à l'information du public sur l'utilisation de systèmes d'IA.** Il doit être dûment proportionné à ce qui est nécessaire pour les individus afin d'exercer leurs droits et à ce qui est diffusable pour les fournisseurs sans atteindre à des droits de propriété intellectuels ou secrets d'affaires. En particulier, l'information ne doit pas consister en la divulgation du savoir-faire de l'entreprise, ni au détail de la conception et du fonctionnement du système d'IA, qui ne serait pas intelligible pour tout un chacun. Encore une fois, le niveau de technicité de l'information à donner doit être proportionné au bagage technique du public ciblé et de ses propres besoins[947]. La chercheuse Nathalie Smuha estime à cette fin que, dans le cadre de la régulation de l'IA, les individus devraient disposer d'un droit à l'information à trois composantes, qui nous semblent particulièrement pertinentes[948].

455. **L'information des personnes sur l'existence et le fonctionnement des systèmes d'IA.** Premièrement, les individus devraient être informés sur toutes les façons dont le fonctionnement du système d'IA pourrait les affecter, incluant l'explication de comment le système infère des résultats, à partir de quelles données d'entrée et comment ces résultats peuvent avoir des conséquences pour la personne concernée.

---

[947] V. Beaudouin, I. Bloch, D. Bounie, *et al.*, « Identifying the "Right" Level of Explanation in a Given Situation », *Hal archives ouvertes*, 2020, hal-02507316, disponible en ligne : <https://hal.telecom-paristech.fr/hal-02507316>, consulté le 20 avril 2020.

[948] N.A. Smuha, E. Ahmed-Rengers, A. Harkens, *et al.*, « How the EU can achieve legally trustworthy AI: A response to the European Commission's proposal for an Artificial Intelligence Act », *SSRN Electronic Journal*, 5 août 2021, spec. p. 52 et s.

456. **L'information des personnes sur les risques d'atteinte à leurs droits fondamentaux.** Deuxièmement, les personnes devraient pouvoir connaître les potentielles interactions ou interférences entre les effets du fonctionnement du système d'IA et leurs droits fondamentaux (droit à la vie, droit à la vie privée, droit à un environnement sain…), incluant une justification de la balance risque/bénéfice si un risque existait.

457. **L'information des personnes sur les voies de recours possibles.** Troisièmement, il convient d'informer les individus concernés par le fonctionnement d'un système d'IA sur les moyens à leur disposition pour agir en justice, porter plainte ou obtenir réparation lorsqu'ils suspectent que leurs droits fondamentaux ont été bafoués injustement ou de façon disproportionnée ou lorsqu'ils suspectent le fournisseur de ne pas avoir respecté ses obligations règlementaires (obligation de transparence, exigences de conformité, défaut de supervision humain effective…)[949].

458. **Articulation avec le droit à l'information des personnes au titre du RGPD.** Les trois composantes décrites ci-dessus pourraient fonder un droit effectif à l'information des individus relatif aux systèmes d'IA qu'ils utilisent ou dont le fonctionnement peut avoir des conséquences sur eux. L'information des personnes ici proposée doit s'entendre sans préjudice de la transparence obligatoire au titre du RGPD. Cette dernière ne s'applique qu'aux traitements de données à caractère personnel, dont certains peuvent être réalisés au moyen de systèmes d'IA, alors que notre proposition vise à étendre le droit à l'information à l'ensemble des systèmes d'IA, y compris ceux ne traitant pas de données personnelles.

459. **Un exemple prospectif dans le secteur de l'électricité.** Si on prend l'exemple d'un système d'IA autonome pilotant la consommation d'un foyer aux appareils électroménagers connectés, notre proposition imposerait d'abord au fournisseur de fournir une information de premier niveau sur l'existence d'un système d'IA, l'explication générale de son fonctionnement et les éventuels risques qui auraient été identifiés lors de l'analyse d'impact préalable. Ensuite, si les individus considéraient que l'information était incomplète, ils pourraient exercer leur droit à l'information directement auprès du fournisseur en leur demandant précisément toutes les

[949] *Ibid.*

conséquences que le fonctionnement du système peut avoir pour eux (ce qui, normalement, a déjà été étudié par le fournisseur lors de l'étude d'impact), les interactions identifiées entre ces conséquences et leurs droits fondamentaux et enfin les moyens par lesquels ils peuvent obtenir réparation s'ils estiment que leurs droits ont été bafoués ou signaler une non-conformité suspectée. De façon plus précise, on pourrait également imaginer que les individus puissent demander une information contextualisée sur un événement précis. Dans notre exemple, si l'individu a été victime d'un incendie dans son domicile, il pourrait exercer son droit à l'information auprès du fournisseur du système d'IA pilotant ses équipements connectés afin d'obtenir une explication des décisions prises par le système au moment de l'événement. L'exercice d'un tel droit (et l'obligation d'y répondre pour le fournisseur) permettrait ainsi de lever les problématiques de preuve en matière de responsabilité civile. En effet, les individus bénéficieraient là de moyens pour obtenir des données précieuses afin de prouver l'origine du dommage[950].

460. **Transition.** À ce titre, le droit à l'information des individus tel que nous l'avons présenté serait une condition indispensable à l'exercice du troisième droit individuel que nous proposons : le droit d'accès à la justice.

[950] EXPERT GROUP ON LIABILITY AND NEW TECHNOLOGIES, *Liability for AI and other emerging digital technologies*, 21 novembre 2019, spec. p. 49 et s., disponible en ligne : <https://op.europa.eu/en/publication-detail/-/publication/1c5e30be-1197-11ea-8c1f-01aa75ed71a1/language-en/format-PDF>, consulté le 18 janvier 2020.

§3 : Le droit d'accès à la justice

461. **Le contenu du droit d'accès à la justice dans le Droit de l'environnement.** Le droit d'accès à la justice a initialement été théorisé et concrétisé en Droit de l'environnement. À l'échelle internationale, il figure dans les textes fondateurs que sont la Déclaration de Rio de 1992[951] et la Convention d'Aarhus[952]. Dans ce contexte, il vise à garantir aux individus des moyens de recours effectifs pour faire valoir leurs droits et exercer un contrôle sur les décisions et projets susceptibles d'avoir une incidence sur l'environnement. Le droit d'accès à la justice vise ainsi à garantir les droits à la participation et à l'accès à l'information, à faire respecter les règles de protection de l'environnement et améliorer l'efficacité des lois environnementales, à améliorer la qualité du processus décisionnel en renforçant le contrôle de l'administration publique, ou encore à promouvoir la légitimité dans la prise de décision et ainsi promouvoir la confiance envers le gouvernement[953]. L'article 9 de la Convention d'Aarhus en précise les conditions. Les États signataires doivent notamment veiller :

- A garantir aux individus des moyens effectifs de recours judiciaires si leur demande d'information a été ignorée, injustement rejetée ou insuffisamment prise en compte[954] ;
- A garantir, à toute personne ayant un intérêt à agir ou faisant valoir une atteinte à un droit, des moyens effectifs de recours judiciaires contre toute décision ayant des conséquences environnementales négatives[955] ;

[951] *Déclaration de Rio sur l'environnement et le développement adoptée par la conférence des Nations unies sur l'environnement et le développement*, Rio de Janeiro, 1992, Principe 10. Disponible en ligne : <https://www.un.org/french/events/rio92/rio-fp.htm>, consulté le 30 avril 2021.
[952] Convention d'Aarhus, article 1er.
[953] J. Ebbesson, « L'accès à la justice en matière d'environnement en droit international : pourquoi et comment ? », in *Le droit d'accès à la justice en matière d'environnement*, dir. J. Bétaille, Presses de l'Université Toulouse 1 Capitole, LGDJ – Lextenso Editions, 2016, pp. 63-75 ; V. aussi COMMISSION ÉCONOMIQUE DES NATIONS UNIES POUR L'EUROPE, *La Convention d'Aarhus : guide d'application*, Rapport, 2ème édition, 2014, 280 p., disponible en ligne : <https://unece.org/DAM/env/pp/Publications/Aarhus_Implementation_Guide_FRE_interactive.pdf>, consulté le 11 décembre 2021.
[954] Convention d'Aarhus, article 9, §1.
[955] Convention d'Aarhus, article 9, §2.

- A ce que les voies de recours ainsi créées constituent des recours suffisants et effectifs et soient objectives, équitables et rapides sans que leur coût soit prohibitif[956] ;
- A ce que le public soit informé de la possibilité qui lui est donnée d'engager des procédures de recours administratif ou judiciaire, et à mettre en place des mécanismes appropriés visant à éliminer ou à réduire les obstacles financiers ou autres qui entraveraient l'accès à la justice[957].

462. **Le droit d'accès à la justice en matière environnementale dans les textes.** D'abord décliné dans le droit de l'UE[958], ce droit a ensuite été concrétisé en droit français, dans le Code de l'environnement. Bien que son intitulé « droit d'accès à la justice » n'y soit pas mentionné, plusieurs dispositions viennent créer ces fameuses voies de recours, garantir leur effectivité et lutter contre les barrières, notamment économiques, à leur exercice. C'est le cas par exemple de l'article L142-1 du Code de l'environnement assouplissant le critère de l'intérêt à agir pour les associations de protection de l'environnement. Les voies « d'accès à la justice » créées consistent principalement en des procédures de référé administratif, permettant la suspension de projets, plans ou programmes qui n'auraient pas suivis les procédures imposées, le cas-échéant d'évaluation environnementale[959], d'enquête publique[960] ou d'étude d'impact[961]. Enfin, le droit d'accès à l'information fait lui aussi l'objet d'une voie de recours dédiée en vue de garantir son effectivité, auprès de la Commission d'Accès aux Documents Administratifs

[956] Convention d'Aarhus, article 9, §4.
[957] Convention d'Aarhus, article 9, §5.
[958] V. notamm. COMMISSION EUROPÉENNE, *Communication de la Commission du 28 avril 2017 sur l'accès à la justice en matière d'environnement*, C(2017) 2616 final ou CONSEIL DE L'UNION EUROPEENNE, Décision du Conseil du 17 février 2005 relative à la conclusion, au nom de la Communauté européenne, de la convention sur l'accès à l'information, la participation du public au processus décisionnel et l'accès à la justice en matière d'environnement, 2005/370/CE, publiée au JOUE L124/1 du 17 mai 2005.
[959] Code de l'environnement, article L122-11.
[960] Code de l'environnement, article L123-16, permettant d'obtenir la suspension d'une décision autorisant la réalisation d'aménagements, d'ouvrages ou de travaux soumise à une enquête publique préalable, si les moyens invoqués présentent un doute sérieux quant à la légalité de la décision, lorsque les conclusions du commissaire enquêteur sont défavorables, ou lorsque l'enquête publique requise n'a pas eu lieu.
[961] Code de l'environnement, article L122-2, permettant d'obtenir la suspension d'une autorisation ou d'une décision d'approbation relative à la réalisation d'aménagements ou d'ouvrages, qui par l'importance de leur dimension ou leur incidence sur le milieu naturel, peuvent porter atteinte à ce dernier, dès lors que le projet est soumis à étude d'impact et que cette étude n'a pas été réalisée.

(CADA)[962] et selon la procédure décrite à l'article L311-2 du Code des Relations entre le Public et l'Administration (CRPA).

463. **Conséquences de la reconnaissance du droit d'accès à la justice.** Les voies de recours ainsi créées visent essentiellement à permettre aux individus d'accéder à l'information à laquelle ils ont légalement droit, de contester toute décision environnementale affectant leurs droits et d'obtenir la réparation ou la cessation du préjudice causé. Les contours du droit d'accès à la justice environnementale ainsi formulés nous semblent être une source d'inspiration pertinente dans le cadre de notre réflexion sur la régulation de l'IA.

464. **L'intégration d'un droit d'accès à la justice dans la régulation de l'IA.** Transposé à la régulation de l'IA, le droit d'accès à la justice devrait viser à garantir les droits des individus relatifs aux systèmes d'IA qu'ils utilisent ou dont le fonctionnement peut avoir des conséquences sur eux, ainsi que tous leurs autres droits existants, en particulier celui d'obtenir réparation en cas de préjudice causé par un système d'IA. Ce dernier point est en principe déjà couvert par les régimes de responsabilité civile qui, en droit français, sont en grande partie adaptés aux systèmes d'IA, à condition d'en encadrer la conception[963]. En revanche, il est primordial que le cadre juridique de l'IA prévoie des voies de recours pour garantir aux individus l'effectivité de leurs droits.

465. **Le droit de porter plainte.** En premier lieu, les individus devraient avoir la possibilité de porter plainte s'ils estimaient qu'un système d'IA portait atteinte à leurs droits fondamentaux ou de signaler une non-conformité suspectée auprès de l'autorité nationale de supervision. Un tel mécanisme permettrait d'impliquer les personnes sujettes à un traitement d'IA dans la mise en œuvre de la régulation, de leur donner de la visibilité sur la procédure à suivre lorsqu'ils soupçonnent une atteinte à leurs droits fondamentaux mais aussi de faciliter le travail de l'autorité de supervision dans l'identification des systèmes d'IA non conformes[964]. Il pourrait

[962] La CADA est une autorité administrative indépendante créée par la loi n° 78-753 du 17 juillet 1978. Sa mission est de veiller au respect de la liberté d'accès aux documents administratifs et elle est également compétente pour l'accès à l'information relative à l'environnement.
[963] Voir Supra, 100.
[964] N.A. Smuha, E. Ahmed-Rengers, A. Harkens, *et al.*, « How the EU can achieve legally trustworthy AI: A response to the European Commission's proposal for an Artificial Intelligence Act », *SSRN Electronic Journal*, 5 août 2021, spec. p. 77.

être calqué sur le modèle de l'article 77 du RGPD : « *toute personne concernée a le droit d'introduire une réclamation auprès d'une autorité de contrôle, en particulier dans l'État membre dans lequel se trouve sa résidence habituelle, son lieu de travail ou le lieu où la violation aurait été commise, si elle considère que le traitement de données à caractère personnel la concernant constitue une violation du présent règlement* »[965]. Toutefois, un tel droit de plainte ne serait effectif que si l'autorité de contrôle en charge de la mise en œuvre de la régulation de l'IA est dotée des moyens techniques, humains et financiers, ainsi que des pouvoirs d'enquête, de contrôle et de sanction pour instruire les plaintes.

466. **Le droit à la réparation.** En deuxième lieu, les individus devraient se voir reconnaître un droit à un recours effectif en vue d'obtenir la réparation du dommage causé par un système d'IA ou l'exécution forcée des obligations de la règlementation. Concernant les dommages réparables, il est important d'adopter une logique de subsidiarité. Les États membres dont les régimes de responsabilité peuvent déjà s'appliquer aux dommages causés par des systèmes d'IA ne devraient pas avoir à modifier leur législation. À l'inverse, ceux dont le droit national n'offre pas un recours effectif en cas de dommage causé par un système d'IA devraient le modifier en conséquence. Cette voie permet d'éviter d'imposer, au nom de l'harmonisation, la modification de droits nationaux dont les principes sont pleinement applicables aux systèmes d'IA. À ce titre, l'article 82 du RGPD intitulé « Droit à réparation et responsabilité » peut être pris en exemple : « *Toute personne ayant subi un dommage matériel ou moral du fait d'une violation du présent règlement a le droit d'obtenir du responsable du traitement ou du sous-traitant réparation du préjudice subi. […] Les actions judiciaires engagées pour exercer le droit à obtenir réparation sont intentées devant les juridictions compétentes en vertu du droit de l'État membre visé à l'article 79, paragraphe 2* »[966]. Un article rédigé de la sorte dans un règlement européen permet de faire de la violation du règlement un fondement à une éventuelle action en responsabilité civile extracontractuelle. Par ailleurs, le droit à un recours juridictionnel effectif de l'article 79 du RGPD donne également une idée de ce à quoi il pourrait ressembler dans le cadre de la régulation de l'IA : « *Sans préjudice de tout recours administratif ou extrajudiciaire qui lui est ouvert, y compris le droit d'introduire une réclamation auprès d'une autorité de contrôle […],*

[965] RGPD, article 77.
[966] RGPD, article 82.

chaque personne concernée a droit à un recours juridictionnel effectif si elle considère que les droits que lui confère le présent règlement ont été violés du fait d'un traitement de ses données à caractère personnel effectué en violation du présent règlement »[967]. Cette formulation gagnerait cependant à être précisée quant à la qualification du « recours juridictionnel effectif ». Selon nous, l'effectivité doit s'apprécier selon la capacité du recours à aboutir soit à la réparation du préjudice subi comme c'est l'objet de l'article 82 du RGPD, soit à l'exécution forcée des obligations du fournisseur d'IA au titre de la règlementation[968]. Si une telle articulation était mise en place, l'initiative européenne pour harmoniser les régimes de responsabilité en matière d'IA[969] pourrait être abandonnée, puisque seuls les États membres devraient s'assurer de l'effectivité des voies de recours offertes par leur droit national. Une nouvelle couche règlementaire serait ainsi évitée au niveau européen.

467. **Le droit de contester les décisions de l'IA : des risques déjà couverts en Europe par le RGPD.** En troisième lieu, la doctrine américaine a proposé la création d'un droit de contester les décisions prises par un système d'IA[970]. Partant du principe que l'IA était utilisée pour prendre des décisions importantes telles que l'orientation des bacheliers[971], l'octroi de crédit financier[972] ou la distribution de vaccin de Covid-19[973], les chercheuses Margot Kaminski et Jennifer Urban soutiennent que les personnes devraient avoir la possibilité de remettre en cause

[967] RGPD, article 79.
[968] N.A. Smuha, E. Ahmed-Rengers, A. Harkens, *et al.*, « How the EU can achieve legally trustworthy AI: A response to the European Commission's proposal for an Artificial Intelligence Act », *SSRN Electronic Journal*, 5 août 2021, spec. p. 76.
[969] Voir la consultation lancée par la Commission européenne sur le sujet : COMMISSION EUROPÉENNE, « Commission collects views on making liability rules fit for the digital age, Artificial Intelligence and circular economy », *Site officiel de la Commission européenne*, 19 octobre 2021, disponible en ligne : <https://digital-strategy.ec.europa.eu/en/news/commission-collects-views-making-liability-rules-fit-digital-age-artificial-intelligence-and>, consulté le 13 février 2022 ; COMMISSION EUROPÉENNE, *Rapport sur les conséquences de l'intelligence artificielle, de l'internet des objets et de la robotique sur la sécurité et la responsabilité*, Rapport de la Commission au Parlement européen, au Conseil et au Comité économique et social européen, 19 février 2020, COM(2020) 64 final.
[970] M.E. Kaminski, J.M. Urban, « The Right to Contest AI », *Columbia Law Review*, 16 novembre 2021, vol. 121, n°7.
[971] A. Boudinar-Zabaleta, « Algorithmes et lignes directrices. Réflexions sur la codification automatisée des motifs des décisions administratives », *Droit Administratif*, avril 2019, n° 4, étude 7.
[972] S. Lacroix De Sousa, « L'octroi de crédit à l'épreuve des droits fondamentaux », Revue de Droit bancaire et financier, novembre 2018, n° 6, dossier 41 ; T. de Ravel d'Esclapon, « La gouvernance des algorithmes dans le secteur financier : le point de vue de l'ACPR », *Revue de Droit bancaire et financier*, juillet 2020, n° 4, étude 12.
[973] M.E. Kaminski, J.M. Urban, *op. cit.*, spec. p. 1.

ces décisions individuelles en application du principe du « *due process* »[974]. Leur proposition s'inspire de l'article 22 du RGPD encadrant les décisions individuelles automatisées et garantissant aux personnes concernées le droit « d'obtenir une intervention humaine […], d'exprimer son point de vue et de contester la décision »[975]. Ce droit était déjà bien ancré en droit français, en particulier dans le régime des décisions administratives individuelles[976], et s'exporte désormais outre-Atlantique[977]. Il ne nous apparaît pas nécessaire de reconnaître un nouveau droit à la contestation des décisions prises sur le fondement d'un système d'IA, puisque les risques y relatifs sont déjà couverts par la législation existante, en particulier le RGPD.

468. **Conclusion de la Section 3 sur la création de droits individuels dans la régulation de l'IA.** En somme, pour garantir une mise en œuvre efficace de la régulation de l'IA et une protection optimale des droits fondamentaux des individus, plusieurs nouveaux droits individuels pourraient être expressément reconnus par les textes, sur le modèle du Droit de l'environnement. D'abord, un droit à la participation permettrait d'impliquer les individus dans la régulation de la technologie. Ensuite, un droit à l'information permettrait au public de disposer des informations nécessaires à la protection de leurs droit fondamentaux, notamment. Enfin, ces droits ne seraient effectifs que si la régulation de l'IA garantissait aux personnes un véritable droit d'accès à la justice, consistant en un droit de plainte et un droit à un recours effectif.

469. **Conclusion du Chapitre 2 relatif aux propositions de mécanismes de régulation de l'IA**. La réflexion théorique sur les modes de régulation et l'étude des expériences de la

---

[974] D.K. Citron, « Technological Due Process », *Wahsington University Law Review*, 2008, vol. 85, issue 6, 1249.
[975] RGPD, article 22 (3).
[976] E. Untermaier-Kerléo, « Les nouveaux visages de la décision administrative : d'une administration assistée à une administration automatisée », *La Semaine Juridique Administrations et Collectivités territoriales*, 17 décembre 2018, n° 50, 2339 ; T. Dautieu, E. Gabrié, « Analyse de l'apport de la loi pour une République numérique à la protection des données à caractère personnel (2e partie) : Les droits des personnes et les missions et pouvoirs de la CNIL », *Communication Commerce électronique*, janvier 2017, n° 1, étude 1.
[977] GOVERNMENT OF CANADA, *Directive on Automated Decision-Making*, 1er avril 2019, disponible en ligne : <https://www.tbs-sct.gc.ca/pol/doc-eng.aspx?id=32592>, consulté le 16 octobre 2019.

régulation dans d'autres domaines ont permis de mettre en évidence plusieurs mécanismes qui seraient pertinents au regard des objectifs et contraintes de la régulation de l'IA.

D'abord, l'approche de co-régulation, consistant en la délégation d'une partie de la régulation directement aux acteurs régulés, permettrait à la fois de garantir un niveau plancher de sécurité pour les systèmes d'IA tout en responsabilisant les entreprises quant aux moyens pour l'atteindre. La responsabilisation des acteurs régulés peut elle-même passer par plusieurs mécanismes tels que la création d'exigences de conformité, différenciées suivant le niveau de risque du système d'IA concerné pour alléger la contrainte sur applications peu risquées, d'obligations de transparence ou encore par l'implication des acteurs régulés dans le processus de création de la norme.

Ensuite, comme toute action de régulation, la création d'une autorité de contrôle, accompagnée de la menace de sanctions financières importantes, paraît essentielle. Sur un sujet aussi technique que l'IA, il conviendra de veiller à ce qu'elle dispose des moyens techniques, humains et financiers de remplir ses missions de contrôle, d'accompagnement à la conformité et d'information du public. À cet égard, se reposer sur le réseau des autorités existantes en matière de sécurité de systèmes d'information ou de protection des données ne peut pas être une mauvaise idée.

Enfin, si un des objectifs de la régulation est la prévention des atteintes aux droits et libertés des individus, en exclure ces derniers serait une erreur. Au contraire, leur reconnaitre des droits de plainte, de signalement, d'information ou de participation à la décision permettrait de les impliquer dans la régulation, de les sensibiliser aux risques de l'IA et d'accroître la pression sur les entreprises peu vertueuses.

470. **Conclusion du Titre 1 relatif à la proposition d'une régulation proportionnée de l'IA par le Droit.** La création d'un cadre juridique pour l'IA apparaît nécessaire pour répondre aux risques spécifiques posés par son développement et contribuer à la régulation économique du marché de l'IA. Toutefois, afin de préserver la capacité d'innovation des entreprises, il convient de veiller à adopter une approche équilibrée et à ne pas entraver le développement d'applications vertueuses, notamment dans le secteur de l'électricité. Les entreprises accueillent cette perspective plutôt positivement puisqu'elles reconnaissent aisément les risques que peuvent générer les systèmes qu'elles développent, en particulier quand ce n'est pas leur cœur de métier comme c'est le cas dans le secteur étudié. La construction théorique d'un tel cadre

est autant stimulante que difficile tant le spectre des approches de régulation est large et varié. L'expérience de la régulation dans d'autres domaines tels que la finance ou l'environnement nous permet néanmoins de tirer plusieurs leçons à cet égard. La réflexion sur les moyens à mettre en œuvre pour parvenir à un cadre proportionné a abouti dans les développements précédents à la formulation de nombreuses propositions de mécanismes qui apparaissent pertinents au regard de l'objectif poursuivi. L'IA est un domaine éminemment technique et il convient de laisser suffisamment de marge de manœuvre aux entreprises pour ne pas priver la société de tous les bénéfices pouvant être apportés par l'IA. Dans le même temps, la technologie présentant des risques, il convient de protéger les individus contre de potentielles dérives. La responsabilisation des acteurs, la création d'un cadre de régulation contrôle par une autorité indépendante et la reconnaissance de droits aux individus au regard des systèmes d'IA dont le fonctionnement affecte leur quotidien ou leurs libertés sont autant d'instruments juridiques pouvant contribuer, selon nous à l'établissement d'une régulation proportionnée du développement de l'IA.

Toutefois, une fois dépassé le stade de la réflexion théorique, il convient de réfléchir à la façon de mettre en œuvre ces propositions dans la pratique.

Titre 2 : La mise en œuvre d'une régulation *sui generis* adaptée à l'IA

472. **Une réflexion nécessaire sur la mise en œuvre du modèle de régulation proposé.** Les risques nouveaux générés par les caractéristiques spécifiques des systèmes d'IA (complexité, opacité, autonomie et dépendance à la donnée) rendent nécessaire la création d'un cadre juridique spécifique. Ce dernier doit être dûment proportionné afin de ne pas entraver l'innovation et le développement d'applications vertueuses comme il en existe dans le secteur de l'électricité. Pour ce faire, le Titre précédent a mis en évidence plusieurs grands principes et mécanismes juridiques pouvant être mobilisés pour concilier prévention des risques et promotion de l'innovation, tels que la co-régulation ou l'approche par les risques. Toutefois, outre la question du contenu, la concrétisation de ce nouveau cadre de régulation pose de nombreuses questions : A quelle échelle doit-il être créé ? Qui doit en être à l'initiative ? Les projets actuels de règlementation de l'IA sont-ils trop ou pas assez contraignants ? Est-il juridiquement possible de réguler une technologie tout en encourageant son développement ?

473. **Plan.** Notre réflexion théorique sur les moyens de la régulation proportionnée de l'IA doit maintenant s'ancrer dans une réalité pratique. En effet, les projets de régulation de l'IA se multiplient à l'échelle internationale, si bien que le nouveau cadre juridique pour lequel nous plaidons se concrétise déjà. On le constate aux États-Unis[978] et en Chine[979] ou, plus proche de nous, au Royaume-Uni[980]. C'est dans ce contexte que l'Union européenne a choisi d'agir rapidement en initiant, en 2018, un long processus ayant abouti le 21 avril 2021 à une première proposition de texte par la Commission européenne connu sous le nom d'*Artificial Intelligence*

[978] Voir notamm. l'*Algorithmic Accountability Act* proposé aux États-Unis dès 2018 et débattu dans sa nouvelle version en 2022 : *Algorithmic Accountability Act of 2022*, 117th Congress, H.R.6580 ; M. MacCarthy, « An Examination of the Algorithmic Accountability Act of 2019 », *The Transatlantic Working Group Papers Series*, 24 octobre 2019.
Voir aussi : H. Pack, « Regulation of Artificial Intelligence in the United States », in *Interactive Robotics: Legal, Ethical, Social and Economic Aspects: Selected contributions to the INBOTS conference 2021*, dir. M.A. Grau Ruiz, Springer, coll. Biosystems & Biorobotics, 2022, vol 30, pp. 233–237.
[979] J. Conrad, W. Knight, « China Is About to Regulate AI—and the World Is Watching », *Forbes (blog)*, 22 février 2022, disponible en ligne : <https://www.wired.com/story/china-regulate-ai-world-watching/>, consulté le 18 juillet 2022 ; H. Roberts, J. Cowls, J. Morley, *et al.*, « The Chinese approach to artificial intelligence: an analysis of policy, ethics, and regulation », *AI & Society*, 2021, vol. 36, 59–77.
[980] UK OFFICE FOR ARTIFICIAL INTELLIGENCE, « New UK initiative to shape global standards for Artificial Intelligence », *Communiqué de presse*, 12 janvier 2022, disponible en ligne : <https://www.gov.uk/government/news/new-uk-initiative-to-shape-global-standards-for-artificial-intelligence>, consulté le 18 juillet 2022.

Act (*AI Act*)[981]. L'objectif souhaitable de la régulation de l'IA étant l'harmonisation des pratiques de développement et des standards de protection des droits et libertés des individus, il est évident que les normes devraient être partagées par le plus grand nombre d'acteurs. Les projets internationaux portés par l'OCDE[982] ou l'UNESCO[983] sont également bien avancés mais ils visent surtout à proposer un cadre de réflexion, à défaut de pouvoir créer des obligations contraignantes. Par conséquent, le niveau européen semble être le plus adéquat pour concrétiser nos propositions. En effet, la Commission européenne souhaite créer, à travers sa proposition de règlement, un cadre juridique contraignant, transversal (tout secteur confondu) et spécifique à l'IA. Cette proposition est bienvenue et contient plusieurs mécanismes proches des propositions présentées précédemment, telles qu'une approche différenciée selon le risque des applications d'IA ou la création d'exigences de conformité. Toutefois, la proposition européenne est particulièrement contraignante, peinant à convaincre qu'elle puisse concilier efficacement prévention des risques et préservation de la capacité d'innovation des entreprises. En l'état, son application au secteur de l'électricité aurait un effet dissuasif et freinerait le développement d'applications vertueuses. L'*AI Act* est un bon point de départ pour concrétiser un cadre de régulation proportionné mais il devrait être amendé afin de bâtir une règlementation réaliste et acceptable (**Chapitre 1**).

A travers cette proposition, la Commission européenne ambitionne d'adresser les risques générés par l'IA et liés à la sécurité, à la cybersécurité, à la discrimination algorithmique et aux atteintes à la vie privée des individus. Un risque est cruellement absent de cette liste : l'empreinte environnementale des systèmes d'IA. Ces derniers sont effectivement énergivores par nature[984] et peuvent également être utilisés à des fins néfastes pour l'environnement[985].

---

[981] COMMISSION EUROPÉENNE, *Proposition de règlement du Parlement européen et du Conseil établissant des règles harmonisées concernant l'intelligence artificielle (législation sur l'intelligence artificielle) et modifiant certains actes législatifs de l'union*, 21 avril 2021, 2021/0106 (COD), ci-après *AI Act*.

[982] OCDE, *Recommandation du Conseil de l'OCDE sur l'Intelligence Artificielle*, 22 mai 2019, OECD/LEGAL/0449.

[983] UNESCO, *Recommandation sur l'éthique de l'intelligence artificielle*, 23 novembre 2021, SHS/BIO/PI/2021/1, disponible en ligne : <https://unesdoc.unesco.org/ark:/48223/pf0000381137_fre>, consulté le 18 juillet 2022.

[984] E. Strubel, « Energy and Policy Considerations for Deep Learning in NLP », *57th Annual meeting of the Association for Computational Linguistics*, 5 juin 2019, p. 1 et p. 4.

[985] Voir par exemple une offre d'AWS visant à utiliser l'IA pour identifier de nouvelles zones géographiques propices à l'extraction d'hydrocarbures : AMAZON WEB SERVICES, « Your Guide to AI and machine learning at re:Invent 2018", *Amazon (blog)*, 27 septembre 2018, disponible en ligne :

Pourtant, ils peuvent également être mis au service des objectifs de développement durable[986], ce qui est d'autant plus vrai dans le secteur de l'électricité où leur utilisation peut servir à l'intégration des énergies renouvelables dans le réseau électrique ou à limiter la consommation énergétique des bâtiments. La problématique de l'empreinte environnementale de l'IA est peu mature et partagée avec l'ensemble du secteur numérique. En raison de son ambivalence, la réponse juridique à y apporter est complexe puisqu'elle doit une nouvelle fois permettre de trouver un équilibre entre la nécessaire minimisation des conséquences environnementales du développement de l'IA et la promotion des systèmes utiles à la transition écologique. Au vu de sa spécificité, de sa nouveauté et de sa complexité, la question de la concrétisation d'un cadre de régulation environnementale du développement de l'IA sera adressée dans un chapitre dédié (**Chapitre 2**).

<https://aws.amazon.com/fr/blogs/machine-learning/your-guide-to-ai-and-machine-learning-at-reinvent-2018/>, consulté le 8 avril 2021.

[986] R. Vinuesa, H. Azizpour, I. Leite, *et al.*, « The Role of Artificial Intelligence in Achieving the Sustainable Development Goals », *Nature Communications*, décembre 2020, 11, n°1, 233.

Chapitre 1 : Le manque de maturité du projet de règlement européen sur l'IA

474. **La régulation de l'IA : une réponse de l'Union européenne à l'enjeu de la souveraineté numérique.** La volonté de créer un cadre juridique pour le développement de l'IA est relativement récente et s'inscrit dans un contexte géopolitique complexe. Le numérique est un secteur stratégique tant sur le plan économique que politique. Les géants du numérique, américains[987] et chinois[988] notamment, écrasent le marché à l'échelle internationale, si bien que certains parlent même d'une nouvelle forme de colonialisme, numérique[989]. Dans ce contexte de domination américaine et chinoise, l'UE accuse un retard technologique considérable et peine à faire émerger des champions du numérique[990]. Ce retard rend également l'Union dépendante des technologies étrangères, sans pouvoir contrôler efficacement la façon dont les données des citoyens européens sont exploitées. Face à ce constat, certains observateurs se sont essayés à analyser l'incidence du contexte géopolitique sur les stratégies locales de développement du numérique. Le philosophe Gaspard Koenig met par exemple en évidence la divergence des approches américaine, chinoise et européenne en matière d'IA[991]. Selon lui, les États-Unis ont été les premiers à dominer le marché du numérique grâce à leur approche libérale, à l'absence de régulation et à leur culture de l'entrepreneuriat. Dans le modèle américain, la régulation n'intervient qu'en dernier recours, lorsque la domination économique est déjà établie et que la pression internationale en faveur d'une règlementation se fait trop forte. La Chine et la Russie sont parvenues à rattraper une partie de leur retard par une approche différente, marquée par un fort interventionnisme de l'État. L'émergence des géants chinois a également été permise par une vision particulièrement laxiste en matière de protection des données personnelles et par la culture du contrôle de la population qui justifie la collecte et l'exploitation du plus grand nombre de données sur les individus. D'un point de vue purement technologique, l'Union européenne semble avoir déjà perdu la bataille sur le plan international.

[987] Google, Apple, Facebook, Amazon, Microsoft.
[988] Baidu, Alibaba, Tencent, Xiaomi.
[989] M. Kwet, « Digital colonialism: US empire and the new imperialism in the Global South », *Race & Class*, 2019, vol. 60, n°4, 3-26.
[990] E. Rugraff, « La politique industrielle de l'UE face à son décrochage technologique », *Bulletin de l'Observatoire des Politiques Économiques en Europe*, 2019, vol. 41, 33–45.
[991] G. Koenig, *La fin de l'individu : voyage d'un philosophe au pays de l'intelligence artificielle*, L'Observatoire, coll. De Facto, 189 p., spec. chap. 5.

En réponse à ce contexte géopolitique complexe, l'Union a fait le choix de la différenciation par l'éthique. La stratégie européenne vise à bâtir une souveraineté technologique et numérique, fondée sur trois principes : *« une technologie au service des personnes », « une économie juste et compétitive »* et *« une société ouverte démocratique et durable »*[992]. L'Union, prenant acte d'un retard qu'elle semble assumer ne pas pouvoir rattraper[993], choisit de promouvoir un numérique « à l'européenne » fondé sur l'éthique et la confiance, plutôt que de chercher à identifier et lever les obstacles techniques et économiques qui empêchent l'émergence de champions technologiques. Pour remplir cet objectif global de promotion d'un numérique de confiance, l'Union européenne a initié la construction d'un cadre de régulation constitué d'un corpus de textes législatifs. En effet, la Commission européenne a commencé à poser les jalons de ce cadre en novembre 2020 avec sa proposition de texte sur la gouvernance des données[994] visant à encadrer l'activité d'intermédiaire dans les transferts de données, suivie des propositions du *Digital Markets Act*[995] et du *Digital Services Act*[996] en décembre 2020 visant notamment à réguler les grandes plateformes.

475. **L'IA « digne de confiance » à l'européenne.** En matière d'IA, l'Union promeut le concept d'IA « digne de confiance » fondé sur le respect des droits fondamentaux et des valeurs européennes qu'elle compte réaliser à travers la création d'un cadre juridique contraignant. En effet, à la suite d'un long processus de concertation, la Commission européenne a publié une première proposition de règlement européen sur l'IA intitulé *AI Act* le 21 avril 2021[997]. Cette proposition vient donc compléter le corpus de textes législatifs précités et confirme la volonté

[992] COMMISSION EUROPÉENNE, *Shaping Europe's digital future*, Communication de la Commission au Parlement européen, au Conseil, au Comité économique et social européen et au Comité des régions, 19 février 2020, COM (2020) 67 final ; voir sur le sujet : B. Bertrand, « Chronique de Droit européen du numérique : La souveraineté technologique européenne », *RTD eur.*, 2021, 139,
[993] A. Buzelay, « A propos du secteur des hautes technologies en Europe », *Revue de l'Union européenne*, 2022, p. 167.
[994] COMMISSION EUROPÉENNE, *Proposition de règlement sur la gouvernance européenne des données (Data Governance Act)*, 25 novembre 2020, COM (2020) 767.
[995] COMMISSION EUROPÉENNE, *Proposition de règlement relatif aux marchés contestables et équitables dans le secteur numérique (Digital Markets Act)*, 15 décembre 2020, COM (2020) 842 final.
[996] COMMISSION EUROPÉENNE, *Proposition de règlement relatif à un marché intérieur des services numériques (Digital Services Act) et modifiant la directive 2000/31/CE*, 15 décembre 2020, COM (2020) 825 final.
[997] COMMISSION EUROPÉENNE, *Proposition de règlement du Parlement européen et du Conseil établissant des règles harmonisées concernant l'intelligence artificielle (législation sur l'intelligence artificielle) et modifiant certains actes législatifs de l'union*, 21 avril 2021, 2021/0106 (COD).

non dissimulée de l'Union de se placer parmi les précurseurs en matière de régulation du numérique à l'échelle internationale[998].

476. **Une proposition critiquable.** L'initiative de la Commission est louable et nécessaire au regard des risques nouveaux que peut générer l'utilisation de systèmes d'IA et qui ne sont pas couverts par le corpus juridique existant. Il faut saluer la volonté des institutions européennes de créer un cadre règlementaire en faveur d'une IA digne de confiance et proportionné en différenciant le degré de contrainte suivant le niveau de risque de l'application concernée. Néanmoins, si l'intention est louable, les choix effectués par la Commission dans sa première proposition de texte restent discutables à bien des égards. Seuls quelques articles académiques ont été publiés sur le sujet[999], mettant en évidence un certain nombre de lacunes. La majorité salue le choix de la Commission d'adopter une définition large des systèmes d'IA[1000] et s'accorde sur le défaut d'effectivité de certaines dispositions, notamment au regard des exceptions trop nombreuses à l'interdiction des systèmes de reconnaissance faciale[1001] ou du contenu des exigences de conformité imposées pour les systèmes d'IA à haut risque[1002]. Malgré l'objectif affiché de protection des droits fondamentaux, le texte adopte une approche particulièrement bureaucratique avec un examen de conformité *ex ante* et une lourde documentation à mettre en place pour certains systèmes d'IA. À ce titre, la doctrine majoritaire fait état d'un manque de garanties suffisantes en matière de droits fondamentaux, en pointant

---

998 V. notamm. l'exposé des motifs de l'*AI Act*, p. 7 : *« La proposition renforce aussi considérablement la contribution de l'Union à la définition de normes mondiales […]. Elle fournit à l'Union une base solide pour dialoguer davantage avec ses partenaires extérieurs, y compris avec des pays tiers et dans le cadre d'échanges internationaux sur des questions liées à l'IA »*.

999 V. notamm. N.A. Smuha, E. Ahmed-Rengers, A. Harkens, *et al.*, « How the EU can achieve legally trustworthy AI: A response to the European Commission's proposal for an Artificial Intelligence Act », *SSRN Electronic Journal*, 5 août 2021, disponible en ligne : <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3899991>, consulté le 10 août 2021 ; N.A. Smuha, « Beyond the Individual: Governing AI's Societal Harm », *Internet Policy Review*, 30 septembre 2021, vol. 10, n°3 ; M. Veale, F. Zuiderveen Borgesius, « Demystifying the Draft EU Artificial Intelligence Act », *Computational Law Review International*, juillet 2021, vol. 22, n°4 ; M. Ebers, V.R.S. Hoch, F. Rosenkranz, *et al.*, « The European Commission's Proposal for an Artificial Intelligence Act—A Critical Assessment by Members of the Robotics and AI Law Society (RAILS) », *Multidisciplinary Scientific Journal*, 2021, 4, 589-603, 15 ; B. Mathis, « Proposition de règlement européen sur l'intelligence artificielle : le regard d'un praticien », *RLDI*, 2022, n°192, 4179, pp. 40–44.

1000 C. Crichton, « Projet de règlement sur l'IA (I) : des concepts larges retenus par la Commission », *Dalloz Actualité*, 3 mai 2021 ; B. Mathis, « Proposition de règlement européen sur l'intelligence artificielle : le regard d'un praticien », *RLDI*, 2022, n°192, 4179, pp. 40–44.

1001 N.A. Smuha, E. Ahmed-Rengers, A. Harkens, *et al.*, *op. cit.*, pp. 24–28 ; C. Crichton, « Artificial Intelligence Act : avis conjoint des CEPD », *Dalloz Actualité*, 2 juillet 2021.

1002 N.A. Smuha, E. Ahmed-Rengers, A. Harkens, *et al.*, *op. cit.*, p. 33.

l'absence de droits de recours reconnus aux individus[1003]. En outre, certainss commentateurs soulèvent une problématique relative au choix d'une règlementation d'harmonisation totale dans l'Union, qui risquerait de remettre en cause l'existence de certaines législations nationales comportant des dispositions relatives aux algorithmes telles que la Loi pour une République numérique en France[1004]. Nous pourrons regretter que l'absence de prise en compte du risque environnemental lié à l'utilisation de l'IA ne soit mentionnée que par un seul article juridique à l'échelle internationale[1005] et que les études de l'impact du projet de règlement sur des secteurs d'activité précis soient malheureusement trop rares[1006].

477. **Délimitation de l'étude.** Nous n'adresserons pas ici l'ensemble des critiques que l'on pourrait opposer à la proposition de la Commission. Le présent Chapitre n'a pas non plus pour vocation de présenter une liste exhaustive des observations émises par la doctrine. Dans la suite de nos développements, seules seront pointées les dispositions susceptibles d'avoir une incidence à l'égard de l'utilisation de systèmes d'IA dans le secteur de l'électricité. Par conséquent, les critiques relatives aux dispositions concernant les IA présentant un risque inacceptable ne seront pas abordées.

478. **Intérêt pratique de l'étude.** Les propositions développées dans le présent Chapitre sont le fruit de travaux de recherche ayant contribué à la définition de la position d'EDF sur le texte et à la rédaction de propositions d'amendements qui ont été présentés à différentes parties prenantes dans le cadre de rendez-vous institutionnels. Ces amendements seront présentés et justifiés tout au long de nos développements. Ils ont été présentés respectivement aux représentants de la DG Connect le 14 avril 2021, à la Représentation Permanente de la France auprès de l'UE (RPUE) le 26 octobre 2021, ainsi qu'à six députés européens influents en leur qualité de rapporteurs au sein des commissions du Parlement européen compétentes pour amender le projet d'AI Act entre novembre 2021 et mars 2022. Les propositions de thèse contenues dans ce Chapitre ont ainsi été présentées à Dragos Tudorache (groupe Renew), Axel

[1003] N.A. Smuha, E. Ahmed-Rengers, A. Harkens, *et al.*, *op. cit.*, pp. 44–54 ; M. Veale, F. Zuiderveen Borgesius, *op. cit.*, 98–100 ; M. Ebers, V.R.S. Hoch, F. Rosenkranz, *et al.*, *op. cit.*, 600.
[1004] M. Veale, F. Zuiderveen Borgesius, *op. cit.*, 90.
[1005] M. Ebers, V.R.S. Hoch, F. Rosenkranz, *et al.*, *op. cit.*, 601.
[1006] Voir la seule étude sectorielle réalisée pour le secteur financier : N. Martial-Braz, « L'intelligence artificielle bientôt régulée. L'incidence du AI Act dans le secteur financier », *Revue de Droit bancaire et financier*, mai 2021, n°3, comm. 81.

Voss[1007] (groupe PPE) et Petar Vitanov (groupe S&D) pour la commission LIBE ; à Miapetra Kumpala-Natri (groupe S&D) et Eva Maydell (groupe PPE) pour la Commission ITRE ; ainsi qu'à Deirdre Clune (groupe PPE) pour la commission IMCO. Les échanges ont mis en évidence l'intérêt des députés européens pour les exemples concrets des usages de systèmes d'IA dans l'industrie, et en particulier dans le secteur de l'énergie. En effet, le point de vue des industriels utilisateurs de l'IA n'est que très peu présent dans la littérature scientifique.

479. **Plan.** Avant d'étudier les principales lacunes de la proposition de règlement européen sur l'IA, il convient de présenter, à titre liminaire, son contenu (**Section liminaire**). Notre analyse du texte ainsi que les études publiées sur le sujet ont permis de mettre en évidence deux lacunes principales. La première tient à la confusion de sa portée (**Section 1**), la seconde à son contenu qui n'est pas à la hauteur de l'objectif affiché (**Section 2**). Chacune de ces lacunes fera l'objet de propositions d'amendement.

Section liminaire : Présentation du projet de règlement européen sur l'IA

480. **Le cheminement vers la régulation européenne de l'IA.** La proposition de règlement européen sur l'IA est le fruit d'un long processus qui n'est pas encore terminé à l'heure où est rédigée la thèse. Il faut remonter en 2018 pour trouver l'origine de cette initiative dans l'annonce par la Commission de sa stratégie pour l'IA en Europe[1008] ainsi que, quelques mois plus tard, de son plan coordonné dans le domaine de l'IA[1009]. En 2019, le projet de régulation est identifié comme une priorité politique pour le mandat de la nouvelle Commission d'Ursula Van der Leyen[1010]. En parallèle, un groupe d'experts indépendants et de haut niveau en IA est établi pour travailler sur la définition d'une éthique européenne en la matière. Ce travail a conduit à la publication des lignes directrices du 8 avril 2019 pour une IA « digne de

[1007] Axel Voss est également membre de la commission JURI dotée de compétences spéciales sur le texte.
[1008] COMMISSION EUROPÉENNE, *L'intelligence artificielle pour l'Europe*, Communication de la Commission au Parlement européen, au Conseil européen, au Conseil, au Comité économique et sociale européen et au Comité des régions, 25 avril 2018, COM (2018) 237 final.
[1009] COMMISSION EUROPÉENNE, *Un plan coordonné dans le domaine de l'intelligence artificielle*, Communication de la Commission au Parlement européen, au Conseil européen, au Conseil, au Comité économique et social européen et au Comité des régions, 7 décembre 2018, COM (2018) 795 final.
[1010] U. Van der Leyen, *A Union that strives for more. My agenda for Europe*, Political guidelines for the next European Commission 2019-2024, 2019, p. 13.

confiance », définissant le cadre de ce que doit être l'IA « à l'européenne »[1011]. Cette dernière repose sur trois piliers : la robustesse et la sûreté (volet technique), la conformité à la règle de droit (volet juridique) et la conformité aux valeurs européennes (volet éthique). Les lignes directrices se concentrent sur ce dernier aspect en édictant des principes généraux tels que le respect de l'autonomie humaine, la prévention des préjudices ou la transparence, ainsi qu'en proposant une grille d'évaluation[1012]. Ce premier texte, non contraignant et finalement peu concret, n'a eu que peu d'effet sur les entreprises qui se dotaient, en parallèle, de leurs propres chartes éthiques sur l'IA[1013]. En réalité, ce travail préliminaire sur l'éthique européenne de l'IA a servi de base aux réflexions sur la création d'un cadre juridique visant à répondre au volet juridique de l'IA dite « digne de confiance ».

481. **Les appels à la régulation de l'IA par le Parlement européen.** En parallèle, le Parlement européen a également été proactif sur le sujet de l'IA avec notamment l'adoption de trois résolutions le 20 octobre 2020 respectivement sur les aspects éthiques[1014], de responsabilité civile[1015] et de propriété intellectuelle[1016]. La première prône la nécessité d'une règlementation proportionnée et à la hauteur des enjeux éthiques du développement de l'IA, la deuxième plaide pour la création d'un régime de responsabilité civile spécifique et la troisième traite de la problématique de la propriété intellectuelle des créations par des systèmes d'IA. C'est par la première que le Parlement a explicitement appelé la Commission européenne à proposer un premier texte.

[1011] GROUPE D'EXPERTS DE HAUT NIVEAU EN IA, *Lignes directrices du 8 avril 2019 pour une IA digne de confiance*, 8 avril 2019.
[1012] Voir la version mise à jour après la période de consultation : GROUPE D'EXPERTS DE HAUT NIVEAU EN IA, *The assessment list for trustworthy artificial intelligence (ALTAI) for self-assessment*, 14 septembre 2020.
[1013] Voir par exemple les chartes adoptées par Thalès (Thales, *Charte éthique du numérique*, document officiel, septembre 2021, disponible en ligne : <https://www.thalesgroup.com/sites/default/files/2021-10/Charte%20%C3%A9thique%20du%20num%C3%A9rique.pdf>, consulté le 11 octobre 2021) ou IBM (*IBM's Principles for trust and transparency*, 30 mai 2018, disponible en ligne : <https://www.ibm.com/blogs/policy/wp-content/uploads/2018/06/IBM_Principles_SHORT.V4.3.pdf>, consulté le 20 octobre 2019).
[1014] *Résolution* 2020/2012(INL) *du Parlement européen du 20 octobre 2020 portant recommandations à la Commission sur un cadre aux aspects éthiques de l'intelligence artificielle, la robotique et autres technologies*.
[1015] *Résolution 2020/2014(INL) du Parlement européen du 20 octobre 2020 portant recommandations à la Commission sur un régime de responsabilité civile pour l'intelligence artificielle*.
[1016] *Résolution 2020/2015(INL) du Parlement européen du 20 octobre 2020 sur les droits de propriété intellectuelle pour le développement des technologies d'intelligence artificielle*.

482. **Le livre blanc sur l'IA de la Commission européenne.** Quelques mois plus tard, la Commission a précisé son ambition de créer un véritable cadre juridique dédié à l'IA dans un livre blanc publié le 19 février 2020[1017]. Cette publication relève d'une certaine originalité puisque la Commission y présente des options de régulation, allant de l'autorégulation totale à la règlementation la plus stricte en passant par des approches hybrides. Les propositions consistent en la mise en place d'obligations et d'examen de la conformité avant la mise sur le marché (*ex ante*) ou après (*ex post*), ou encore en la promotion de l'adoption de codes de bonne conduite sur la base du volontariat. S'inscrivant dans une démarche innovante de co-construction de la norme, le livre blanc a été publié en vue d'être soumis à consultation pour recueillir l'avis des parties prenantes pendant quatre mois, de février à juin 2020. Les résultats de la consultation ont permis d'orienter le choix de la Commission pour bâtir un cadre juridique pertinent, proportionné et acceptable. Il peut être surprenant de constater que très peu d'entreprises se sont positionnées en opposition au projet de régulation des institutions européennes, sans doute pour des questions d'image. Néanmoins, les grandes entreprises du numérique ont évidemment cherché à limiter le degré de contrainte de la régulation en construction[1018]. Les résultats de la consultation ont conforté l'initiative de la Commission puisque près de 75% des 1200 répondants composés pour moitié d'associations professionnelles, d'entreprises, d'universitaires et d'organisations non gouvernementales considéraient qu'une législation spécifique était nécessaire ou que le corpus existant présentait des lacunes[1019]. L'hybridation entre obligations contraignantes et mesures d'autorégulation est également ressortie comme une priorité pour les parties prenantes[1020]. Il est également intéressant de constater la divergence d'opinions entre les entreprises et associations professionnelles, naturellement réticentes à la contrainte règlementaire, et les citoyens

[1017] COMMISSION EUROPÉENNE, *Livre blanc du 19 février 2020 sur l'Intelligence Artificielle – Une approche européenne axée sur l'excellence et la confiance*, 19 février 2020, COM (2020) 65.
[1018] Voir par exemple la réponse de Google au Livre blanc du 19 février 2020 : GOOGLE, *Consultation on the white paper on AI - a European approach, Google's submission*, 28 mai 2020, disponible en ligne : <https://www.blog.google/documents/77/Googles_submission_to_EC_AI_consultation_1.pdf/>, consulté le 5 juin 2022.
[1019] COMMISSION EUROPÉENNE, *Public consultation on the AI White Paper : Final report*, novembre 2020, p. 8, disponible en ligne : < https://digital-strategy.ec.europa.eu/en/white-paper-artificial-intelligence-public-consultation-towards-european-approach-excellence-and>, consulté le 5 juin 2022.
[1020] *Ibid.*, pp. 8–9 ; p. 13.

répondants à titre individuel, en faveur d'une règlementation exhaustive. A la suite de cette consultation, la Commission a bâti la première mouture de son règlement européen sur l'IA.

483. **Le contenu de la proposition de l'*AI Act*.** La Commission a fait le choix de proposer un règlement européen, d'application directe dans les États membres, en vue de bâtir un cadre juridique harmonisé et d'éviter la fragmentation du marché intérieur. Le texte proposé poursuit quatre objectifs, présentés dans l'exposé des motifs[1021]. Premièrement, il vise à garantir que les systèmes d'IA mis sur le marché de l'UE soient sûrs et respectent les législations en vigueur en matière de droits fondamentaux ainsi que les valeurs de l'UE. Deuxièmement, le cadre proposé est censé apporter de la sécurité juridique pour encourager les investissements et l'innovation dans le domaine de l'IA. Troisièmement, l'*AI Act* doit être un outil pour renforcer l'application effective de la législation existante, notamment au regard de la sécurité des produits. Dernièrement, la Commission souhaite empêcher la fragmentation du marché par une démarche d'harmonisation totale des règles applicables à l'IA.

484. **Contenu de l'*AI Act*.** Le projet de règlement européen sur l'IA est la première proposition de règlementation transversale et contraignante portant sur l'IA à l'échelle internationale. Son premier titre est consacré aux définitions retenues et à son champ d'application[1022]. Le corps du texte comprend une définition très large des systèmes d'IA complétée en annexe I par une liste de techniques englobant les approches symboliques, l'apprentissage automatique et les approches statistiques. Le cadre est applicable à l'ensemble des systèmes d'IA, ce qui englobe un grand nombre d'applications, y compris les systèmes-experts classiques[1023]. Le premier titre définit également les acteurs concernés, à savoir les acteurs principaux de la chaîne de valeur de l'IA : fournisseurs, distributeurs, importateurs, exportateurs, utilisateurs. Toutefois, l'*AI Act* ne prévoit pas les mêmes obligations pour tous les systèmes d'IA, ni pour tous les acteurs définis. Le cadre proposé par la Commission prévoit une approche par les risques, en créant

[1021] *AI Act*, p. 4.
[1022] *AI Act*, Titre 1.
[1023] C. Crichton, « Projet de règlement sur l'IA (I) : des concepts larges retenus par la Commission », *Dalloz Actualité*, 3 mai 2021.

quatre niveaux de risques[1024] : risque inacceptable[1025], haut risque[1026], risque limité[1027], sans risque[1028]. Chaque catégorie est encadrée par des dispositions spécifiques. Les systèmes d'IA considérés comme « inacceptables » sont prohibés sauf exceptions[1029] ; ceux jugés « à haut risque » doivent respecter un corpus d'exigences de conformité[1030] ; ceux présentant un « faible risque » sont soumis à un principe accru de transparence[1031] ; enfin, pour les autres systèmes d'IA, la Commission recommande la création volontaire de codes de bonne conduite[1032]. Les États membres devront désigner une ou plusieurs autorités nationales compétentes et, parmi elles, l'autorité de contrôle nationale chargée de contrôler l'application et la mise en œuvre du règlement. Au niveau européen, l'application des nouvelles règles sera coordonnée par une nouvelle entité, le « Comité européen de l'IA » composé de représentants des États membres et de la Commission[1033]. L'Union européenne cherche ainsi à instaurer un véritable cadre de régulation de l'IA, non sans rappeler le modèle de la protection des données à caractère personnel avec le réseau des autorités nationales de contrôle, supervisées par la Comité européen à la protection des données (CEPD). À cet égard, le non-respect des dispositions prévues par le texte pourrait être sanctionné par des amendes administratives allant jusqu'à 2% ou 6% du chiffre d'affaires annuel mondial suivant l'article concerné[1034].

485. **Les conséquences du règlement IA sur le secteur de l'électricité.** Le contenu de l'*AI Act*, tant par la lourdeur des obligations qu'il impose que par les modalités de sa mise en œuvre et du montant des sanctions encourues, traduit la volonté des institutions européennes de créer un équivalent du RGPD pour réguler les systèmes d'IA. Pour les entreprises, la mise en

[1024] A. Bensamoun, « Artificial Intelligence Act : l'Union européenne invente la pyramide des risques de l'intelligence artificielle », *Le Club des juristes (blog)*, 21 mai 2021, disponible en ligne : < https://blog.leclubdesjuristes.com/artificial-intelligence-act-lunion-europeenne-invente-la-pyramide-des-risques-de-lintelligence-artificielle/>, consulté le 9 juin 2022.
[1025] *AI Act*, Titre 2.
[1026] *AI Act*, Titre 3.
[1027] *AI Act*, Titre 4.
[1028] Non définis par le texte, il s'agit des applications qui entrent dans la définition générale des systèmes d'IA prévue à l'article 3 mais dans aucune des catégories précitées. Ces applications sont seulement concernées par le Titre 9 relatif à la création de codes de bonne conduite pour tous systèmes d'IA.
[1029] *AI Act*, article 5.
[1030] *AI Act*, articles 6 à 51.
[1031] *AI Act*, article 52.
[1032] *AI Act*, article 69.
[1033] *AI Act*, Titres 6, 7 et 8.
[1034] *AI Act*, article 71.

conformité avec les dispositions de ce texte va représenter un coût conséquent, jusqu'à 20% des investissements dans les projets d'IA[1035], et va nécessiter de transformer l'intégralité du processus de conception des systèmes d'IA. Plusieurs cas d'usage dans le secteur de l'électricité pourraient être concernés par les dispositions de l'*AI Act* relatives aux systèmes d'IA à haut risque. En effet, l'annexe III vise explicitement les systèmes utilisés dans la gestion et l'exploitation des infrastructures critiques, et en particulier *« les systèmes d'IA destinés à être utilisés en tant que composants de sécurité […] dans la fourniture d'eau, de gaz, de chauffage et d'électricité »*[1036]. Les autres cas d'usage pourraient rentrer dans la catégorie des IA à faible risque, notamment les assistants virtuels, ou, à défaut, dans les IA sans risque. Les applications étudiées dans le cadre de la présente thèse sont relativement éloignées des applications visées par l'interdiction des systèmes jugés inacceptables. Ainsi, bien que ces dispositions soient hautement critiquables, en particulier au regard des exceptions larges retenues pour l'utilisation de la reconnaissance faciale dans l'espace public, elles ne seront pas étudiées dans ce Chapitre. Il convient de garder à l'esprit que le texte étudié ne constitue que la première proposition de la Commission, laquelle va être amenée à évoluer au cours du processus législatif européen.

486. **L'évolution attendue de la proposition d'*AI Act*.** La première proposition de texte par la Commission marque le point de départ de la procédure législative ordinaire. En effet, l'exposé des motifs précise que les bases juridiques de la proposition sont d'une part l'article 114 du Traité sur le fonctionnement de l'Union européenne (TFUE) qui prévoit l'adoption de mesures destinées à assurer le fonctionnement du marché intérieur et, d'autre part, l'article 16 relatif à l'adoption de règles spéciales en matière de protection des données à caractère personnel. Dans chacun de ces domaines, le TFUE prescrit expressément le recours à la procédure législative ordinaire, prévue à l'article 294. À ce titre, le texte proposé par la Commission doit être débattu et amendé par le Conseil de l'Union européenne et le Parlement européen afin d'être adopté par chacune des institutions en des termes identiques. En cas de désaccord, le projet de texte peut faire l'objet d'une deuxième lecture, d'une procédure de

[1035] CENTER FOR DATA INNOVATION, *How much will the Artificial Intelligence Act cost Europe ?*, Rapport, juillet 2021, disponible en ligne : <https://www2.datainnovation.org/2021-aia-costs.pdf>, consulté le 9 juin 2022.
[1036] *AI Act*, annexe III, 2).

conciliation ainsi que d'une troisième lecture[1037]. En pratique, en cas de désaccord entre le Conseil et le Parlement en première lecture, les trois parties prenantes de la procédure peuvent organiser des réunions interinstitutionnelles informelles, dites « trilogues », afin de parvenir à un accord informel qui doit ensuite être approuvé conformément au règlement intérieur de chacune des institutions[1038]. Ce procédé a été mobilisé dans les dernières procédures législatives en lien avec le numérique[1039], ce qui laisse supposer qu'il pourrait l'être à nouveau dans le cas de l'*AI Act*.

487. **Les débats au Conseil de l'Union européenne.** Dans le cadre de cette procédure, le texte proposé le 21 avril 2021 a fait l'objet de débats au Conseil de l'Union européenne ayant donné lieu à la publication de deux rapports de compromis sous présidences slovène puis française. Le premier rapport, en date du 29 novembre 2021[1040], introduit une exclusion du champ du règlement pour les systèmes d'IA « à usage général », qui sont commercialisés sans être affectés à une finalité précise, amende la définition de l'IA pour reprendre celle établie par l'OCDE[1041] et modifie à la marge certaines obligations pour les systèmes à haut risque. Les deux rapports de compromis sous présidence française, respectivement des 3 et 15 février 2022[1042], apportent des corrections de forme, précisent les délais de conservation des données et proposent de

---

[1037] TFUE, article 294.

[1038] « La procédure législative ordinaire », *Site officiel du Conseil de l'UE et du Conseil européen (blog)*, disponible en ligne : <https://www.consilium.europa.eu/fr/council-eu/decision-making/ordinary-legislative-procedure/>, consulté le 6 juin 2022.

[1039] V. notamm. lors de la procédure législative ordinaire pour le DSA et le DMA (
L. Bertuzzi, « Les institutions européennes donnent le coup d'envoi des négociations sur une loi visant les Big Tech », *Euractiv (blog)*, 11 janvier 2022, disponible en ligne :
<https://www.euractiv.fr/section/economie/news/eu-institutions-kick-off-negotiations-on-law-targeting-big-tech/>, consulté le 6 juin 2022) ou pour la réforme du règlement e-privacy (C. Pehlivan, P. Church, « EU: The ePrivacy Regulation - Let the trilogue begin! », *Linklaters (blog)*, 12 février 2021, disponible en ligne :
<https://www.linklaters.com/fr-fr/insights/blogs/digilinks/2021/february/eu---the-eprivacy-regulation---let-the-trilogue-begin>, consulté le 6 juin 2022).

[1040] CONSEIL DE L'UE, *Presidency compromise text on the proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (AI Act) and amending certain Union legislative acts*, 29 Novembre 2021, n° 2021/0106(COD), n°14278/21, Présidence slovène.

[1041] OCDE, *Recommandation du Conseil de l'OCDE sur l'Intelligence Artificielle*, 22 mai 2019, OECD/LEGAL/0449.

[1042] CONSEIL DE L'UE, *Presidency compromise text (Articles 16-29) on the proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (AI Act) and amending certain Union legislative acts*, n° 2021/0106(COD), 3 février 2022, n°5756/22, Présidence française ; CONSEIL DE L'UE, *Presidency compromise text (Articles 40-52) on the proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (AI Act) and amending certain Union legislative acts*, n° 2021/0106(COD), 15 février 2022, n°6239/22, Présidence française.

modifier les modalités de mise en œuvre de la régulation relatives notamment aux prérogatives des autorités de régulation et aux processus de certification par des tiers. Les discussions au sein du Conseil ont abouti à la publication en novembre 2022 d'un document dit « d'orientation générale » actant la majeure partie des modifications proposées dans les compromis antérieurs[1043].

488. **Les débats au Parlement européen.** Au sein du Parlement européen, le sujet de l'*AI Act* est traité conjointement par la commission des libertés civiles, de la justice et des affaires intérieures (LIBE) et celle du marché intérieur et de la protection des consommateurs (IMCO). Les commissions des affaires juridiques (JURI) et de l'industrie, de la recherche et de l'énergie (ITRE) ont des compétences spéciales sur certaines parties du texte. Ces commissions parlementaires ont également publié leurs premiers rapports[1044]. La mobilisation de l'ensemble du Parlement européen sur le sujet a abouti à l'adoption d'une résolution en date du 3 mai 2022[1045] au contenu symbolique. Les parlementaires y rappellent leurs priorités, à savoir éviter la fragmentation du marché intérieur, préserver la cohérence des définitions utilisées notamment au regard des efforts de normalisation en cours à l'OCDE et éviter la création d'un fardeau administratif disproportionné.

489. **Justification de l'étude de certains amendements proposés.** Il n'est pas certain que l'ensemble des propositions d'amendement contenues dans les différents documents précités,

---

[1043] CONSEIL DE L'UE, *Orientation générale sur la Proposition de règlement du Parlement européen et du Conseil établissant des règles harmonisées concernant l'intelligence artificielle (législation sur l'intelligence artificielle) et modifiant certains actes législatifs de l'Union*, 25 novembre 2022, 2021/0106(COD).

[1044] Pour la commission ITRE : COMMITTEE ON INDUSTRY, RESEARCH AND ENERGY (ITRE), *Draft opinion on the proposal for a regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts (COM(2021)0206 – C9-0146/2021 – 2021/0106(COD))*, 3 mars 2022, 2021/0106(COD).
Pour la commission JURI : COMMITTE ON LEGAL AFFAIRS (JURI), *Draft opinion on the proposal for a regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union Legislative Acts (COM(2021)0206 – C9-0146/2021 – 2021/0106(COD)*, 2 mars 2022, 2021/0106(COD).
Pour les commissions conjointes IMCO et LIBE : COMMITTEE ON THE INTERNAL MARKET AND CONSUMER PROTECTION (IMCO) AND COMMITTEE ON CIVIL LIBERTIES, JUSTICE AND HOME AFFAIRS (LIBE), *Draft report on the proposal for a regulation of the European Parliament and of the Council on harmonised rules on Artificial Intelligence (Artificial Intelligence Act) and amending certain Union Legislative Acts (COM2021/0206 – C9-0146/2021 – 2021/0106(COD)*, 20 avril 2022, 2021/0106(COD).

[1045] *Résolution 2020/2266(INI) du Parlement européen du 3 mai 2022 sur l'intelligence artificielle à l'ère du numérique*.

qu'ils émanent du Conseil de l'UE ou du Parlement européen, figurent effectivement dans la version finale du texte. Toutefois, certaines propositions communes peuvent être identifiées, comme celles relatives aux évolutions de la définition de l'IA ainsi que celles relatives à l'exemption des systèmes d'IA « généraux ». Le consensus autour de ces propositions augmente la probabilité qu'elles figurent dans la version finale du texte, c'est pourquoi elles seront également analysées et critiquées dans nos développements. Les débats promettent d'être encore longs et les divergences d'opinions entre les États membres ou les groupes parlementaires les rendent encore plus difficiles. Selon Dragos Tudorache, rapporteur parlementaire sur l'*AI Act* et président de la commission spéciale sur l'IA, la définition des systèmes d'IA, en particulier ceux considérés « à haut risque », l'usage de la reconnaissance faciale, l'application des exigences relatives à la sécurité des systèmes ou encore les modalités d'encadrement de la production et de la collecte des jeux de données font partie des principaux points de désaccords[1046]. L'absence de consensus peut faire craindre un allongement de la durée des débats avant l'adoption du texte, qui pourraient prendre encore plusieurs mois.

490. **Transition et rappel du plan.** Le texte n'ayant pas encore fait l'objet d'un accord au niveau de Parlement européen, il est encore temps de l'amender pour corriger ses lacunes. En effet, de nombreuses dispositions ne nous semblent pas adaptées pour parvenir à une régulation proportionnée de l'IA, conciliant prévention des risques et promotion de l'innovation. En particulier, sa portée semble encore trop confuse (**Section 1**) et son contenu inadapté à la réalité pratique (**Section 2**).

[1046] « Exclusive Interview with MEP Dragoș Tudorache: Artificial Intelligence Act », *Vote Watch Europe (blog)*, 22 février 2022, disponible en ligne : <https://www.votewatch.eu/blog/exclusive-interview-with-mep-dragos-tudorache-artificial-intelligence-act/>, consulté le 8 juin 2022.

Section 1 : Un texte à la portée confuse

491. **Plan.** Avant même de rentrer dans l'analyse du contenu des dispositions encadrant le développement des systèmes d'IA, la première version de l'*AI Act* présente un certain nombre de lacunes relatives à la fois à son champ d'application (**§1**) et à son articulation avec le corpus législatif existant (**§2**). Elles traduisent une certaine méconnaissance de l'état de l'art dans le domaine de l'IA, due en partie au flou qui entoure cette notion y compris dans la littérature scientifique, et une forme de précipitation dans la règlementation pouvant aboutir à des conflits de règles de droit.

§1 : Une inadéquation du champ d'application

492. **Plan.** L'*AI Act* est applicable à l'ensemble des systèmes d'IA et prévoit des règles différentes selon le niveau de risque de l'application envisagée. Les systèmes d'IA considérés « à haut risque » seraient ainsi soumis à un corpus d'exigences très contraignantes qui nécessiteront une lourde mise en conformité. La démarche est justifiée pour les applications les plus risquées. Les industriels ont par ailleurs engagé de nombreux travaux sur l'encadrement de l'utilisation de l'IA dans les systèmes critiques, nécessitant le développement de bonnes pratiques en matière de sécurité, d'explicabilité voire de certification[1047]. En revanche, au vu du coût prévisible de la mise en conformité avec les obligations de l'*AI Act*, il est très important de bien circonscrire leur champ d'application aux systèmes les plus risqués. À ce titre, les définitions retenues dans le projet de texte ne semblent pas satisfaisantes. D'une part, la définition générale des systèmes d'IA est particulièrement large et risque d'inclure tous les logiciels classiques qui ne présentent pas de risques nouveaux (**A**). D'autre part, les définitions employées pour désigner les systèmes d'IA à haut risque démontrent une forme de méconnaissance des cas d'usage de l'IA dans l'industrie et des risques qu'ils peuvent réellement générer (**B**).

[1047] Voir par exemple la création du premier laboratoire industriel commun en IA par EDF, Thalès et Total : EDF, « Thales et Total forment un trio dans l'intelligence artificielle », *Les Echos*, 6 février 2020, disponible en ligne : <https://www.lesechos.fr/industrie-services/energie-environnement/edf-thales-et-total-forment-un-trio-dans-lintelligence-artificielle-1169776>, consulté le 17 juin 2022.

Par ailleurs, les premiers rapports émanant du Conseil de l'UE ou du Parlement européen font état d'un souhait d'évolution relative au champ d'application du texte plutôt surprenant. En effet, on retrouve dans plusieurs documents des propositions d'amendements visant à exclure du champ de la régulation les « *general purpose AI systems* »[1048] qui correspondent aux systèmes capables de réaliser plusieurs fonctions et pouvant être utilisés pour différentes finalités[1049]. Ce terme sera traduit dans la suite de notre propos par l'expression « systèmes d'IA à usage général » ou « systèmes d'IA généraux »[1050]. Cette exclusion, évidemment défendue par les puissants lobbys des géants du numérique et des entreprises développant des produits commerciaux fondés sur l'IA, enverrait un très mauvais signal à l'égard des industriels souhaitant avoir recours à l'IA dans leur secteur d'activité et qui devrait porter seuls les coûts de la mise en conformité (**C**).

A/ Une définition extensive des systèmes d'IA

493. **Une définition large**. La Commission européenne a fait le choix d'une acception particulièrement extensive dans sa définition des systèmes d'IA afin de respecter le principe de neutralité technologique et de pouvoir être applicable aux futurs développements et avancées techniques dans le domaine de l'IA. L'article 3 de l'*AI Act* définit les systèmes d'IA de la façon suivante :

> *« un logiciel qui est développé au moyen d'une ou plusieurs des techniques et approches énumérées à l'annexe I et qui peut, pour un ensemble donné d'objectifs définis par l'homme, générer des résultats tels que des contenus, des prédictions, des*

[1048] CONSEIL DE L'UE, *Presidency compromise text on the proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (AI Act) and amending certain Union legislative acts*, 29 Novembre 2021, n° 2021/0106(COD), n°14278/21, Présidence slovène, p. 34.
[1049] Voir les propositions d'amendements portées par la Commission des affaires juridiques (JURI) au Parlement européen : COMMITTE ON LEGAL AFFAIRS (JURI), *Draft opinion on the proposal for a regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union Legislative Acts (COM(2021)0206 – C9-0146/2021 – 2021/0106(COD)*, 2 mars 2022, 2021/0106(COD), p. 23.
[1050] Traductions libres de l'expression *« general purpose AI systems »*.

> *recommandations ou des décisions influençant les environnements avec lesquels il interagit »*[1051]

494. **Le regrettable choix d'une définition autonome.** On pourra regretter que la Commission ait choisi d'utiliser une définition différente de celle retenue par l'OCDE, qui n'est pas elle-même exempte de critiques mais qui a le mérite d'être le fruit d'un long travail de consultations et d'un consensus politique à l'échelle internationale. Dans sa recommandation sur l'IA datant de 2020, donc antérieure à la proposition de la Commission, l'OCDE avait en effet retenu la définition suivante :

> *« un système automatisé qui, pour un ensemble donné d'objectifs définis par l'homme, est en mesure d'établir des prévisions, de formuler des recommandations, ou de prendre des décisions influant sur des environnements réels ou virtuels. Les systèmes d'IA sont conçus pour fonctionner à des degrés d'autonomie divers »*[1052]

495. **Comparaison entre les définitions.** Si l'on compare les deux définitions, on observe d'abord que celle retenue dans l'*AI Act* limite l'IA à un « logiciel » sans prendre en compte la nature composite des systèmes d'IA, là où celle de l'OCDE parle bien de « système ». Ensuite, la Commission a fait disparaître toute notion d'autonomie, contrairement à la définition antérieure qui la mentionne à deux reprises. En outre, la définition européenne ajoute les « contenus » aux « prévisions, recommandations et décisions » dans les résultats produits par les systèmes d'IA. En l'état, cet ajout risquerait d'englober toute production de données, de quelque nature que ce soit, par un logiciel. Tout système produisant des résultats se verrait *de facto* couvert par la définition et non pas seulement les systèmes générant des prédictions, des recommandations ou des actions, ce qui est pourtant une caractéristique essentielle des systèmes d'IA. Enfin, malgré l'objectif affiché de neutralité technologique, la Commission fait le choix de renvoyer à une liste d'approches considérées par le texte comme relevant du champ de l'IA. La démarche aurait pu être bienvenue s'il existait un consensus absolu dans la communauté

[1051] *AI Act*, article 3, (1).
[1052] OCDE, *Recommandation du Conseil de l'OCDE sur l'Intelligence Artificielle*, 22 mai 2019, OECD/LEGAL/0449, p. 7.

scientifique sur les techniques qui relèvent, ou non, de l'IA. Malheureusement, à défaut, l'exercice présente un intérêt limité.

496. **Une liste de techniques contestable.** On retrouve cette liste à l'annexe I de l'*AI Act*. Son premier point concerne les *« approches d'apprentissage automatique »* en précisant qu'elles incluent l'apprentissage supervisé, non supervisé, par renforcement, et l'apprentissage profond[1053]. Ces méthodes d'apprentissage sont celles présentant le plus grand degré d'originalité par rapport à la programmation classique donc il n'est pas surprenant de les trouver ici. La Commission vise dans ce premier point les approches dites connexionistes de l'IA, fondées principalement sur l'exploitation de données et pouvant fonctionner de manière plus ou moins autonome et opaque. Ce point est donc parfaitement justifié. Le deuxième élément de la liste inclut les approches dites symboliques de l'IA fondées sur *« la logique et les connaissances, y compris la représentation des connaissances, la programmation inductive (logique), les bases de connaissances, les moteurs d'inférence et de déduction, le raisonnement (symbolique) et les systèmes experts »*[1054]. On peut légitimement s'interroger sur la pertinence de l'accent mis sur la « logique » tant c'est le fondement de toute la programmation informatique[1055]. De plus, la mention des « systèmes experts » est surprenante car ils sont utilisés depuis des décennies et déjà bien appréhendés par les règlementations sectorielles pour leurs applications les plus dangereuses, notamment dans la production d'électricité à travers de lourds processus de certification. À cet égard, la Commission semble ici vouloir englober l'intégralité des systèmes algorithmiques, qu'importent leurs caractéristiques. Le dernier point de l'annexe I vise les *« approches statistiques, estimation bayésienne [et les] méthodes de recherche et d'optimisation »*[1056]. La littérature a déjà pu questionner la pertinence de cette troisième catégorie au motif qu'elle engloberait des techniques comme le calcul de moyenne ou l'informatique décisionnelle basique utilisée sur les sites internet de réservation de voyage par exemple[1057]. L'étude des cas d'usage dans le secteur de l'électricité montre que l'estimation

[1053] *AI Act*, Annexe I, a).
[1054] *AI Act*, Annexe I, b).
[1055] B. Mathis, « Proposition de règlement européen sur l'intelligence artificielle : le regard d'un praticien », *RLDI*, 2022, n°192, pp. 40–44, 4179.
[1056] *AI Act*, Annexe I, c).
[1057] B. Mathis, *op. cit.*

bayésienne peut effectivement être utilisée dans des systèmes de décision ou d'aide à la décision pour sélectionner l'option à recommander parmi une multitude de possibilités[1058]. En revanche, à défaut de précision sur ce que la Commission entend par *« approches statistiques »*, il serait préférable de ne pas les mentionner.

497. **L'absence de justification par la Commission.** On pourrait regretter que la Commission n'ait pas justifié ses choix de techniques figurant à l'annexe I, par exemple dans les considérants, afin que l'on puisse comprendre pourquoi elle a choisi d'inclure tant de techniques dans cette liste, fussent-elles anciennes et bien maîtrisées. En effet, le contenu de l'annexe I illustre bien la volonté de la Commission de retenir l'acception la plus large possible des systèmes d'IA en ne se limitant pas seulement aux algorithmes d'apprentissage automatique. Pourtant, ce sont bien eux qui ont permis un nouvel essor de l'IA depuis les années 90 et posent des problématiques en matière d'explicabilité et d'usage des données. C'est d'autant plus surprenant que le reste du texte est éminemment focalisé sur les techniques d'apprentissage automatique sur des données et non sur les approches « symboliques » fondées sur la programmation de règles[1059]. En effet, le texte s'attache à définir les données d'entraînement[1060], ce qui n'a de sens que lorsque des techniques d'apprentissage automatique ont été employées. De la même manière, les exigences relatives aux systèmes d'IA à haut risque mentionnent l'annotation, l'étiquetage, l'examen des biais[1061], les boucles de rétroaction[1062] ou l'accès aux données d'entraînement[1063]. Si l'on peut comprendre que la définition de l'IA relève également de considérations politiques, économiques et commerciales[1064], il est surprenant de constater que certains auteurs soutiennent l'approche extensive retenue par la Commission[1065],

[1058] Voir le brevet déposé par EDF pour une technologie de maintenance prédictive commercialisée par sa filiale Metroscope : Brevet n° WO2018115646, *Procédé de caractérisation d'une ou plusieurs défaillances d'un système*, déposé le 14 décembre 2017.
[1059] B. Mathis, « Proposition de règlement européen sur l'intelligence artificielle : le regard d'un praticien », *RLDI*, 2022, n°192, pp. 40–44, 4179.
[1060] *AI Act*, article 3, (29).
[1061] *AI Act*, article 10, (2).
[1062] *AI Act*, article 15, (6).
[1063] *AI Act*, article 64.
[1064] N. Boujemaa, « La définition de l'intelligence artificielle, enjeu juridico-commercial », *Le Monde*, 15 décembre 2021.
[1065] Voir notamm. C. Crichton, « Projet de règlement sur l'IA (I) : des concepts larges retenus par la Commission », *Dalloz Actualité*, 3 mai 2021.

voire propose de l'élargir encore[1066]. À l'inverse, d'un point de vue opérationnel et pratique, on regrettera plutôt que la Commission n'ait pas cherché, à travers l'exercice de la définition, à identifier les caractéristiques particulières des systèmes d'IA qui les différencient des logiciels classiques. L'ensemble de ces considérations nous poussent à croire que la définition des systèmes d'IA retenue dans l'*AI Act* mériterait d'être amendée tant au niveau de la définition générale que dans la liste des techniques présente en annexe I.

Il est possible de pallier les lacunes des définitions retenues par la Commission en modifiant plusieurs dispositions.

[1066] N.A. Smuha, E. Ahmed-Rengers, A. Harkens, *et al.*, « How the EU Can Achieve Legally Trustworthy AI: A Response to the European Commission's Proposal for an Artificial Intelligence Act », *SSRN Electronic Journal*, 5 août 2021, p. 55.

498. **Proposition d'amendement de la définition générale des systèmes d'IA (article 3).** D'une part, il convient d'amender la définition générale des systèmes d'IA figurant au premier alinéa de l'article 3. Une première option serait de modifier complètement la définition pour faire figurer les caractéristiques particulières des systèmes d'IA qui justifient la création d'un cadre juridique *ad hoc*. À cette fin, les quatre caractéristiques identifiées en introduction de la thèse peuvent être utilisées : la complexité structurelle, la dépendance à la donnée, la potentielle autonomie et l'opacité du fonctionnement[1067]. Ce faisant, l'article 3 (1) pourrait être amendé ainsi :

Article 3
Définitions

Aux fins du présent règlement, on entend par:

(1)*système d'intelligence artificielle » (système d'IA),* ~~*un logiciel qui est développé au moyen d'une ou plusieurs des techniques et approches énumérées à l'annexe I et qui peut, pour un ensemble donné d'objectifs définis par l'homme, générer des résultats tels que des contenus, des prédictions, des recommandations ou des décisions influençant les environnements avec lesquels il interagit*~~

des systèmes logiciels, pris isolément ou incorporés à un produit, ayant la capacité de générer, à partir de données d'entrée, des prédictions, recommandations ou des décisions susceptibles d'influencer leur environnement. Ils sont conçus à partir de règles fixées par l'humain ou d'un processus d'apprentissage automatique sur des données. Selon les techniques employées, ces systèmes peuvent fonctionner à différents niveaux d'autonomie et de façon plus ou moins opaque.

[…]

499. **Solution alternative.** La définition de l'IA faisant l'objet de nombreux débats, il est probable qu'elle ne puisse pas évoluer autant. Une deuxième option, par défaut, serait

[1067] Voir Supra, 11-16.

d'harmoniser la définition retenue avec celle de l'OCDE qui a le mérite de placer l'accent sur la notion d'autonomie.

500. **Proposition d'amendement de la liste des techniques d'IA (annexe I).** D'autre part, le règlement européen sur l'IA devrait selon nous s'attacher à ne couvrir que les systèmes nouveaux dont les risques ne seraient pas déjà couverts par les règlementations existantes. À ce titre, la définition des techniques considérées comme de l'IA doit faire l'objet d'une attention particulière pour ne pas inclure des logiciels classiques et utilisés depuis des décennies. Il est important de replacer la notion d'IA « symbolique » au centre du deuxième point tout en supprimant la mention des techniques de programmation logique traditionnelle qui ne posent pas de problèmes particuliers en matière d'opacité ou de complexité. Concernant les systèmes experts, ils peuvent être audités, contrôlés et certifiés de façon satisfaisante. Leur intégration dans le champ de la nouvelle règlementation ne devrait intervenir que lorsqu'ils se fondent sur des techniques nouvelles, opaques ou qu'ils fonctionnent avec un degré d'autonomie plus important. La suppression du terme « systèmes experts » dans le point b) n'empêcherait pas leur inclusion dans le champ d'application du règlement s'ils étaient conçus à partir des autres techniques visées dans les définitions.

Au vu des éléments développés précédemment, l'annexe I pourrait être amendée comme suit :

(1) *Approches d'apprentissage automatique, y compris d'apprentissage supervisé, non supervisé et par renforcement, utilisant une grande variété de méthodes, y compris l'apprentissage profond.*

(2)*Approches* ***symboliques*** *fondées sur la logique et les connaissances, y compris la représentation des connaissances, ~~la programmation inductive (logique)~~, les bases de connaissances, les moteurs d'inférence et de déduction~~,~~* ***et*** *le raisonnement ~~(symbolique) et les systèmes experts~~.*

(3)*~~Approches statistiques,~~ estimation bayésienne~~, méthodes de recherche et d'optimisation.~~*

501. **Transition.** Ces propositions d'amendements permettraient d'éviter d'inclure tous les systèmes informatiques ou algorithmiques dans le champ d'application de la règlementation en se concentrant sur les systèmes présentant des caractéristiques nouvelles et problématiques. On comprend aisément que la Commission ait du mal à trancher sur ce sujet tant la littérature scientifique n'est, elle-même, pas homogène. La volonté de construire un cadre horizontal,

commun à tous les secteurs d'activité, se heurte à un obstacle d'envergure : réguler de façon uniforme des applications totalement différentes et nombreuses, allant des algorithmes de sélection de CV aux systèmes utilisés dans le pilotage d'infrastructures critiques. L'exercice, difficile, a abouti à des définitions des IA à haut risque inadaptées au regard de la réalité des cas d'usage dans les différents secteurs d'activité, et *a fortiori* dans le secteur de l'électricité.

B/ Une définition inappropriée des systèmes d'IA à haut risque

502. **Le champ d'application des dispositions les plus contraignantes.** Le Titre III du projet de règlement est consacré aux systèmes d'IA considérés comme à haut risque et soumis en conséquence à un corpus d'exigences de conformité. Les systèmes concernés sont d'abord ceux destinés à être utilisés comme composants de sécurité de produits déjà soumis à des processus d'évaluation *ex ante* par un tiers en vertu de l'un des textes européens énumérés à l'annexe II régissant notamment les domaines de l'aviation civile ou de l'automobile[1068]. Aucun de ces textes ne régit les produits ou composants de sécurité utilisés dans le secteur de l'électricité[1069]. L'article 6 renvoie ensuite à l'annexe III qui contient une liste des systèmes considérés d'emblée comme à haut risque. Les articles 8 et suivants dressent l'inventaire de l'ensemble des exigences très contraignantes auxquelles les fournisseurs de systèmes d'IA entrant dans cette catégorie devront se plier. L'annexe III relève ainsi d'une importance capitale puisqu'elle conditionne l'applicabilité des coûteuses contraintes juridiques.

503. **Le secteur de l'électricité directement visé.** Les cas d'usage dans le secteur de l'électricité sont directement ou indirectement visés dans l'annexe susmentionnée. En effet, le point 2 concerne le domaine de la *« gestion et de l'exploitation des infrastructures critiques »* en précisant que sont visés les *« les systèmes d'IA destinés à être utilisés en tant que composants de sécurité dans la gestion et l'exploitation du trafic routier et dans la fourniture d'eau, de gaz, de chauffage et d'électricité »*[1070]. De plus, le point 5, concernant les systèmes utilisés dans *«*

[1068] *AI Act*, article 6, 1.
[1069] *AI Act*, annexe II.
[1070] *AI Act*, annexe III, 2.

l'accès aux services privés essentiels, aux services publics et aux prestations sociales », et en particulier *« les systèmes d'IA destinés à être utilisés pour évaluer la solvabilité des personnes physiques […] »*, pourrait également inclure un cas d'usage dans le secteur de l'électricité. En effet, le Considérant 37, qui précise ce que la Commission entend par *« services essentiels »*, mentionne expressément l'accès à l'électricité. Elle considère en effet que *« lorsque les systèmes d'IA sont utilisés pour déterminer si ces services devraient être refusés, réduits, révoqués ou récupérés par les autorités, ils peuvent […] porter atteinte à leurs droits fondamentaux »*. Même si le texte vise principalement les services fournis par les autorités publiques, on peut légitimement penser que les opérateurs agissant sur délégation de service public (Enedis et RTE) seraient concernés. À ce titre, un système d'IA qui viserait à analyser des données relatives aux clients finals afin de déterminer leur solvabilité et qui pourrait, en fonction des résultats, couper leur accès à l'électricité entrerait dans le champ de cette définition. Bien qu'une telle utilisation semble plutôt irréaliste, il serait légitime de la considérer comme à haut risque. Par conséquent, cette définition ne sera pas commentée dans la suite de notre propos, tout comme le reste de la liste figurant à l'annexe III qui concerne des domaines de cas d'usage très éloignés des utilisations dans le secteur de l'électricité.

504. **Une définition inadaptée des systèmes d'IA à haut risque dans les infrastructures critiques.** Dans les entreprises, une démarche de mise en conformité à une nouvelle règlementation passe en premier lieu par l'identification des situations pour lesquelles de nouveaux processus, dits de *compliance*, doivent être prévus. Pour étudier l'impact que pourrait avoir l'*AI Act* sur les entreprises du secteur de l'électricité, cette première étape a été réalisée chez EDF S.A. dans le cadre de la présente thèse. L'objectif était d'identifier parmi les 150 cas d'usage d'IA ceux qui seraient considérés à haut risque en vertu de la nouvelle règlementation. Toutefois, la formulation retenue par la Commission n'a pas permis aux experts sollicités de déterminer avec certitude les cas d'usage concernés. Cette expérience a permis d'identifier deux principales lacunes qui pourraient être corrigées.

505. **L'imprécision de la référence au secteur de l'électricité.** Premièrement, le terme « fourniture » d'électricité n'apporte que de la confusion puisqu'il ne couvre que l'une des activités composant le secteur de l'électricité, y compris au sens de la directive européenne sur

le marché intérieur de l'électricité[1071]. Le terme « approvisionnement en électricité » pourrait lui être préféré pour englober les systèmes d'IA utilisés dans tout le secteur de l'électricité, d'autant plus que les applications les plus risquées se trouveront naturellement dans la gestion des sites de production et du réseau.

506. **L'inclusion de systèmes déterministes ne présentant pas de risques nouveaux.** Deuxièmement, le milieu industriel et notamment les secteurs d'activité considérés comme critiques sont des domaines déjà fortement règlementés. Dans la production d'électricité, y compris nucléaire, et dans la gestion des infrastructures de réseau, les risques liés à l'usage de systèmes informatiques, notamment en termes de cybersécurité, ou de systèmes experts dans des fonctions critiques sont déjà régulés par de nombreux textes[1072] et supervisés par des autorités de régulation sectorielles[1073]. Des risques nouveaux apparaissent avec le recours à des systèmes d'IA lorsqu'ils présentent un certain niveau d'autonomie. En effet, dans les environnements critiques hautement régulés, toutes les actions entreprises répondent d'un protocole dûment validé par les autorités de surveillance. Un système qui ne serait utilisé qu'en tant qu'aide à la décision ne dispenserait pas les employés de leurs obligations règlementaires. Le risque n'apparaît finalement que lorsque le système est capable d'initier une action ou qu'il est intégré dans le processus de décision en remplacement de l'humain. C'est pourquoi il est essentiel d'ajouter la notion d'autonomie ou de prise de décision dans la définition figurant à l'annexe III de l'*AI Act*.

507. **L'exemple de la maintenance prédictive en centrale nucléaire.** Un exemple peut être donnée dans la production nucléaire d'électricité avec un cas d'usage de maintenance prédictive. Un système d'IA peut être utilisé pour analyser les données générées par des capteurs présents dans une centrale nucléaire et identifier des pistes d'optimisation de

[1071] *Directive (UE) 2019/944 du Parlement européen et du Conseil du 5 juin 2019 concernant des règles communes pour le marché intérieur de l'électricité et modifiant la directive 2012/27/UE*, publiée au JOUE n°L158/125 du 14 juin 2019, article 2, 12) définissant la *« fourniture d'électricité »* comme *« la vente, y compris la revente, d'électricité à des clients »* et ne couvrant donc pas les activités de transport, de distribution et de production.
[1072] Voir notamment le Code de la défense en matière de sécurité nucléaire.
[1073] L'ASN pour la production nucléaire d'électricité ; l'ANSSI pour la surveillance de la cybersécurité des infrastructures critiques du réseau électrique.

l'exploitation. Une telle application peut être utilisée de trois façons différentes, présentant trois niveaux de risque distincts.

Dans un premier cas, le système pourrait simplement produire un diagnostic identifiant les capteurs présentant des mesures anormales. L'utilisateur aura la charge d'analyser ce diagnostic pour identifier la cause de l'alerte et les actions possibles pour y remédier, avant de prendre une décision qui ne pourra de toute façon pas être contraire aux règles de sécurité en vigueur dans la conduite ou la maintenance d'une centrale nucléaire. Le niveau de risque est ici minime puisque le recours à l'IA ne s'accompagne pas d'un affranchissement des règlementations de sécurité en vigueur.

Dans une deuxième situation, l'application de maintenance prédictive pourrait, en plus de produire un diagnostic, proposer à l'utilisateur des recommandations d'opérations à mener pour résoudre les problèmes identifiés. Dans ce cas, la décision finale de choisir parmi les recommandations et réaliser une des opérations recommandées revient à l'utilisateur. Ce dernier est toujours contraint par les règles de sécurité encadrant l'exploitation de l'infrastructure critique et, le cas-échéant, par les processus validés par les autorités de régulation. Juridiquement, il ne peut donc pas prendre une décision qui mettrait en danger le bon fonctionnement de la centrale.

Le dernier cas envisageable correspondrait à un système d'IA qui réaliserait un diagnostic, recommanderait des actions et aurait la possibilité d'en initier une automatiquement. Une telle application présenterait un risque important si les règles de sécurité inhérentes à son environnement d'exploitation n'étaient pas intégrées dans le programme informatique. Un tel système devrait naturellement être considéré comme à haut risque et voir sa conception strictement encadrée.

508. **La nécessité d'une appréciation des risques partagées avec les parties prenantes et régulateurs sectoriels.** On constate ainsi que l'appréciation du risque lié à l'utilisation de l'IA diffère suivant l'environnement dans lequel elle est utilisée et l'usage précis qui en est fait. De ce point de vue, l'approche binaire de la Commission, qui consiste à porter elle-même l'évaluation des risques et dresser la liste exhaustive des systèmes considérés comme à haut

risque, est discutable[1074]. Certains considèrent également que cette approche peut aboutir à un effet de seuil incitant les industriels à développer uniquement des applications qui n'entreraient pas dans cette catégorie, en dépit des potentiels bénéfices qu'elles pourraient apporter[1075]. Dans une logique de co-régulation, il serait pertinent de partager l'exercice de l'évaluation des risques avec les parties prenantes sectorielles (entreprises et régulateurs) ou à tout le moins de faire figurer la notion d'autonomie dans la définition, qui est le critère principal du risque du point de vue des industriels.

509. **Proposition d'amendement de l'annexe III.** Pour répondre aux deux critiques présentées précédemment, nous proposons d'amender l'annexe III comme suit :

> *[…]*
>
> *Gestion et exploitation des infrastructures critiques:*
>
> *1. les systèmes d'IA* ***pouvant initier automatiquement des actions influençant leur environnement et*** *destinés à être utilisés en tant que composants de sécurité dans la gestion et l'exploitation du trafic routier et dans ~~la fourniture d'~~* ***l'approvisionnement en*** *eau, ~~de~~ gaz, ~~de~~ chauffage et ~~d'~~électricité.*
>
> *[…]*

510. **Une lacune due à un défaut de prise en compte des spécificités sectorielles.** L'ensemble des critiques adressées jusqu'ici traduisent un défaut de maîtrise des secteurs régulés par la Commission. Il s'agit d'un écueil attendu dans une démarche de régulation transversale qui peut difficilement prendre en considération toutes les spécificités sectorielles. Néanmoins, sans évolution du texte pour prendre en compte les retours des acteurs régulés, les choix effectués risquent d'aboutir à des définitions inappropriées et des contraintes injustifiées sur des applications peu risquées.

511. **Transition.** Outre le contenu du projet d'*AI Act* en lui-même, les probables évolutions qu'il va subir méritent notre attention. Les débats au Conseil de l'UE et au Parlement européen doivent aboutir à l'adoption commune d'un texte amendé. Bien que le travail législatif ne soit

[1074] B. Mathis, *op. cit.*
[1075] M. Veale, F. Zuiderveen Borgesius, *op.cit*.

pas encore finalisé, on peut constater dans les amendements proposés par les différentes institutions que certains semblent faire consensus. L'une de ces propositions vise à exclure du champ d'application du règlement des systèmes pouvant être utilisés dans divers contextes[1076]. Cette exclusion nous semble malvenue et pourrait avoir des conséquences négatives sur le développement de l'IA dans l'industrie.

C/ Une exclusion discutable des systèmes d'IA à usage général

512. **La volonté du Conseil et du Parlement européen d'exclure les systèmes d'IA généraux du champ d'application de l'*AI Act*.** La lecture des premiers rapports publiés par le Conseil de l'UE et les différentes commissions du Parlement européen compétentes sur le sujet de l'*AI Act* présente un intérêt non négligeable. Elle permet en effet d'avoir un aperçu du contenu des débats dans les différentes institutions, de leur position respective mais aussi de celle des groupes d'influence cherchant à faire évoluer la régulation en construction.

La proposition d'exclure les systèmes d'IA à usage général – ou généraux – du champ d'application du règlement européen sur l'IA est apparue la première fois dans le texte de compromis publié par la présidence slovène du Conseil de l'UE le 29 novembre 2021[1077]. Dans ce document, le Conseil propose la création d'un considérant[1078] et d'un article[1079] prévoyant que la commercialisation ou la mise en service de systèmes d'IA capables de réaliser des tâches générales telles que la reconnaissance d'image et de texte, la traduction ou la production de contenus ne devrait être soumises à aucune des dispositions contraignantes de l'*AI Act*. En revanche, la personne qui commercialiserait ou mettrait en service un tel système pour une finalité précise devrait être considérée comme le fournisseur du système et donc être le garant du respect de l'ensemble des obligations du règlement. Cette proposition figure dans l'avis final

[1076] Par exemple, des systèmes généraux de reconnaissance d'image, de traitement du langage naturel.
[1077] CONSEIL DE L'UE, *Presidency compromise text on the proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (AI Act) and amending certain Union legislative acts*, 29 Novembre 2021, n° 2021/0106(COD), n°14278/21, Présidence slovène.
[1078] *Ibid.*, p. 26.
[1079] *Ibid.*, pp. 65–66.

du Conseil, publié le 25 novembre 2022[1080], à la seule différence qu'il est prévu un renvoi à un acte délégué pour préciser ultérieurement les modalités d'application de certaines dispositions aux systèmes à usage général.

Cette proposition a été reprise dans deux documents émanant de commissions du Parlement européen compétentes sur l'*AI Act*. L'un a été publié par la commission des affaires juridiques (JURI) le 2 mars 2022[1081], l'autre par la commission de l'industrie, de la recherche et de l'énergie (ITRE) le 3 mars 2022[1082]. Les deux publications comportent, dans leurs propositions d'amendements, une définition similaire des systèmes d'IA à usage général[1083] qui laisse supposer soit une concertation soit une influence commune. De la même manière, les deux commissions proposent des amendements très proches visant à exclure les systèmes d'IA généraux et à faire peser la charge de la mise en conformité à la personne qui, en utilisant le système général, lui attribue une finalité précise[1084]. Les formulations retenues rejoignent en grande partie celles employées dans le texte de compromis publié en novembre 2021 par le Conseil.

513. **Une proposition partiellement justifiée.** Dans l'*AI Act*, le corpus d'obligations contraignantes encadrant la conception des systèmes d'IA à haut risque exige des fournisseurs qu'ils justifient leurs choix (de données d'entraînement, de mesures de sécurité et d'autres) au regard de la destination du système[1085], définie comme « *l'utilisation à laquelle un système d'IA est destiné par le fournisseur, y compris le contexte et les conditions spécifiques*

[1080] CONSEIL DE L'UE, *Orientation générale sur la Proposition de règlement du Parlement européen et du Conseil établissant des règles harmonisées concernant l'intelligence artificielle (législation sur l'intelligence artificielle) et modifiant certains actes législatifs de l'Union*, 25 novembre 2022, 2021/0106(COD).
[1081] COMMITTEE ON INDUSTRY, RESEARCH AND ENERGY (ITRE), *Draft opinion on the proposal for a regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts (COM(2021)0206 – C9-0146/2021 – 2021/0106(COD))*, 3 mars 2022, 2021/0106(COD).
[1082] COMMITTE ON LEGAL AFFAIRS (JURI), *Draft opinion on the proposal for a regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union Legislative Acts (COM(2021)0206 – C9-0146/2021 – 2021/0106(COD)*, 2 mars 2022, 2021/0106(COD).
[1083] COMMITTEE ON INDUSTRY, RESEARCH AND ENERGY (ITRE), *op. cit.*, amendement n°27, p. 19 ; COMMITTEE ON LEGAL AFFAIRES (JURI), *op. cit.*, amendement n°38, p. 25.
[1084] COMMITTEE ON INDUSTRY, RESEARCH AND ENERGY (ITRE), *op. cit.*, amendement n°69, p. 37 ; COMMITTEE ON LEGAL AFFAIRES (JURI), *op. cit.*, amendement n°32, p. 23 et amendement n° 158, p. 77.
[1085] Voir *AI Act*, articles 9 à 15.

d'utilisation »[1086]. Il est vrai que certains logiciels ou modèles d'IA sont commercialisés sans être affectés à une utilisation précise. Les commercialisateurs de tels produits, majoritairement les grandes entreprises du numérique et les start-ups fournissant des services d'informatique en nuage, seraient dans l'impossibilité de documenter la conformité aux articles 9 à 15 de l'*AI Act*, contenant les exigences applicables aux IA à haut risque. Il n'est donc pas surprenant que ces entreprises se soient mobilisées pour faire entendre leur voix et influencer le processus législatif européen afin d'exclure les systèmes généraux du champ du règlement[1087].

514. **Les possibles effets pervers de l'exclusion des systèmes d'IA à usage général.** Si l'exclusion des systèmes d'IA généraux est nécessaire pour préserver la dynamique d'innovation, elle est avant tout à la faveur des entreprises de services numériques qui développent des systèmes d'IA sur étagère. Il peut s'agir, par exemple, de modèles de reconnaissance d'image pré-entraînés pouvant être utilisés tant pour des applications récréatives que militaires. Les géants du numérique tels que Google ou Amazon Web Services sont les principaux fournisseurs de tels systèmes logiciels, utilisables dans tous les secteurs d'activité. Les entreprises développant ce type de produits n'auraient ainsi à supporter aucun des coûts de mise en conformité, lesquels incomberaient exclusivement à l'entreprise utilisant le système acheté pour une finalité précise et qui serait alors considérée comme fournisseur au sens de l'*AI Act*. Outre le coût, l'utilisateur devenu fournisseur aura également la charge de contrôler la qualité des données d'entraînement, la sécurité du système et bien d'autres aspects au titre des articles 9 à 15. N'ayant pas conçu lui-même l'IA, le fournisseur pourrait alors se retrouver à son tour dans l'impossibilité de se mettre en conformité puisqu'il ne disposerait pas de l'ensemble des informations nécessaires, telles que les jeux de données ayant servi à l'entraînement de l'IA par exemple. Une solution aurait été d'imposer au fournisseur du système général de fournir l'ensemble des informations nécessaires à la mise en conformité de son client s'il est soumis aux exigences de l'*AI Act* en raison de l'application du système à une

[1086] *AI Act*, article 3, (12).

[1087] Voir par exemple la réponse de Google à la suite de la publication de l'*AI Act* consacrant plusieurs pages à la nécessité d'exclure les produits « sur étagère » et les systèmes généraux du champ d'application du texte : GOOGLE, *Consultation on the EU AI Act Proposal : Google's submission*, 15 juillet 2021, pp. 3–6, disponible en ligne : <https://ec.europa.eu/info/law/better-regulation/have-your-say/initiatives/12527-Artificial-intelligence-ethical-and-legal-requirements/F2662492_en>, consulté le 19 juin 2022.

finalité précise. Toutefois, ce n'est pas l'esprit des amendements proposés qui adoptent une formulation conditionnelle et non-contraignante en prévoyant seulement que le fournisseur de l'IA générale *« devrait travailler avec les utilisateurs de leurs produits pour les aider à respecter certaines exigences du règlement en leur offrant une expertise technique »*[1088]. Les parlementaires confirment dans leur justification que l'objectif n'est pas de contraindre le fournisseur initial mais seulement d'encourager un échange de bonne foi, fondé sur le respect des principes du marché et de la protection du secret des affaires[1089]. Si l'on comprend que le secret des affaires doit être une limite à ne pas franchir, le refus de coopérer du fournisseur du système d'IA général se répercuterait en réalité sur l'utilisateur du système qui serait dans l'impossibilité de remplir les exigences de conformité de l'*AI Act* et s'exposerait à de lourdes sanctions.

515. **L'effet contre-productif d'une exclusion des systèmes d'IA à usage général.** Tous les acteurs de la chaîne de valeur de l'IA devraient poursuivre le même objectif : développer des systèmes d'IA sûrs. La conformité aux exigences prévues dans l'*AI Act* pour les systèmes d'IA à haut risque (notamment en matière de qualité des données) doit être garantie à toutes les étapes du cycle de vie des systèmes, de sa conception à son utilisation. La proposition initiale de la Commission établissait une approche proportionnée en faisant peser les exigences sur les fournisseurs et non sur les utilisateurs des systèmes d'IA qui ne disposent pas des informations techniques, des données et de l'expertise nécessaires à la mise en conformité. La proposition d'amendement visant à exclure les systèmes d'IA à usage général renverse la situation en faisant peser la charge de la conformité sur l'utilisateur qui en détermine la finalité. Une telle modification ne serait acceptable que si l'utilisateur avait la certitude que le fournisseur du système général lui donnera toutes les informations nécessaires pour se mettre en conformité et dont il est le seul détenteur telles que les jeux de données d'entraînement (article 10) ou la documentation technique (article 11) relatives aux *« méthodes et étapes suivies pour le développement »*[1090] et aux *« spécifications de conception du système, à savoir la logique*

[1088] COMMITTEE ON INDUSTRY, RESEARCH AND ENERGY (ITRE), *op. cit.*, amendement n°69, article 28a, al. 1er, p. 37, traduction libre.
[1089] COMMITTEE ON INDUSTRY, RESEARCH AND ENERGY (ITRE), *op. cit.*, p. 38.
[1090] *AI Act*, annexe IV, 2, a).

générale du système d'IA et des algorithmes; les principaux choix de conception »[1091] par exemple.

516. **Proposition d'amendement.** L'exclusion des systèmes d'IA à destination générale ne figurait pas dans la proposition initiale de la Commission européenne. Proposée par Google[1092] puis reprise par le Conseil de l'UE et deux rapports parlementaires, elle répond à une problématique réelle. Bien que la proposition ne figure pas dans le rapport des deux commissions co-responsables de l'*AI Act* au Parlement européen[1093], il est fort probable qu'elle revienne dans les débats visant à amender le projet de texte. Si son inclusion venait à constituer un point de blocage, nous proposons ici un amendement de compromis, plus acceptable que l'exclusion stricte et déresponsabilisante qui était discutée jusqu'à maintenant :

Article 28a[1094]

General purpose AI systems

1. The placing on the market, putting into service or use of general purpose AI systems shall not, on its own, make those systems subject to this Regulation.

Providers of general purpose AI systems shall ***actively*** *work with users of their products to aid them in fulfilling* ***all*** ~~***certain***~~ *requirements set out in this Regulation by providing technical expertise,* ***as well as all relevant data***. ~~***The shift from user to provider in Article 28 paragraph1, point (ca), still applies***~~*. Such an exchange shall be in full respect of trade secrets* ~~***and current market indicators***~~ *and shall have within its scope only those obligations, relating to the technical design and development of the system before an intended purpose is attributed to it. The provider of the general purpose AI system shall register the system in the Union database as referred to in Article 60.*

[1091] *AI Act*, annexe IV, 2, b).
[1092] Voir GOOGLE, *Consultation on the EU AI Act Proposal : Google's submission*, *op. cit.*, note 105.
[1093] COMMITTEE ON THE INTERNAL MARKET AND CONSUMER PROTECTION (IMCO) AND COMMITTEE ON CIVIL LIBERTIES, JUSTICE AND HOME AFFAIRS (LIBE), *Draft report on the proposal for a regulation of the European Parliament and of the Council on harmonised rules on Artificial Intelligence (Artificial Intelligence Act) and amending certain Union Legislative Acts (COM2021/0206 – C9-0146/2021 – 2021/0106(COD)*, 20 avril 2022, 2021/0106(COD), p. 159 : *« no AI system should be excluded ex-ante, either from the definition of "artificial intelligence" or by carving out exceptions for particular types of AI systems, including general purpose AI »*.
[1094] Amendement construit à partir de celui proposé par Eva Maydell (ITRE) dans son rapport pour opinion et conservé dans sa version anglaise en l'absence de traduction officielle : COMMITTEE ON INDUSTRY, RESEARCH AND ENERGY (ITRE), *op. cit.*, amendement n°69, article 28a, p. 37.

*2. Any person who **develops** or places on the market or puts into service under its own name or trademark **~~or uses~~** a general purpose AI system made available on the market or put into service for an intended purpose that makes it subject to this Regulation shall be considered the provider of the AI system in accordance with this Regulation.*

*3. ~~Paragraph 2 shall apply, mutatis mutandis, to any person who~~ **In the case where an actor** integrates a general purpose AI system made available on the market, with or without modifying it, into an AI system whose intended purpose makes it subject to this Regulation, **the provider shall work closely with the user of the AI system to ensure its compliance with all requirements set out in this Regulation.***

517. **Conclusion sur l'inadéquation du champ d'application de l'*AI Act*.** La question du champ d'application et des définitions retenues cristallisent une grande partie des débats dans la procédure législative européenne. L'étendue de la définition des systèmes d'IA, le défaut de précision dans les formules employées pour délimiter le champ des IA à haut risque et la probable introduction d'une exclusion pour les systèmes d'IA commercialisés sans être destinés à une finalité précise sont les trois principales préoccupations pour les acteurs du secteur de l'électricité. En effet, des définitions trop larges ou imprécises risqueraient de faire entrer dans le champ d'application du règlement des cas d'usage qui ne présentent qu'un degré de risque très limité ou des systèmes déjà bien encadrés par les règlementations sectorielles. L'exclusion des systèmes généraux, quant à elle, ferait peser la charge de la conformité sur les entreprises du secteur de l'électricité utilisatrices de systèmes d'IA sur étagère, en déresponsabilisant le fournisseur initial. Des propositions concrètes d'amendements ont été formulées dans le présent Paragraphe afin de répondre à ces problématiques.

518. **Transition.** L'*AI Act* innove sur de nombreux aspects. Il introduit de nouvelles définitions et entend créer des exigences inédites relatives au contrôle humain ou à la qualité des données notamment. Toutefois, comme on a pu le démontrer en première Partie, le développement de l'IA ne s'effectue pas dans un vide juridique et de nombreux textes lui sont déjà applicables. Le secteur de l'électricité est un bon cas d'étude puisqu'il fait l'objet lui-même d'une règlementation très fournie. La cohérence de l'*AI Act* avec le corpus juridique existant est donc primordiale.

§2 : Un défaut de cohérence avec les règlementations existantes

519. **Plan.** Comme le souligne la chercheuse Nathalie Smuha, la question de l'articulation de l'*AI Act* avec les législations existantes (ou en cours de construction) n'est pas clairement adressée dans le texte[1095]. Elle identifie, au niveau européen, plusieurs textes dont les dispositions pourraient entrer en contradiction avec les dispositions du projet de règlement sur l'IA : la Charte des droits fondamentaux dans l'Union européenne[1096], le RGPD[1097], la directive dite « police-justice »[1098] et la directive « MiFID II » concernant les marchés d'instruments financiers[1099]. Naturellement, tous ne concernent pas les cas d'usage de l'IA dans le secteur de l'électricité et ne seront donc pas étudiés pour cette raison dans la suite de notre propos. Il convient d'y ajouter également les textes européens en construction ou récemment adoptés, qui devraient être applicables plus ou moins au même moment que le règlement sur l'IA : le Data Act[1100], le Data Governance Act[1101], le DSA[1102] et le DMA[1103]. Le cadre de régulation du numérique souhaité par l'Union européenne promet d'être particulièrement complexe. Il est

[1095] N.A. Smuha, E. Ahmed-Rengers, A. Harkens, *et al.*, « How the EU Can Achieve Legally Trustworthy AI: A Response to the European Commission's Proposal for an Artificial Intelligence Act », *SSRN Electronic Journal*, 5 août 2021, p. 41.
[1096] *Charte des droits fondamentaux de l'Union européenne*, 2000/C 364/01, publiée au JOCE n°C364/3 du 18 décembre.
[1097] *Règlement (UE) 2016/679 du Parlement européen et du Conseil du 27 avril 2016 relatif à la protection des personnes physiques à l'égard du traitement des données à caractère personnel et à la libre circulation de ces données, et abrogeant la directive 95/46/CE (règlement général sur la protection des données)*, publié au JOUE n°L119/1 le 4 mai 2016, ci-après le « RGPD ».
[1098] *Directive (UE) 2016/680 du Parlement européen et du Conseil du 27 avril 2016 relative à la protection des personnes physiques à l'égard du traitement des données à caractère personnel par les autorités compétentes à des fins de prévention et de détection des infractions pénales, d'enquêtes et de poursuites en la matière ou d'exécution de sanctions pénales, et à la libre circulation de ces données, et abrogeant la décision-cadre 2008/977/JAI du Conseil*, publiée au JOUE n°L119/89 le 4 mai 2016, dite directive « police-justice ».
[1099] *Directive 2014/65/UE du Parlement européen et du Conseil du 15 mai 2014 concernant les marchés d'instruments financiers et modifiant la directive 2002/92/CE et la directive 2011/61/UE*, publiée au JOUE n°L173/349 le 12 juin 2014.
[1100] COMMISSION EUROPÉENNE, *Proposition de règlement du Parlement européen et du Conseil sur des règles harmonisées relatives à l'accès équitable aux données et à leur utilisation équitable (Data Act)*, 23 février 2022, COM (2022) 68 final.
[1101] COMMISSION EUROPÉENNE, *Proposition de règlement sur la gouvernance européenne des données (Data Governance Act)*, 25 novembre 2020, COM (2020) 767.
[1102] COMMISSION EUROPÉENNE, *Proposition de règlement relatif à un marché intérieur des services numériques (Digital Services Act) et modifiant la directive 2000/31/CE*, 15 décembre 2020, COM (2020) 825 final.
[1103] COMMISSION EUROPÉENNE, *Proposition de règlement relatif aux marchés contestables et équitables dans le secteur numérique (Digital Markets Act)*, 15 décembre 2020, COM (2020) 842 final.

donc particulièrement important que l'*AI Act* s'inscrive en cohérence avec le droit européen du numérique (**A**).

Outre le droit européen, l'articulation de l'*AI Act* avec les législations nationales relatives aux algorithmes et l'intégration de ses exigences dans les règlementations sectorielles doivent également faire l'objet d'une attention particulière (**B**).

A/ Une articulation complexe avec le droit européen

520. **Plan.** La création d'une règlementation spécifique aux systèmes d'IA ne peut se faire sans penser sa cohérence avec le RGPD (**1**). De plus, l'Union européenne devrait également veiller à la bonne articulation entre l'*AI Act* et tous ses autres projets de textes visant à réguler le numérique (**2**).

1. L'articulation avec le RGPD

521. **Les systèmes d'IA à haut risque traitant des données à caractère personnel : le risque d'une double mise en conformité.** Il est tout à fait probable que certains systèmes d'IA utilisés dans le secteur de l'électricité soient considérés comme à haut risque au sens de l'*AI Act* et opèrent dans le même temps des traitements de données à caractère personnel. En effet, il est envisageable qu'un système d'IA soit un jour utilisé dans des fonctions critiques du pilotage du réseau de distribution et qu'il nécessite le traitement en temps réel des courbes de charge des individus afin d'optimiser son fonctionnement[1104]. Une telle application devrait se conformer à la fois aux exigences de conformité de l'*AI Act* pour les systèmes d'IA à haut risque et aux dispositions du RGPD.

522. **Plan.** L'articulation entre les dispositions des deux textes peut dans certains cas s'avérer complexe (**a**), ce qui a conduit les autorités européennes de protection des données à exprimer un certain nombre de réserves sur le projet d'*AI Act* (**b**). De plus, outre les conflits probables

[1104] A. van der Mei, J.-P. Doomernik, « Artificial intelligence potential in power distribution system planning », *24th International Conference & Exhibition on Electricity Distribution (CIRED)*, 12–15 juin 2017, session 5.

entre les dispositions des deux textes, l'adoption de l'*AI Act* risque de conduire à la multiplication des autorités de contrôle au niveau européen et, par conséquent, générer des difficultés relatives à leur coordination (**c**).

a) Des difficultés d'articulation du fait du contenu des textes

523. **Le principe : la primauté du RGPD sur l'*AI Act*.** La proposition de la Commission tient compte de cette éventualité puisqu'elle mentionne à quelques reprises l'articulation de l'*AI Act* avec le RGPD. Ainsi, la Commission établit dans l'exposé des motifs que sa proposition de texte est construite de façon à garantir sa cohérence avec l'ensemble des textes européens qui pourraient être applicables aux systèmes d'IA à haut risque, qu'importe le secteur d'activités concerné. Elle serait à ce titre *« sans préjudice du règlement général sur la protection des données »*[1105] et le complèterait avec *« un ensemble de règles harmonisées concernant la conception, le développement et l'utilisation de certains systèmes d'IA à haut risque ainsi que des restrictions portant sur certaines utilisations de systèmes d'identification biométrique à distance »*[1106]. L'objectif de la Commission est donc de ne pas remettre en cause les dispositions du RGPD, voire de les compléter pour renforcer la protection des droits des individus sur les données les concernant. Pourtant, une partie de la doctrine s'interroge sur le fait de savoir si l'*AI Act* ne viendrait pas au contraire affaiblir les standards de protection existant[1107].

524. **L'exception : la primauté de l'*AI Act* sur le RGPD.** La proposition de règlement européen sur l'IA contient plusieurs références explicites au RGPD, preuve s'il en est que la Commission a bien saisi l'importance de l'articulation entre les deux textes. L'article 10 de l'*AI Act*, contenant les exigences relatives à la gouvernance des données et applicables aux systèmes d'IA à haut risque, crée une nouvelle base légale pour le traitement des données dites « sensibles »[1108], en principe prohibé par l'article 9 du RGPD[1109]. Là où certains y voient un

[1105] *AI Act*, exposé des motifs, p. 4.
[1106] *Ibid.*
[1107] N.A. Smuha, E. Ahmed-Rengers, A. Harkens, *et al.*, *op. cit.*, p. 42.
[1108] *AI Act*, article 10 (5).
[1109] RGPD, article 9 (1), posant comme principe l'interdiction du traitement des données à caractère personnel révélant *« l'origine raciale ou ethnique, les opinions politiques, les convictions religieuses ou philosophiques ou l'appartenance syndicale, ainsi que le traitement des données génétiques, des données biométriques aux fins*

recul de la protection de la vie privée des individus[1110], cette exception nous paraît être le corollaire logique de l'obligation imposée aux fournisseurs de systèmes d'IA à haut risque de mettre en place des mesures pour lutter contre les biais algorithmiques. La Commission semble l'avoir bien compris puisqu'elle accompagne cette possibilité d'un encadrement strict. Les fournisseurs ne peuvent en effet traiter des données sensibles que dans la mesure où cela est strictement nécessaire aux fins de la surveillance, de la détection et de la correction des biais et sous réserve de garanties appropriées pour les droits et libertés fondamentaux des personnes physiques[1111]. Par *« garanties appropriées »*, la Commission entend notamment *« des limitations techniques relatives à la réutilisation ainsi que l'utilisation des mesures les plus avancées en matière de sécurité et de protection de la vie privée, telles que la pseudonymisation, ou le cryptage lorsque l'anonymisation peut avoir une incidence significative sur l'objectif poursuivi »*[1112]. Des gardes-fous sont donc prévus pour limiter le risque pour la vie privée des individus tout en préservant la capacité des fournisseurs à se mettre en conformité avec les exigences relatives à la lutte contre la présence de biais discriminatoires dans les systèmes d'IA à haut risque. Dans le secteur de l'électricité, notre étude n'a pas révélé de cas d'usage nécessitant le traitement de données sensibles pour vérifier l'absence de biais. Les seules applications susceptibles d'être concernées sont les systèmes utilisés dans la relation entre les fournisseurs d'énergie et leurs clients. En effet, l'utilisation d'agents conversationnels artificiels (*chatbots*) ou de systèmes automatiques de réponse aux mails devront être testés pour vérifier que leur fonctionnement ne diffère pas suivant les origines raciales ou ethniques, les convictions religieuses ou l'orientation sexuelle des utilisateurs[1113]. Or, ces applications ont peu de chance d'être considérées comme des IA à haut risque en l'état actuel des définitions. Les fournisseurs d'agents conversationnels artificiels ne pourraient donc pas se prévaloir de l'exception présente à l'article 10 (5) de l'*AI Act* afin de mettre en œuvre les mesures nécessaires

d'identifier une personne physique de manière unique, des données concernant la santé ou des données concernant la vie sexuelle ou l'orientation sexuelle d'une personne physique ».

[1110] N.A. Smuha, E. Ahmed-Rengers, A. Harkens, *et al.*, *op. cit.*

[1111] *AI Act*, article 10 (5).

[1112] *Ibid.*

[1113] A. Schlesinger, K.P. O'Hara, A.S. Taylor, « Let's Talk About Race: Identity, Chatbots, and AI », *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, Association for Computing Machinery, New York, 2018, Paper 315, 1–14 ; W.D. Heaven, « How to Make a Chatbot That Isn't Racist or Sexist », in *Ethics of Data and Analytics*, dir. K. Martin, Auerbach Publications, New York, 1ère ed., 2022 chap. 2.2.

pour lutter contre les biais discriminatoires, ce qui est regrettable. L'articulation entre l'*AI Act* et le RGPD mériterait d'être clarifiée sur ce point. D'autres lacunes sont identifiées par la doctrine notamment au regard du traitement de la reconnaissance faciale qui serait contraire aux standards de protection présents dans le RGPD en raison des exceptions larges reconnues dans le texte[1114]. Ces dernières ne seront pas développées dans la suite de notre propos puisque ces dispositions concernent exclusivement l'usage de l'identification biométrique à des fins répressives, donc sans lien avec les applications dans le secteur de l'énergie. Les autorités de contrôle compétentes en matière de protection des données ont elles aussi exprimé leurs interrogations relatives à l'articulation entre l'*AI Act* et le RGPD.

b) Des réserves exprimées par les autorités de protection des données

525. **Le scepticisme des autorités européennes de protection des données sur l'*AI Act*.** Le Comité et le Contrôleur européens de la protection des données (les CEPD) ont publié le 18 juin 2021 un avis conjoint sur la proposition de règlement européen sur l'IA[1115]. Dans cet avis, les autorités européennes regrettent également le défaut d'articulation claire entre l'*AI Act* et le RGPD. Elles pointent de nombreux manques dans le projet de règlement qui aurait dû, selon elles, reprendre d'avantage les mécanismes prévus dans la législation européenne sur les données à caractère personnel. Pour renforcer la cohérence entre les deux textes, les CEPD font de nombreuses recommandations[1116].

526. **Les recommandations des CEPD portant sur le contenu des exigences de l'*AI Act*.** Concernant le contenu des exigences encadrant la conception et les conditions d'utilisation des systèmes d'IA, les autorités recommandent d'abord d'intégrer la notion de protection des données dès la conception (*privacy by design*) pour tous les systèmes, en particulier ceux conçus à partir de techniques d'apprentissage automatique et ayant été entraînés sur des données à

[1114] N.A. Smuha, E. Ahmed-Rengers, A. Harkens, *et al.*, *op. cit.*, p. 43.
[1115] COMITÉ EUROPÉEN DE LA PROTECTION DES DONNÉES ET CONTRÔLEUR EUROPÉEN DE LA PROTECTION DES DONNÉES, *Avis conjoint 5/2021 sur la proposition de règlement établissant des règles harmonisées sur l'intelligence artificielle*, 18 juin 2021.
[1116] Voir sur le sujet C. Crichton, « Artificial Intelligence Act : avis conjoint des CEPD », *Dalloz Actualité*, 2 juillet 2021.

caractère personnel[1117]. Ensuite, elles souhaitent également que l'*AI Act* reprenne la notion de « personnes concernées » en précisant leurs droits et les recours à leur disposition[1118]. Les droits visés renvoient au droit de ne pas faire l'objet d'une prise de décision automatisée[1119], aux droits à l'effacement et de rectification[1120] ainsi qu'au droit à l'effacement[1121], déjà présents dans le RGPD et qui devraient également être garantis dès la conception des systèmes d'IA. Enfin, les CEPD souhaitent renforcer les dispositions relatives à l'information des personnes, autre mesure phare du RGPD, lorsque leurs données personnelles sont traitées par une IA[1122]. Toutes ces propositions sont alignées avec les lignes directrices du G29 sur l'application du RGPD au profilage et autres décisions automatisées, qui reposent le plus souvent sur des systèmes d'IA[1123].

527. **Les recommandations des CEPD portant sur les modalités de mise en œuvre de l'*AI Act*.** Concernant les mécanismes de conformité, les autorités européennes compétentes en matière de protection des données souhaitent l'intégration de critères relatifs à la protection des données à caractère personnel dans les procédures d'examen de la conformité *ex ante*. Dans un objectif de simplicité, elles souhaitent éviter qu'un système d'IA marqué CE (donc certifié conforme aux exigences de l'*AI Act*) ne soit pas utilisé *a posteriori* d'une manière non conforme aux règles de protection des données personnelles[1124]. Elles recommandent à ce titre l'inclusion du concept de *privacy by design* dans le marquage CE prévu par l'*AI Act*[1125] et l'implication des CEPD dans la construction des normes harmonisées qui seront publiées par la Commission[1126].

[1117] *Ibid.*, 58 et 76.
[1118] *Ibid.*, 18.
[1119] RGPD, article 22.
[1120] COMITÉ EUROPÉEN DE LA PROTECTION DES DONNÉES ET CONTRÔLEUR EUROPÉEN DE LA PROTECTION DES DONNÉES, *op. cit.*, 58.
[1121] *Ibid.*, 60.
[1122] *Ibid.*
[1123] GROUPE DE TRAVAIL « ARTICLE 29 » SUR LA PROTECTION DES DONNÉES, *Lignes directrices relatives à la prise de décision individuelle automatisée et au profilage aux fins du règlement (UE) 2016/679*, 3 octobre 2017, n°WP251.
[1124] COMITÉ EUROPÉEN DE LA PROTECTION DES DONNÉES ET CONTRÔLEUR EUROPÉEN DE LA PROTECTION DES DONNÉES, *op. cit.*, 75.
[1125] *Ibid.*, 23.
[1126] *Ibid.*, 76.

528. **Le risque de superposition des textes.** Il est évident que les institutions européennes doivent veiller la bonne articulation entre l'*AI Act* et la législation en matière de protection des données. Toutefois, il ne nous semble pas pertinent, comme le suggèrent les CEPD dans leur dernière recommandation, de mentionner une nouvelle fois les droits reconnus aux individus par le RGPD, ce qui pourrait ajouter de la confusion quant à leur champ d'application. En effet, les droits du RGPD sont reconnus à toutes les personnes concernées par un traitement de données personnelles, ce qui englobe déjà tous les systèmes d'IA susceptibles de traiter de telles données. Il serait néanmoins très utile de pouvoir bénéficier de normes harmonisées précisant les mesures techniques à mettre en œuvre pour se conformer aux deux corpus simultanément. Cela nécessiterait la création d'un double standard : l'un relatif à la mise en conformité des systèmes d'IA à haut risque ne traitant pas de données à caractère personnel, l'autre relatif aux systèmes à haut risque traitant des données personnelles et soumis à ce titre aux dispositions du RGPD qui viendraient s'ajouter à celles de l'*AI Act*. La pertinence de la majorité des remarques adressées par les CEPD dans leur avis conjoint démontre l'utilité de l'implication des autorités de régulation dans le processus de création de la norme. La proposition de la Commission prévoit également la création d'une nouvelle entité au niveau européen (le Comité européen de l'IA) chargée notamment de contribuer à la coopération des autorités de contrôle nationales et d'assurer l'application harmonisée du texte dans l'Union européenne[1127]. Cette organisation risque toutefois de démultiplier les autorités de contrôle et d'aboutir à des divergences locales dans la mise en œuvre de l'*AI Act*.

c) Une regrettable multiplication des autorités de contrôle

529. **La création de nouvelles autorités de contrôle par l'*AI Act*.** Dans sa première proposition du 21 avril 2021, la Commission européenne a fait le choix d'une mise en œuvre décentralisée du règlement sur l'IA[1128]. Concrètement, les États membres devront chacun désigner une autorité de contrôle nationale chargée de l'application du règlement, du rôle de

[1127] *AI Act*, article 56.
[1128] C. Crichton, « Projet de règlement sur l'IA (II) : une approche fondée sur les risques », *Dalloz Actualité*, 4 mai 2021.

point de contact unique auprès des institutions européennes et de la représentation de l'État membre au sein du Comité européen de l'IA[1129]. Les examens de conformité au règlement pourront être conduits par des organismes désignés par une autorité dite « notifiante » chargée de mettre en place et d'accomplir les procédures nécessaires à l'évaluation, à la désignation et au contrôle des tiers indépendants pouvant être chargés d'évaluer la conformité aux exigences de l'*AI Act*[1130]. Le modèle de gouvernance présente de nombreux points communs avec celui en place pour l'application du RGPD.

530. **Les conséquences négatives de la multiplication des autorités de contrôle.** Pour cette raison, plusieurs commentateurs s'inquiètent des potentiels effets négatifs que pourrait avoir cette mise en œuvre décentralisée et qui ont déjà été constatés dans l'application du RGPD. Le chercheur indépendant Bruno Mathis rappelle à cet égard que l'application du RGPD a révélé d'importantes divergences tant dans l'interprétation du règlement par les autorités de contrôle nationales que dans les pratiques de supervision[1131]. L'application d'un modèle similaire à un domaine encore plus technique – l'IA – risque en effet de conduire à des écarts encore plus importants. Cette préoccupation est partagée dans la doctrine internationale[1132] qui s'inquiète également des potentielles disparités dans les moyens à la disposition des autorités de contrôle nationales[1133]. Une idée de solution pourrait être tirée de l'exemple en matière de lutte anti-blanchiment. Pour éviter que des entreprises choisissent leur pays d'établissement en fonction de la faiblesse de la supervision en matière de lutte anti-blanchiment, la Commission européenne a proposé la création d'une autorité de contrôle unique au niveau européen[1134]. L'encadrement des systèmes d'IA est un domaine au moins aussi technique que la supervision des systèmes d'information dans le secteur financier. Il serait donc tout à fait pertinent de

[1129] *AI Act*, article 59 et article 3 (42).
[1130] *AI Act*, article 30 et article 3 (19).
[1131] B. Mathis, « Proposition de règlement européen sur l'intelligence artificielle : le regard d'un praticien », *RLDI*, 2022, n°192, 4179, p. 44.
[1132] M. Veale, F. Zuiderveen Borgesius, « Demystifying the Draft EU Artificial Intelligence Act », *Computational Law Review International*, juillet 2021, vol. 22, n°4, 101–102.
[1133] N.A. Smuha, « Beyond the Individual: Governing AI's Societal Harm », *Internet Policy Review*, 30 septembre 2021, vol. 10, n°3, p. 21 ; ACCESS NOW, *Two years under the EU GDPR: an implementation progress report*, Rapport, 2020, p. 10, disponible en ligne : <https://www.accessnow.org/cms/assets/uploads/2020/05/Two-Years-Under-GDPR.pdf>, consulté le 24 juin 2022.
[1134] COMMISSION EUROPÉENNE, *Proposition de règlement n°2021/0240 instituant l'Autorité de lutte contre le blanchiment de capitaux et le financement du terrorisme*, COM (2021) 421 final, 20 juillet 2021.

reproduire ce schéma dans la régulation de l'IA[1135] sans attendre de constater les mêmes écueils que l'on a pu constater à la fois dans le secteur financier et dans l'application du RGPD.

531. **Synthèse des propositions sur l'articulation entre *l'AI Act* et le RGPD.** Partant du constat que de nombreux systèmes d'IA, y compris à haut risque, dans le secteur de l'électricité peuvent nécessiter le traitement de données à caractère personnel, il nous apparaît important de renforcer la cohérence entre l'*AI Act* et le RGPD. Cela peut passer, de façon contre-intuitive, par le refus d'intégrer dans l'*AI Act* des dispositions déjà présentes dans le RGPD, contrairement à ce que recommande les CEPD. Il convient en revanche d'harmoniser les standards et certificats de conformité pour éviter de créer de la confusion pour les fournisseurs de systèmes d'IA et leurs utilisateurs. En plus de la nécessaire harmonisation des contenus des deux textes, il est important que la Commission tire les enseignements de la mise en œuvre du RGPD. À ce titre, la décentralisation de l'application du règlement semble peu opportune. Une réflexion doit être menée pour déterminer si la gouvernance de l'*AI Act* ne pourrait pas plutôt s'appuyer sur une autorité de contrôle unique au niveau européen sur le modèle de l'Autorité de lutte contre le blanchiment des capitaux dont la création a été proposée en 2021 pour répondre aux divergences locales dans l'application de la législation européenne. À défaut, il est primordial que la Commission apporte des garanties supplémentaires pour que les autorités de contrôle nationales disposent des moyens financiers, techniques et humains pour assurer une mise en œuvre cohérente de l'*AI Act* sur le territoire européen.

2. L'articulation avec les projets européens de régulation du numérique

532. **La nécessaire cohérence entre l'*AI Act* et les autres projets européens de régulation du numérique.** Le règlement sur l'IA est le fruit d'une dynamique européenne plus générale visant à créer un cadre juridique pour le numérique. Ce cadre juridique est composé d'une

[1135] B. Mathis, *op. cit.* ; Y. Paquier, *Le principe de transparence des traitements algorithmiques : de l'étude juridique d'un enjeu démocratique*, Thèse pour le doctorat en droit public, 10 novembre 2021, Université de Caen Normandie, p. 294 et s.

multitude de textes aujourd'hui au stade de projets plus ou moins avancés, ou récemment adoptés. Ils visent à la fois à stimuler l'innovation sur le territoire européen et à promouvoir des technologies éthiques au service des citoyens[1136]. Certains contribuent également à la politique européenne en matière de transition écologique, jumelle de la transition numérique[1137]. Etant donné que la plupart des textes ne sont pas encore pleinement applicables, l'apport principal de notre propos est avant tout d'identifier les textes dont l'articulation avec l'*AI Act* doit être réfléchie plutôt que d'analyser leurs interactions dans le détail, ce qui n'aurait pas de sens au vu de leur faible niveau de maturité.

533. **Le *Data Act*.** Le premier texte identifié est le *Data Act*[1138], un projet de règlement européen visant à faciliter les échanges de données dans la droite ligne du règlement sur la libre circulation des données non personnelles[1139]. Le texte impose notamment aux fabricants d'objets connectés et de services associés de concevoir leurs produits de telle sorte que les données produites par ces derniers soient facilement accessibles par les utilisateurs[1140]. De plus, le *Data Act* créerait un droit pour les utilisateurs de demander au fabricant de l'objet connecté de transférer les données générées à un tiers en vue de la fourniture de services complémentaires[1141]. On comprend aisément l'objectif de la Commission qui est de libérer les données en favorisant leur partage entre entreprises (*BtoB*) et des entreprises vers les utilisateurs (*BtoC*)[1142]. Le texte contient également des dispositions permettant aux autorités publiques d'imposer aux entreprises la communication de données pour des motifs exceptionnels d'intérêt général (*BtoG*)[1143]. Dans la pratique, il est courant que la fourniture de services sur des données

[1136] COMMISSION EUROPÉENNE, *Shaping Europe's digital future*, Communication de la Commission au Parlement européen, au Conseil, au Comité économique et social européen et au Comité des régions, 19 février 2020, COM (2020) 67 final ; B. Bertrand, « Chronique Droit européen du numérique - La politique européenne du numérique : une vision politique européenne ? », *RTD Eur.*, 2022, p. 449.
[1137] B. Bertrand, « The Twin Digital and Green Transition », *RTD eur.*, 2022, p. 619.
[1138] COMMISSION EUROPÉENNE, *Proposition de règlement du Parlement européen et du Conseil sur des règles harmonisées relatives à l'accès équitable aux données et à leur utilisation équitable (Data Act)*, 23 février 2022, COM (2022) 68 final, ci-après le *Data Act*.
[1139] *Règlement (UE) 2018/1807 du Parlement européen et du Conseil du 14 novembre 2018 établissant un cadre applicable au libre flux des données à caractère non personnel dans l'Union européenne*, publié au JOUE n°L303/59 le 28 novembre 2018.
[1140] *Data Act*, article 3.
[1141] *Data Act*, article 5.
[1142] J. Keller, « Le Data Act : de nouvelles règles de partage des données », *Dalloz Actualité*, 8 mars 2022.
[1143] *Data Act*, articles 14 et s.

passe par le recours à des systèmes d'IA. Dans le secteur de l'électricité, les consommateurs peuvent se voir proposer des services d'analyse de leur consommation ou d'optimisation énergétique[1144]. Dès lors, si le *Data Act* entrait en application, il serait tout à fait envisageable que des systèmes d'IA à haut risque soient conçus par des tiers à partir de données obtenues auprès du fabricant des objets connectés avec le consentement de l'utilisateur. Le tiers destinataire de données, déjà soumis à certaines dispositions du *Data Act*[1145], pourrait dans ce cas être également qualifié de fournisseur d'IA à haut risque et devoir se conformer aux exigences de conformité de l'*AI Act*. Parmi elles figurent plusieurs dispositions relatives à la qualité des données utilisées dans la conception du système d'IA. Pour se mettre en conformité, le fournisseur doit notamment pouvoir disposer d'informations relatives aux modalités de collecte, aux mesures de sécurité mises en œuvre ou aux éventuelles modifications des données par le fabricant de l'objet connecté[1146]. Le *Data Act* devrait prendre en compte cette éventualité dans ses dispositions en ajoutant ces informations aux données devant être mises à disposition d'un tiers fournisseur de service sur demande de l'utilisateur.

534. **Le *Data Governance Act*.** Un deuxième texte dont la cohérence avec l'*AI Act* devrait être étudiée est le *Data Governance Act*[1147]. Ce dernier vise notamment à faciliter la réutilisation des données protégées du secteur public et à promouvoir le partage de données volontaire à des fins d'intérêt général[1148]. La volonté de la Commission de lever les obstacles juridiques et techniques à l'ouverture des données est bienvenue. Toutefois, pour que les données partagées puissent être exploitées, il est important que les fournisseurs de systèmes d'IA disposent de l'ensemble des informations nécessaires à leur mise en conformité avec l'*AI Act*. Comment un fournisseur d'IA pourrait-il justifier de la qualité des données utilisées lorsqu'elles ont été obtenues auprès de tiers s'ils ne disposent pas de l'historique des opérations réalisées ainsi que du détail des mesures de sécurité mises en œuvre pour protéger leur intégrité ? Ces points

[1144] A. Sozontov, M. Ivanona, A. Gibadullin, « Implementation of artificial intelligence in the electric power industry », *E3S Web of Conferences*, 2019, vol. 114, n°01009, spec. p. 2.
[1145] *Data Act*, Chapitre 3, articles 8 à 12.
[1146] *AI Act*, article 10 et annexe IV.
[1147] COMMISSION EUROPÉENNE, *Proposition de règlement sur la gouvernance européenne des données (Data Governance Act)*, 25 novembre 2020, COM (2020) 767.
[1148] CONSEIL DE L'UE, « Le Conseil approuve l'acte sur la gouvernance des données », *Communiqué de presse*, 16 mai 2022.

peuvent sembler relever du détail mais ils sont en réalité cruciaux si l'on veut convertir le partage de données en développement de systèmes d'IA vertueux et sécurisés.

535. **La regrettable création d'une comitologie européenne complexe sur les sujets numériques.** Il convient de noter également que le *Data Act*, le *Data Governance Act* et l'*AI Act* proposent tous les trois la création de comités au niveau européen[1149] ou d'autorités de contrôle au niveau national[1150]. La multiplication des comités consultatifs, des comités d'harmonisation de la régulation et des autorités de contrôle risque de créer un cadre de régulation du numérique particulièrement complexe au niveau européen. Ces nouvelles entités disposeraient d'ailleurs de missions proches (expertise auprès des institutions européennes, surveillance de la mise en œuvre harmonisée de la régulation, sanction des manquements) dans des domaines très techniques mais pourtant très liés (l'IA et le partage des données). La mutualisation des ressources et des compétences par la création d'une entité de supervision unique au niveau européen nous semblerait plus adaptée.

536. **Pour un moratoire en vue d'assurer la cohérence du cadre européen en construction.** Le *Data Act* et le *Data Governance Act* ne sont que deux exemples de textes au niveau européen et dont la cohérence avec l'*AI Act* devrait être mûrement réfléchie. Le *Digital Services Act*[1151] et le *Digital Markets Act*[1152] doivent également faire l'objet d'une attention particulière puisqu'ils contiennent des dispositions encadrant l'activité des grandes plateformes (y compris des algorithmes qu'elles utilisent) et des réseaux sociaux (notamment dans le cadre de la lutte contre la haine en ligne et la désinformation qui repose en partie sur le fonctionnement des algorithmes). Néanmoins, leur faible lien avec les applications de l'IA dans le secteur de l'électricité justifie le fait que nous ne nous y attardions pas dans la présente thèse. Certains

[1149] Voir notamment : *Data Governance Act*, articles 26 et 27 proposant la création d'un Comité européen de l'innovation chargé de conseiller la Commission et de favoriser l'interopérabilité des données au niveau européen ; AI Act, articles 56 à 58 proposant la création du Comité européen de l'IA.
[1150] *Data Act*, article 31 établissant un schéma de mise en œuvre de la régulation similaire à celui retenu dans l'*AI Act* avec la nomination d'une « autorité compétente » dans chaque État membre. Toutefois, contrairement à l'autorité de contrôle pour l'IA, l'autorité compétente en matière de données aura également la charge d'instruire les plaintes et réclamations qui lui seront adressées (*Data Act*, article 32).
[1151] COMMISSION EUROPÉENNE, *Proposition de règlement relatif à un marché intérieur des services numériques (Digital Services Act) et modifiant la directive 2000/31/CE*, 15 décembre 2020, COM (2020) 825 final.
[1152] COMMISSION EUROPÉENNE, *Proposition de règlement relatif aux marchés contestables et équitables dans le secteur numérique (Digital Markets Act)*, 15 décembre 2020, COM (2020) 842 final.

textes, le *Data Act* et l'*AI Act* notamment, sont encore à un stade du processus législatif qui laisse présager de nombreuses évolutions avant leur entrée en vigueur. Il est donc encore temps de lancer une réflexion d'ampleur quant à leur articulation avec les textes déjà promulgués. Une étude détaillée devrait être lancée par la Commission afin d'étudier les liens potentiels entre l'ensemble des textes qu'elle propose afin de s'assurer de leur compatibilité et de leur cohérence. Cette étude devrait également permettre à la Commission de justifier son choix de travailler sur cinq textes distincts, prévoyant chacun une gouvernance distincte, sur des sujets connexes. L'entrée en vigueur concomitante de l'ensemble de ces règlements pourrait en effet envoyer un très mauvais signal aux acteurs régulés, en Europe et à l'international. Ces derniers pourraient craindre une règlementation très contraignante, complexe et bureaucratique pouvant entraver leur activité, ce qui pourrait pousser les entreprises et les talents du numérique à quitter le marché européen en privilégiant les pays à la régulation plus souple.

537. **Transition.** Un tel phénomène pourrait également se produire au sein même du marché unique européen si les régulations européennes du numérique, et en particulier l'*AI Act*, ne s'articulaient pas correctement avec le droit national.

B/ Une articulation complexe avec le droit national

538. **L'effet préemptif d'une législation d'harmonisation totale.** Les deux chercheurs Michael Veale et Frederik Zuiderveen Borgesius ont été les premiers à alerter sur les risques liés au potentiel effet préemptif de l'*AI Act* sur les lois nationales[1153]. Leur analyse montre notamment que le choix d'une législation d'harmonisation totale, combinée avec une définition très large de l'IA dans le texte, devrait conduire à l'abrogation de toutes les règles nationales relatives aux algorithmes. Certains craignent alors que les États membres ne puissent adopter des standards de protection plus contraignants que ceux établis dans l'*AI Act*, par exemple, en renforçant les exigences relatives aux IA à haut risque ou en interdisant plus largement les

[1153] M. Veale, F. Zuiderveen Borgesius, « Demystifying the Draft EU Artificial Intelligence Act », *Computational Law Review International*, juillet 2021, vol. 22, n°4, 81–90.

systèmes d'identification biométrique[1154]. Nos recherches n'ont pas permis d'identifier de lois nationales applicables au secteur de l'électricité dont le champ d'application pourrait entrer en contradiction avec le contenu de l'*AI Act*. En revanche, si l'on sort du domaine de la loi pour entrer dans le domaine règlementaire, de nombreuses dispositions existantes sont déjà applicables aux systèmes d'IA les plus risqués. C'est le cas notamment des dispositions règlementaires applicables aux systèmes informatiques utilisés dans les activités d'importance vitale ou dans les centrales nucléaires par exemple. Il est important qu'une parfaite cohérence soit assurée entre ces corpus. Les règlementations sectorielles dans le secteur de l'énergie ont prouvé leur pertinence et doivent à ce titre conserver leur primauté, quitte à les amender à la marge pour intégrer les exigences spécifiques de l'*AI Act* au cas par cas.

539. **La nécessaire incorporation des exigences de l'*AI Act* dans les règlementations sectorielles.** On comprend facilement la raison pour laquelle la Commission européenne a porté une attention particulière aux systèmes d'IA utilisés dans les infrastructures critiques. En effet, il existe de nombreuses activités dans lesquelles la défaillance d'un système peut conduire à de terribles conséquences pour la santé des individus ou pour l'environnement. Ce fait, bien connu du droit, a déjà justifié la création d'un cadre juridique spécifique, dédié aux opérateurs d'importance vitale pour la vie de la Nation[1155]. Certains secteurs, tels que l'aviation, l'automobile, la santé ou encore l'énergie font ainsi partie des domaines dans lesquels la sécurité des systèmes d'information fait l'objet d'une attention particulière. Ce constat justifie que ces secteurs fassent l'objet de lourdes règlementations sectorielles. À première vue, la Commission européenne semble tenir compte de cet élément puisqu'elle exclut de son champ d'application toute une série de produits relevant de certaines législations sectorielles au niveau européen, listées à l'article 2 de l'*AI Act*[1156]. En plus de ces exclusions, l'intention des auteurs du texte semble limpide dans l'exposé des motifs. Ce dernier précise en effet que les

[1154] N.A. Smuha, E. Ahmed-Rengers, A. Harkens, *et al.*, « How the EU Can Achieve Legally Trustworthy AI: A Response to the European Commission's Proposal for an Artificial Intelligence Act », *SSRN Electronic Journal*, 5 août 2021, p. 41 ; M. Hildebrandt, « Commentary on the proposal for an EU AI Act of 21 April 2021 », *Réponse à la consultation publique de la Commission européenne sur l'AI Act*, 19 juillet 2021, Feedback n°F2662611.
[1155] ANSSI, « Protection des OIV en France », *Site institutionnel (blog)*, disponible en ligne : <https://www.ssi.gouv.fr/entreprise/protection-des-oiv/protection-des-oiv-en-france/>, consulté le 27 juin 2022.
[1156] Sont notamment visés par l'exclusion les produits soumis aux législations relatives à la sûreté de l'aviation civile, de l'automobile et du système ferroviaire (*AI Act*, article 2).

dispositions applicables aux systèmes d'IA à haut risque constituant des composants de sécurité de produits déjà régulés[1157] devront être intégrées dans la législation sectorielle existante pour *« assurer la cohérence, empêcher les doubles emplois et réduire au minimum les charges supplémentaires »*[1158]. La Commission ajoute vouloir créer une *« cohérence parfaite avec la législation sectorielle »*[1159] en incorporant les exigences de l'*AI Act* aux procédures d'examen de la conformité déjà existantes dans le cadre des législations sectorielles, notamment en matière de sécurité des produits[1160].

540. **L'absence de cohérence assurée avec les législations sectorielles au niveau national.** Toutefois, l'effort de cohérence ne vise que les « législations sectorielles au niveau de l'Union européenne », listées à l'article 2, et non l'ensemble des règlementations sectorielles qui pourraient exister au niveau national. Il n'est pas précisé si les exigences de conformité contenues dans l'*AI Act* doivent être intégrées aux processus de certification déjà établis dans les États membres si ces derniers ne relèvent pas des législations européennes listées à l'article 2 du texte. Cette absence est regrettable puisque les règlementations sectorielles visent à répondre à des risques très spécifiques aux secteurs d'activité concernés. Les autorités de régulation sectorielles sont les mieux placées pour établir, en coopération avec les acteurs régulés, les procédures de certification les plus adaptées. Il serait alors inopportun de les remettre en cause intégralement. Il convient plutôt d'étudier leur articulation et leur cohérence avec les nouvelles exigences pour les systèmes d'IA à haut risque. S'il est établi que les processus de certification existants sont plus exigeants que le contenu de l'*AI Act* alors ils ne doivent pas être modifiés. En revanche, si le standard de sécurité offert par la règlementation sectorielle au vu des risques spécifiques à l'utilisation de systèmes d'IA est plus faible, alors les exigences de l'*AI Act* doivent être intégrées au cas par cas aux procédures de certification existantes. Dans le secteur de l'électricité, les procédures de qualification des systèmes logiciels utilisés dans les centrales nucléaires pourraient être concernées. Bien qu'elles méritent d'être

[1157] Est principalement visée la directive du 3 décembre 2001 relative à la sécurité général des produits : *Directive 2001/95/CE du Parlement européen et du Conseil du 3 décembre 2001 relative à la sécurité générale des produits*, publiée au JOUE n° L 011 du 15 janvier 2002, pp. 4–17.
[1158] *AI Act*, exposé des motifs, p. 4.
[1159] *AI Act*, exposé des motifs, p. 12.
[1160] *AI Act*, exposé des motifs, p. 16.

adaptées aux spécificités des systèmes d'IA[1161], ces procédures présentent le plus haut niveau d'exigence et devraient donc conserver la primauté sur l'*AI Act* dans leur champ d'application.

541. **Proposition d'amendements visant à garantir une meilleure articulation de l'*AI Act* avec les règlementations sectorielles nationales.** Le défaut de prise en compte des règlementations sectorielles nationales dans le champ d'application de l'*AI Act* fait craindre un affaiblissement des standards de sécurité applicables aux systèmes d'IA utilisés dans des secteurs de haute technologie, critiques et hautement régulés. C'est pourquoi il serait utile de lancer une étude d'ampleur pour analyser les conséquences sectorielles du projet d'*AI Act*. Cette étude pourrait avoir pour objectifs d'identifier les règlementations sectorielles nationales applicables aux systèmes d'IA à haut risque et d'évaluer leur degré d'exigence en comparaison avec les dispositions de l'*AI Act*. Si le standard de sécurité offert par les règlementations sectorielles est plus important alors les systèmes d'IA soumis à ces exigences devraient être exclus du champ d'application de l'*AI Act*. À l'inverse, s'il apparaît que les exigences du projet de règlement sur l'IA sont plus strictes ou sont absentes des procédures existantes, alors les critères fixés par l'*AI Act* devront être intégrés aux règlementations sectorielles nationales. En ce sens, nous proposons deux amendements visant à clarifier l'articulation du projet de règlement avec les règlementations sectorielles nationales par l'introduction d'un nouveau considérant et l'ajout d'une exclusion relative à la règlementation en matière de sûreté nucléaire à l'article 2 de l'*AI Act* portant sur son champ d'application :

Considérant 86 (ajout)

(XX) L'objectif du présent règlement étant de garantir un haut niveau de sécurité pour les systèmes d'IA les plus risqués, les règlementations sectorielles relevant des États membres offrant un standard de sécurité plus élevé que le présent règlement ne devraient pas être remises en cause par celui-ci. Les exigences relatives aux systèmes d'IA à haut risque devront être intégrées aux procédures de certification et d'autorisation préexistantes dans les États membres. La Commission initiera, en collaboration avec les autorités compétentes dans les États membres, une étude de vaste ampleur, réalisée à la maille de chaque État membre et

[1161] Voir Supra, 190-207.

en amont de l'entrée en vigueur du présent règlement, pour identifier les règlementations sectorielles présentant un niveau d'exigences équivalent ou supérieur à celui de l'*AI Act*, d'une part, et celles ne comprenant pas de procédures de conformité équivalentes, d'autre part. En application du principe de subsidiarité, les premières ne seront pas affectées par l'entrée en vigueur du présent règlement tandis que les secondes devront être modifiées dans les meilleurs délais pour intégrer les exigences de conformité de l'*AI Act *dans les procédures existantes.

Article 2 :

[…]

2. Seul l'article 84 du présent règlement s'applique aux systèmes d'IA à haut risque qui sont des composants de sécurité de produits ou de systèmes ou qui constituent eux-mêmes des produits ou des systèmes et qui relèvent du champ d'application des actes suivants :

(a)règlement (CE) n° 300/2008;
[…]
(i) directive 2009/71/Euratom.
3. Le présent règlement ne s'applique pas aux systèmes d'IA soumis à des procédures de certification ou d'autorisation en vertu de règlementations sectorielles relevant des États membres lorsque ces dernières présentent un standard de sécurité plus élevé que les exigences relatives aux systèmes d'IA à haut risque énoncées au titre III, chapitre 2 du présent règlement. À défaut, les États membres doivent garantir que les procédures de certification ou d'autorisation prévues dans les règlementations sectorielles et applicables aux systèmes d'IA à haut risque intègrent les exigences énoncées au titre III, chapitre 2 du présent règlement.

[…]

542. **Conclusion de la Section 1 relative à la portée confuse du projet de règlement européen sur l'IA.** De manière générale, il semble que la Commission se presse dans l'adoption d'une règlementation transversale sans avoir pris le temps d'identifier précisément l'ensemble des conséquences qu'un tel texte d'harmonisation pourrait avoir tant au niveau national qu'européen. En effet, l'*AI Act* doit être parfaitement articulé avec les autres

législations européennes, en vigueur ou en construction. Pour ne pas dédoubler les procédures d'évaluation de la conformité pour les systèmes d'IA les plus risqués, il faut également que les exigences de l'*AI Act* soient dûment intégrées dans les procédures de certification et d'autorisation préexistantes. Ces procédures peuvent résulter non seulement des législations sectorielles européennes, qui sont bien exclues du projet de règlement, mais aussi des règlementations nationales, non mentionnées dans le texte. La question de la cohérence avec le corpus existant est d'autant plus importante que la Commission a fait le choix de retenir un champ d'application extrêmement large. Ce choix est discutable au vu de la potentielle inclusion de systèmes logiciels utilisés depuis des décennies, si bien que l'on peut légitimement s'interroger sur le véritable objectif poursuivi par la Commission : protéger les individus des risques nouveaux générés par les systèmes d'IA innovants ou faire de l'Union européenne l'acteur le plus rapide et le plus ambitieux en matière de régulation du numérique ?

543. **Transition.** Cette interrogation se confirme dans notre analyse du contenu des exigences relatives aux systèmes d'IA à haut risque qui sont à de nombreux égards déconnectés de la réalité. Leur adoption en l'état pourrait conduire à un frein considérable au développement de l'IA dans le secteur de l'électricité, notamment en raison des coûts de mise en conformité qu'ils engendreraient. L'*AI Act* relève à cet égard d'une ambition irréaliste.

Section 2 : Un texte à l'ambition démesurée

544. **Les objectifs ambitieux de l'*AI Act*.** La Commission européenne a présenté la création d'un cadre juridique pour l'IA comme l'un des trois piliers du concept d'IA « digne de confiance » qu'elle promeut depuis 2019[1162]. Si l'objectif affiché dans le projet de texte est la préservation du marché intérieur européen en luttant contre la fragmentation des législations nationales[1163], son contenu semble relever d'une ambition bien plus grande. Par son approche fondée sur les risques, son champ d'application large et le haut niveau d'exigence des dispositions relatives aux systèmes d'IA les plus risqués, le projet de règlement sur l'IA tente de réguler de façon transversale toutes les problématiques posées par le développement de l'IA. Toutefois, en souhaitant adresser un trop grand nombre de risques et en adoptant une démarche éminemment bureaucratique basée sur la mise en place de mécanismes de conformité, la Commission s'est déconnectée de ce qui aurait dû être la priorité dans la régulation d'une nouvelle technologie : la protection des droits et libertés des individus et la cohérence des exigences avec la réalité technique de la technologie régulée.

545. **Une ambition inachevée : l'insuffisance des garanties en matière de droits fondamentaux.** Les risques de l'IA pour les droits et libertés fondamentales ont fait l'objet de plusieurs études. Celles du Conseil de l'Europe[1164] et du Conseil d'État en France[1165] font aujourd'hui référence. Bien que la protection des droits fondamentaux ne soit pas l'objectif principal du projet d'*AI Act*, elle apparaît tout de même à quarante-sept reprises dans le corps du texte. La Commission européenne assure même dans son exposé des motifs que la proposition *« vise à garantir un niveau élevé de protection de ces droits fondamentaux »* avant de dresser une liste des droits protégés dans la Charte des droits fondamentaux de l'Union

[1162] GROUPE D'EXPERTS DE HAUT NIVEAU EN IA, *Lignes directrices du 8 avril 2019 pour une IA digne de confiance*, 8 avril 2019, p. 2.
[1163] Voir la base juridique de l'*AI Act* faisant référence à l'article 114 du TFUE, Considérant 2.
[1164] CONSEIL DE L'EUROPE, *Algorithmes et droits humains*, Etude menée par le comité d'experts sur les intermédiaires d'internet MSI-NET, 2017, DGI (2017)12, disponible en ligne : <https ://rm.coe.int/algorithms-and-human-rights-fr/1680795681>, consulté le 8 mars 2020.
[1165] CONSEIL D'ÉTAT, *Le numérique et les droits fondamentaux*, Etude annuelle 2014, La Documentation française, 64, 447 p.

européenne[1166] qu'elle juge renforcés par le texte : le droit à la dignité humaine, le respect de la vie privée et la protection des données à caractère personnel, la non-discrimination et bien d'autres. La longue liste ne fait l'objet d'aucune justification par la Commission et relève finalement plus de l'affichage que d'une réelle volonté de faire de l'*AI Act* un texte de protection des droits et libertés fondamentaux. Certains auteurs regrettent le choix d'une approche décrite comme éminemment technique et bureaucratique[1167] qui risque d'affaiblir les standards de protection des droits fondamentaux[1168]. Plus spécifiquement, il est également critiqué le fait que la Commission ait fait le choix de faire reposer la mise en œuvre exclusivement sur l'autoévaluation d'exigences techniques de conformité plutôt que d'impliquer les individus en leur reconnaissant des droits. De tels droits individuels auraient permis aux individus d'être mieux informés sur l'utilisation des systèmes d'IA, d'obtenir plus facilement réparation pour d'éventuels préjudices causés par l'utilisation d'un système d'IA ou plus généralement de signaler des non-conformités par un mécanisme de plainte[1169]. Nos propositions relatives à la nécessaire reconnaissance de droits individuels, réalisées dans le Titre précédent[1170], s'inscrivent dans la continuité de ces commentaires. Cette absence dans l'*AI Act* est d'autant plus surprenante que de tels droits avaient été reconnus aux personnes concernées par des traitements de données à caractère personnel dans le RGPD[1171] et sont en passe d'être créés pour les utilisateurs d'objets connectés dans le *Data Act*[1172]. Il s'agit d'une véritable lacune du projet d'*AI Act* qui exclut de fait les individus de la régulation d'une technologie amenée à transformer la société toute entière. Concernant le contenu des droits individuels à intégrer au projet de texte, nos précédents développements relatifs au mode de régulation le plus adapté[1173] ont permis de formuler plusieurs propositions concrètes qu'il aurait été pertinent d'intégrer au

[1166] *Charte des droits fondamentaux de l'Union européenne*, 2000/C 364/01, publiée au JOCE n°C364/3 du 18 décembre ; CONSEIL DE L'UE, *Conclusions de la présidence – La charte des droits fondamentaux dans le contexte de l'intelligence artificielle et du changement numérique*, 11481/20, 2020.
[1167] M. Hildebrandt, « Commentary on the proposal for an EU AI Act of 21 April 2021 », *Réponse à la consultation publique de la Commission européenne sur l'AI Act*, 19 juillet 2021, Feedback n°F2662611.
[1168] N.A. Smuha, E. Ahmed-Rengers, A. Harkens, *et al.*, *op. cit.*, p. 44.
[1169] *Ibid.*, p. 45 ; voir aussi dans le même sens M. Veale, F. Zuiderveen Borgesius, « Demystifying the Draft EU Artificial Intelligence Act », *Computational Law Review International*, juillet 2021, vol. 22, n°4, 98–100.
[1170] Voir Supra, 442-468.
[1171] RGPD, Chapitre 3, articles 12 à 23 relatifs aux droits des personnes concernées.
[1172] Voir par exemple : *Data Act*, articles 4 et 5 relatifs aux droits des utilisateurs d'objets connectés au regard des données générées par leur usage (droit d'accéder et droit de les partager à un tiers).
[1173] Voir Supra, 190-207.

projet d'*AI Act*. Toutefois, la Commission a fait le choix d'une régulation technique fondée sur des exigences de conformité pour les systèmes les plus risqués, lesquelles paraîssent inadaptées pour plusieurs raisons.

546. **Une ambition déconnectée de la réalité : l'irréalisme des exigences de conformité pour les systèmes d'IA à haut risque.** En effet, les systèmes d'IA considérés comme à haut risque doivent respecter une série d'exigences de conformité figurant aux articles 9 à 15 du projet d'*AI Act*. Ces exigences imposent la mise en place d'un système de gestion des risques (article 9), le respect de nombreux critères relatifs à la qualité des données utilisées (article 10), la rédaction d'une documentation technique exhaustive (article 11 et détail en annexe IV), la mise en œuvre de fonctionnalités d'enregistrement et de traçabilité du fonctionnement (article 12), la transparence (article 13), la garantie d'un contrôle humain (article 14) ainsi que la robustesse et la sécurité (article 15). Bien que ces exigences soient déjà particulièrement contraignantes, certains chercheurs considèrent malgré tout qu'elles devraient être clarifiées et renforcées[1174]. Si nous les rejoignons sur la nécessaire clarification, il nous semble néanmoins qu'un renforcement conduirait à alourdir encore plus la règlementation dont la mise en conformité sera déjà un véritable défi pour les entreprises. Nous nous contenterons donc ici d'étudier la pertinence et les éventuelles modifications à apporter aux exigences proposées par la Commission qui nous semblent irréalistes à de nombreux égards.

547. **Méthodologie employée.** Outre la connaissance des textes juridiques *stricto sensu*, l'analyse du contenu des exigences de conformité pour les systèmes d'IA à haut risque nécessite des compétences techniques pour évaluer leur faisabilité ainsi que des compétences légistiques pour proposer des amendements visant à rendre la lettre de l'*AI Act* plus adaptée à l'état de la technique. Pour mener à bien cette étude, nous avons réuni un groupe de travail pluridisciplinaire composé de chefs de projets, de développeurs, d'ingénieurs-chercheurs, d'un doctorant juriste et d'un spécialiste de légistique européenne dans le cadre de la *Task Force « IA »* du Groupe EDF. Notre première analyse de l'*AI Act* a permis de cibler précisément les articles du texte pouvant concerner les activités du secteur de l'électricité. Les experts techniques ont ensuite été appelés à classer les exigences pour les systèmes d'IA à haut risque

[1174] N.A. Smuha, E. Ahmed-Rengers, A. Harkens, *et al.*, *op. cit.*, p. 33.

dans l'une des quatre catégories suivantes : réalisable ; réalisable mais très coûteux ; difficilement réalisable ou nécessitant une clarification ; irréalisable. Enfin, la dernière étape de notre méthodologie consistait en la co-construction de propositions d'amendements visant à faire passer les exigences irréalisables, coûteuses ou imprécises dans la catégorie « réalisable ». Les résultats obtenus et les propositions résultant de cette démarche sont présentés dans la suite de nos développements.

548. **Plan.** Les exigences relatives aux systèmes d'IA à haut risque correspondent globalement aux bonnes pratiques dans la conduite de projets informatiques présentant des risques importants. Néanmoins la plupart d'entre elles vont nécessiter de lourds investissements pour créer des procédures spécifiques et risquent d'allonger la durée de la phase de conception. Si les dispositions sont très contraignantes et parfois difficiles à mettre en œuvre, elles apparaissent justifiées pour les systèmes les plus risqués. La suite de notre propos se concentrera sur les exigences pour lesquelles la mise en conformité présenterait un coût disproportionné (**§1**) ou serait tout bonnement irréalisable en l'état (**§2**).

§1 : Des exigences au coût disproportionné

549. **Plan.** Certaines dispositions encadrant la conception et l'utilisation des systèmes d'IA à haut risque pourraient entraîner des coûts de mise en conformité très importants, ce qui pourrait constituer un véritable frein à l'innovation. Elles concernent principalement la rédaction d'une documentation technique exhaustive (**A**) et l'enregistrement de données de fonctionnement aux fins de traçabilité (**B**).

A/ Des exigences relatives à la documentation technique

550. **L'exigence d'une documentation technique exhaustive pour les systèmes d'IA à haut risque.** L'article 11 du projet d'*AI Act* impose la rédaction d'une documentation technique avant la mise sur le marché ou la mise en service d'un système d'IA à haut risque. Elle doit être établie de manière à démontrer que le système satisfait aux exigences de conformité pour les IA à haut risque figurant aux articles 9 à 15 du règlement et doit permettre de fournir aux autorités de contrôle nationales toutes les informations nécessaires pour évaluer la conformité

avec ces exigences[1175]. L'établissement d'une documentation est une pratique courante dans tous les projets informatiques et justifiées avant tout par des considérations techniques. En effet, elle permet de retracer tout le cycle de développement d'un logiciel, les choix de conception ou encore les conditions d'usage pour lesquelles le produit a été conçu. Une documentation exhaustive permet notamment de résoudre plus rapidement des défaillances et de faciliter la preuve en cas de contentieux. Elle peut également comporter un grand nombre d'éléments utiles à l'information des personnes achetant des produits logiciels au titre de l'obligation d'information contractuelle.

Au premier abord, on pourrait donc s'étonner du fait que l'exigence de l'établissement d'une documentation technique dans l'*AI Act* soit limitée aux seuls systèmes d'IA à haut risque. En pratique une telle documentation technique continuera probablement d'être rédigée pour la majorité des systèmes logiciels.

Toutefois, la véritable raison d'être de cette disposition réside dans l'annexe IV du projet de règlement qui liste les éléments devant figurer dans la documentation établie au titre de l'article 11. La documentation technique pour les IA à haut risque devra d'abord contenir une description générale du système explicitant notamment sa destination, les personnes impliquées dans son développement, les modalités d'utilisation ou encore le matériel informatique sur lequel il est destiné à être exécuté[1176]. La documentation devra également contenir une description détaillée incluant, par exemple, les étapes suivies pour le développement du système d'IA, une explication de sa logique générale, une description de son architecture ou encore toutes les procédures mises en place pour garantir la qualité des données utilisées et pour tester le système[1177]. Enfin, l'annexe IV impose que la documentation contienne des informations relatives à la surveillance et au contrôle du système d'IA au cours de son utilisation, au système de gestion des risques mis en place par le fournisseur, aux modifications du système après sa mise en service ou encore aux normes techniques publiées au Journal Officiel de l'Union européenne qui ont été appliquées lors de la conception[1178].

[1175] *AI Act*, article 11, (1).
[1176] *AI Act*, annexe IV, (1).
[1177] *AI Act*, annexe IV, (2).
[1178] *AI Act*, annexe IV, (3) à (9)

Ainsi, le contenu de la documentation technique exigée au titre de l'article 11 de l'*AI Act* est particulièrement lourd. De manière générale, il vise à assurer la traçabilité de tout le cycle de vie des systèmes concernés, de leur conception à leur utilisation, et impose à leur fournisseur de justifier et documenter chaque choix effectué. On comprend aisément l'objectif de la Commission à travers cette exigence d'une documentation technique très complète, qui est de faciliter l'évaluation de la conformité des systèmes à haut risque avec les exigences du règlement. Toutefois, il est important de prendre en considération les conséquences que cette obligation pourrait avoir sur l'innovation.

551. **Les conséquences de l'exigence d'une documentation technique exhaustive.** Plus que l'exigence d'une documentation en elle-même, c'est plutôt l'exhaustivité de son contenu qui risque d'avoir un effet négatif sur l'innovation et *a fortiori* dans le secteur de l'électricité. En effet, l'effort de documentation va nécessairement peser sur les développeurs des systèmes d'IA à haut risque qui vont devoir tracer et justifier chaque étape du processus de conception ainsi que chaque choix qu'ils réalisent. Ce travail risque d'avoir trois effets négatifs pour la promotion de l'innovation en général, et encore plus dans le secteur de l'électricité.

Premièrement, l'effort de documentation va nécessairement engendrer des coûts financiers liés au temps passé par les développeurs à rédiger l'ensemble des éléments listés à l'annexe IV de l'*AI Act*.

Deuxièmement, cette tâche va également conduire à un allongement de la durée de la phase de conception puisqu'il faudra y inclure la rédaction de la documentation et sa validation par la comitologie interne à chaque entreprise.

Troisièmement, la rédaction de la documentation technique est une tâche administrative qui risque de démotiver les talents du numérique. Dans un contexte où les entreprises peinent à attirer les meilleurs profils face à la concurrence internationale, un ajout de contrainte bureaucratique dans l'Union européenne risquerait d'accélérer la fuite des cerveaux. Les talents disposant des compétences requises au développement de systèmes d'IA pourraient alors fuir les secteurs où ils seraient amenés à développer des IA considérées comme « à haut risque » ou se tourner vers des entreprises opérant en dehors du territoire de l'Union européenne. Le secteur de l'électricité pourrait pâtir de cette situation puisqu'une grande partie des cas d'usage explorés peuvent concerner des infrastructures critiques et donc être soumis aux dispositions de l'*AI Act*.

Ces considérations renforcent l'importance de restreindre la définition des systèmes d'IA à haut risque en annexe III pour que la contrainte règlementaire ne concerne que les IA dont le

niveau de risque est réellement critique. Aucun amendement concernant l'article 11 ou l'annexe IV ne sera proposé ici, puisque la question porte surtout sur leur champ d'application. En effet, leur contenu a été bien accueilli par les experts techniques sollicités dans le cadre du groupe de travail établi pour la présente Section car il est particulièrement clair malgré la lourdeur administrative qu'il implique.

552. **Transition.** Si l'obligation d'établir une documentation technique exhaustive risque d'engendrer des coûts disproportionnés liés au temps humain nécessaire à sa réalisation, les exigences relatives à l'enregistrement des données pendant le fonctionnement des systèmes d'IA à haut risque peuvent, elles, générer d'importants coûts financiers liés aux ressources informatiques requises pour s'y conformer.

B/ Des exigences relatives à l'enregistrement des données de fonctionnement

553. **L'obligation de mettre en place un mécanisme d'enregistrement automatique des données de fonctionnement.** L'article 12 du projet de règlement européen sur l'IA impose aux fournisseurs de systèmes d'IA à haut risque de prévoir des fonctionnalités permettant l'enregistrement automatique des évènements (aussi appelés « journaux » ou « logs ») pendant le fonctionnement des systèmes. Là encore, on pourrait s'interroger sur la raison pour laquelle la Commission européenne a souhaité créer cette obligation uniquement pour la catégorie à haut risque tant c'est une pratique courante pour tous les systèmes logiciels. La mise en conformité avec les dispositions de l'article 12 nécessitera la création, dès la conception, d'une architecture logicielle dédiée capable de tracer le fonctionnement du système d'IA et en mesure de les stocker sur une période de temps qui peut, selon l'interprétation que l'on fait des formulations retenues, s'avérer très longue.

554. **Une durée de conservation illimitée ?** Le deuxième alinéa de l'article 12 précise que ces fonctionnalités doivent permettre *« un degré de traçabilité du système d'IA tout au long de son cycle de vie qui soit adapté à sa destination »*[1179]. En revanche, aucune précision n'est

[1179] *AI Act*, article 12, (2).

apportée quant à la durée de conservation de ces données. Doit-on comprendre de la formulation du deuxième alinéa que les données doivent être conservée pendant l'entier cycle de vie ou que la durée de conservation doit être définie en fonction de la destination du système concerné ? La première option nécessiterait le stockage d'importantes quantités de données et l'engagement d'investissements financiers conséquents pour les entreprises. De plus, il est peu probable que des décisions prises par un système d'IA au cours de son fonctionnement se révèlent avoir des conséquences négatives plusieurs mois ou années plus tard. La conservation des données de fonctionnement devrait donc, selon nous, être limitée dans le temps (avec un maximum de six ou douze mois par exemple) et être définie au cas par cas selon la destination du système par les fournisseurs.

555. **La nécessaire rationalisation des données de fonctionnement à tracer.** Le projet d'*AI Act* dresse également la liste des données de fonctionnement que les fournisseurs d'IA à haut risque devront enregistrer et conserver. Elles doivent fournir au minimum *« la période de chaque utilisation »* (date et heure d'utilisation), la *« base de données de référence utilisée par le système pour vérifier les données d'entrée »*, les *« données d'entrée »* ainsi que *« l'identification des personnes physiques participant à la vérification des résultats »*[1180]. Ces données permettent effectivement un haut degré de traçabilité du fonctionnement d'un système. Suivant l'application concernée, cela peut toutefois représenter de grandes quantités de données. Il est donc essentiel de limiter leur durée de conservation dans le temps sans quoi la mise en conformité conduirait à des coûts de développement et de stockage des données dissuasifs pour les fournisseurs. Outre le coût disproportionné de la mise en conformité, la fourniture de la base de données de référence peut s'avérer difficile dans le cas où un système a été conçu à partir de plusieurs bases de données. L'identification par le fournisseur de l'ensemble des bases utilisées par le système pour vérifier les données d'entrée s'avère à ce titre plus raisonnable.

[1180] *AI Act*, article 12, (4).

556. **Proposition d'amendements.** Afin de rationaliser le contenu des données à conserver aux fins de traçabilité et leur durée de conservation, l'article 12 pourrait être amendé ainsi :

Article 12
Enregistrement

1. La conception et le développement des systèmes d'IA à haut risque prévoient des fonctionnalités permettant l'enregistrement automatique des événements («journaux») pendant le fonctionnement de ces systèmes. Ces fonctionnalités d'enregistrement sont conformes à des normes ou à des spécifications communes reconnues.

2. Les fonctionnalités d'enregistrement garantissent un degré de traçabilité du fonctionnement du système d'IA ~~tout au long de son cycle de vie~~ qui soit adapté à la destination du système. ***La durée de conservation des données collectées au titre du présent alinéa doit être limitée dans le temps et ne doit pas excéder 6 (six) mois.***

3. En particulier, ces fonctionnalités permettent de surveiller le fonctionnement du système d'IA à haut risque dans l'éventualité de situations ayant pour effet que l'IA présente un risque au sens de l'article 65, paragraphe 1, ou entraînant une modification substantielle, et facilitent la surveillance après commercialisation visée à l'article 61.

4. Pour les systèmes d'IA à haut risque visés à l'annexe III, paragraphe 1, point a), les fonctionnalités d'enregistrement fournissent, au minimum:

(a) l'enregistrement de la période de chaque utilisation du système (date et heure de début et de fin pour chaque utilisation);

(b) ~~la base de données de référence utilisée~~ ***les bases de données de référence utilisées*** *par le système pour vérifier les données d'entrée;*

(c) les données d'entrée pour lesquelles la recherche a abouti à une correspondance;

(d) l'identification des personnes physiques participant à la vérification des résultats, visées à l'article 14, paragraphe 5.

557. **Transition.** Tandis que les exigences relatives à la documentation technique et à l'enregistrement des données de fonctionnement risqueraient de générer un coût disproportionné de mise en conformité, d'autres dispositions contraignantes pour les systèmes d'IA à haut risque sont tout simplement irréalisables.

§2 : Des exigences irréalisables

558. **Plan.** L'étude des exigences relatives aux systèmes d'IA à haut risque menée dans le cadre du groupe de travail pluridisciplinaire établi spécifiquement a révélé que plusieurs dispositions de l'*AI Act*, en leur formulation initiale, n'étaient pas réalistes d'un point de vue technique. Il ne s'agit pas de remettre en cause l'existence de ces dispositions qui sont justifiées et correspondent aux bonnes pratiques dans le domaine de l'IA, mais plutôt de corriger leur formulation qui les rend impraticables en l'état. Ces dispositions concernent principalement la gouvernance des données utilisées pour concevoir les systèmes d'IA à haut risque (**A**), la garantie d'un contrôle humain sur le fonctionnement du système (**B**) ainsi que la robustesse et la cybersécurité (**C**).

A/ Des exigences relatives à la qualité des données

559. **L'encadrement strict de la qualité des données utilisées dans la conception des systèmes d'IA à haut risque.** L'article 10 du projet d'*AI Act* établit une liste de critères de qualité auxquels les systèmes d'IA à haut risque faisant appel à des techniques d'apprentissage automatique doivent satisfaire. Ces critères visent spécifiquement *« les jeux de données d'entraînement, de validation et de test »*[1181]. Ils consistent, d'abord, en l'adoption de pratiques appropriées relatives aux choix de conception, à la collecte des données ou aux opérations préalables réalisées sur ces dernières[1182]. Le texte impose, ensuite, la réalisation d'une évaluation de la disponibilité et de l'adéquation des jeux de données nécessaires ainsi que d'un examen visant à détecter d'éventuels biais dans les données[1183]. De plus, le troisième paragraphe de l'article 10 dispose que les jeux d'entraînement, de validation et de test doivent être impérativement *« exempts d'erreurs et complets »*[1184] et présenter des *« propriétés statistiques appropriées »*[1185] en ce qui concerne la population auprès de laquelle le système d'IA est destiné à être utilisé. En outre, l'article autorise – et c'est heureux – le traitement de

[1181] *AI Act*, article 10, (1).
[1182] *AI Act*, article 10, (2), a) à d).
[1183] *AI Act*, article 10, (2), e) à g).
[1184] *AI Act*, article 10, (3).
[1185] *Ibid.*

données considérées comme sensibles au sens du RGPD lorsque cela s'avère nécessaire pour détecter les biais dans les données. Enfin, il est précisé que des bonnes pratiques similaires à celles décrites précédemment doivent être mises en œuvre pour les systèmes d'IA qui ne sont pas fondés sur l'apprentissage automatique, sans plus de précisions quant à leur nature.

560. **Des formulations hasardeuses rendant les dispositions impraticables.** Plusieurs critiques peuvent être faites au regard des formulations retenues dans cet article qui est l'un des plus importants du projet de règlement puisqu'il tente d'harmoniser les bonnes pratiques en matière de gouvernance des données utilisées dans la conception des systèmes d'IA.

561. **Une imprécision quant à l'applicabilité de l'article 10 aux systèmes d'IA symbolique.** D'une part, on constate dès son premier paragraphe que l'article 10 se focalise sur les systèmes d'IA conçus à partir de techniques d'apprentissage sur des données. Est-ce à dire que les systèmes d'IA symboliques, fondés sur l'application de règles déterministes, ne sont soumis à aucun critère de qualité ? La Commission semble vouloir répondre à cette question dans le dernier paragraphe disposant que *« des pratiques appropriées en matière de gouvernance et de gestion des données s'appliquent au développement de systèmes d'IA à haut risque autres que ceux qui font appel à des techniques impliquant l'entraînement de modèles afin de garantir que ces systèmes d'IA à haut risque sont conformes au paragraphe 2 »*[1186]. Or, le paragraphe 2 auquel il est fait référence contient des critères de qualité relatifs aux jeux de données d'entraînement, de validation et de test, lesquels n'existent pas forcément dans le cas de la conception d'un système d'IA déterministe. La formulation retenue est donc incohérente en l'état. Sans précision sur la nature des pratiques à appliquer aux IA symboliques, il serait plus judicieux de retirer le paragraphe 6 pour que l'article 10 soit spécifique aux systèmes d'IA dont le développement nécessite l'apprentissage sur des données[1187].

562. **Une obligation de résultat irréaliste techniquement.** D'autre part, l'obligation faite à l'article 10 paragraphe 3 de garantir que les jeux de données d'entraînement, de validation et de test sont *« exempts d'erreurs et complets »* apparaît irréalisable aux yeux des experts sollicités dans le cadre de notre étude. La formulation impérative retenue par la Commission

[1186] *AI Act*, article 10, (6).
[1187] Voir dans le même sens : M. Hildebrandt, *op. cit.*, p. 5.

fait de cette disposition une obligation de résultat, ce qui est difficilement compatible avec l'état actuel de la technique. Par ailleurs, le paragraphe 2 impose déjà la mise en place de mesures appropriées pour lutter contre la présence de biais et d'erreurs ainsi que pour garantir que les données utilisées sont représentatives, si bien que la mention des erreurs et de la complétude des données au paragraphe 3 s'avère redondant. Il est donc légitime que la doctrine académique[1188] et les professionnels[1189] questionnent la pertinence de conserver cette mention. Puisqu'elle est déjà couverte par les dispositions du paragraphe 2, sa suppression au paragraphe 3 n'affectera pas la sécurité des systèmes d'IA à haut risque.

563. **Proposition d'amendements.** Pour répondre à ces deux critiques, les paragraphes 3 et 6 de l'article 10 pourraient être amendés comme suit :

Article 10
Données et gouvernance des données

[…]

*3. Les jeux de données d'entraînement, de validation et de test sont pertinents**, et** représentatifs, **~~exempts d'erreurs et complets~~**. Ils possèdent les propriétés statistiques appropriées, y compris, le cas échéant, en ce qui concerne les personnes ou groupes de personnes à l'égard desquels le système d'IA à haut risque est destiné à être utilisé. Ces caractéristiques des jeux de données peuvent être présentes au niveau des jeux de données pris individuellement ou d'une combinaison de ceux-ci.*

[…]

*6. **~~Des pratiques appropriées en matière de gouvernance et de gestion des données s'appliquent au développement de systèmes d'IA à haut risque autres que ceux qui font appel à des techniques impliquant l'entraînement de modèles afin de garantir que ces systèmes d'IA à haut risque sont conformes au paragraphe 2.~~***

564. **Transition.** Comme pour les dispositions de l'article 10, les exigences de l'article 14 relatives au contrôle humain du fonctionnement des systèmes d'IA à haut risque sont bienvenues et nécessaires, mais leur formulation les rend irréalistes en l'état.

[1188] *Ibid*.
[1189] GOOGLE, *Consultation on the EU AI Act Proposal : Google's submission*, 15 juillet 2021, pp. 3–6, disponible en ligne : <https://ec.europa.eu/info/law/better-regulation/have-your-say/initiatives/12527-Artificial-intelligence-ethical-and-legal-requirements/F2662492_en>, consulté le 19 juin 2022.

B/ Des exigences relatives au contrôle humain

565. **L'exigence bienvenue d'un contrôle humain effectif sur le fonctionnement des systèmes d'IA à haut risque.** L'article 14 de l'*AI Act* impose aux fournisseurs d'IA à haut risque de concevoir leurs systèmes de façon à pouvoir garantir un contrôle effectif par des personnes physiques pendant la période d'utilisation[1190]. Les mesures de contrôle humain peuvent être intégrées dans les fonctionnalités propres du système ou consister en des mesures devant être mises en œuvre par l'utilisateur sous réserve qu'elles figurent dans les instructions d'usage données par le fournisseur[1191]. Ces mesures doivent notamment donner aux personnes physiques chargées du contrôle les possibilités de surveiller le fonctionnement du système, de détecter et traiter rapidement les dysfonctionnements, de pouvoir passer outre ou inverser la décision du système et d'arrêter le système si nécessaire[1192]. En somme, l'article 14 vise à proscrire le fonctionnement complètement autonome des systèmes d'IA à haut risque. Cet apport est bienvenu et justifié pour les systèmes les plus risqués.

L'esprit et la lettre de cette disposition sont en phase avec d'autres règles juridiques européennes. D'une part, on perçoit des similarités avec l'exigence d'un « contrôle humain effectif et non artificiel » sur les traitements de données personnelles constituant des décisions automatisées au sens du RGPD, conformément aux lignes directrices du G29[1193] et à la doctrine de la CNIL[1194]. D'autre part, la volonté de la Commission européenne d'interdire le fonctionnement de systèmes d'IA risqués sans contrôle humain est conforme à la position retenue par la Cour de Justice de l'Union européenne (CJUE) dans un arrêt du 21 juin 2022 relatif à l'exploitation des données sur les passagers des transporteurs aériens aux fins de lutte

[1190] *AI Act*, article 14, (1).
[1191] *AI Act*, article 14, (3).
[1192] *AI Act*, article 14, (4).
[1193] GROUPE DE TRAVAIL « ARTICLE 29 » SUR LA PROTECTION DES DONNÉES, *Lignes directrices relatives à la prise de décision individuelle automatisée et au profilage aux fins du règlement (UE) 2016/679*, 3 octobre 2017, n°WP251.
[1194] CNIL, *Délibération n°2017-299 du 30 novembre 2017 portant avis sur un projet de loi d'adaptation au droit de l'Union européenne de la loi n°78-17 du 6 janvier 1978*, p. 28.

contre le terrorisme[1195]. En effet, dans cet arrêt, la Cour a estimé que les autorités répressives ne pouvaient utiliser de systèmes d'IA fondés sur l'apprentissage automatique sans supervision humaine afin de croiser les fichiers des compagnies aériennes avec les fichiers de personnes surveillées pour des raisons de lutte contre le terrorisme[1196]. Un tel croisement de données ne peut intervenir qu'en présence d'un système fonctionnant avec des paramètres préétablis ou sous la supervision de personnes physiques. On retrouve donc dans la jurisprudence de la CJUE, dans les autres textes européens et dans la proposition d'*AI Act* la même volonté d'éviter que des décisions susceptibles d'avoir des conséquences négatives pour les droits et libertés des individus ne soient prises par des systèmes fonctionnant de façon complètement autonome. À cet égard, l'article 14 de l'*AI Act* apparaît cohérent avec la politique de l'Union et la protection des droits fondamentaux[1197]. Néanmoins, la formulation retenue par la Commission laisse une nouvelle fois à désirer.

566. **L'exigence irréaliste d'une compréhension totale du système d'IA.** Dans sa traduction française officielle, le paragraphe 4 de l'article 14 exige des personnes physiques chargées d'effectuer le contrôle humain *« d'appréhender totalement les capacités et limites du système d'IA à haut risque »*[1198]. Le terme « appréhender » est la traduction du verbe *« understand »* dans la version anglaise qui peut être traduit plus communément par le mot « comprendre ». Ce choix de traduction est surprenant car l'appréhension et la compréhension ont des portées différentes. Si l'appréhension des capacités de l'IA reste accessible, leur compréhension totale relève d'une très grande expertise.

Quelle que soit la traduction retenue dans la version finale, exiger une appréhension ou une compréhension « totale » des capacités et limites du système d'IA à haut risque apparaît irréaliste et injustifié. D'une part, cette exigence semble irréaliste car certaines des spécificités des systèmes d'IA rendent très difficiles leur compréhension. Il s'agit en effet de systèmes ayant recours à des techniques souvent fondées sur des principes mathématiques complexes et dont

[1195] CJUE, aff. C-817/19, 21 juin 2022, *Ligue des droits humains* ; CJUE, *Press release n°105/22*, Case C-817/19, Ligue des droits humains, 21 juin 2022.
[1196] CJUE, aff. C-817/19, *op. cit.*, 194.
[1197] Voir dans le même sens : M. Hildebrandt, *op. cit.*, p. 5.
[1198] *AI Act*, article 14, (4), a).

le fonctionnement peut s'avérer opaque[1199]. Demander une compréhension totale d'un système d'IA revient à exclure des personnes physiques pouvant effectuer un contrôle effectif la majorité des individus qui ne disposent pas d'une formation supérieure en mathématiques et en IA. D'autre part, une telle exigence nous semble également injustifiée car il n'est pas nécessaire de disposer d'une compréhension complète de toutes les capacités et limites d'un système d'IA pour pouvoir surveiller son fonctionnement et détecter d'éventuelles anomalies. D'après les experts sollicités dans le cadre de notre étude, une formation adéquate sur les fonctionnalités principales, la logique sous-jacente de l'algorithme, les performances normales et les paramètres à contrôler suffirait pour effectuer des missions de contrôle du fonctionnement d'un système d'IA. Ainsi, exiger l'appréhension des principales capacités et limites du système d'IA semblerait plus réaliste, d'autant que les autres dispositions de l'article détaillent les informations que devraient maîtriser les personnes physiques chargées du contrôle humain.

567. **Proposition d'amendement.** La majorité de l'article 14 ne suscite pas de difficultés majeures et est parfaitement justifiée au regard de la nécessité d'exercer un contrôle humain sur la technologie. Pour parfaire la formulation de l'article afin qu'il soit pleinement réaliste, son paragraphe 4 pourrait être amendé comme suit :

Article 14
Contrôle humain

[…]

4. Les mesures prévues au paragraphe 3 donnent aux personnes chargées d'effectuer un contrôle humain, en fonction des circonstances, la possibilité:

(a)d'appréhender ~~***totalement***~~ *les* ***principales*** *capacités et les limites du système d'IA à haut risque et d'être en mesure de surveiller correctement son fonctionnement, afin de pouvoir détecter et traiter dès que possible les signes d'anomalies, de dysfonctionnements et de performances inattendues;*

(b)d'avoir conscience d'une éventuelle tendance à se fier automatiquement ou excessivement aux résultats produits par un système d'IA à haut risque («biais d'automatisation»), en particulier pour les systèmes d'IA à haut risque utilisés pour fournir des informations ou des recommandations concernant les décisions à prendre par des personnes physiques;

[1199] R. Yampolskiy, « Unexplainability and Incomprehensibility of AI », *Journal of Artificial Intelligence and Consciousness*, juillet 2020, vol. 7, n°2, 1-15.

(c) d'être en mesure d'interpréter correctement les résultats du système d'IA à haut risque, compte tenu notamment des caractéristiques du système et des outils et méthodes d'interprétation disponibles;

(d) d'être en mesure de décider, dans une situation particulière, de ne pas utiliser le système d'IA à haut risque ou de négliger, passer outre ou inverser le résultat fourni par ce système;

(e) d'être capable d'intervenir sur le fonctionnement du système d'IA à haut risque ou d'interrompre ce fonctionnement au moyen d'un bouton d'arrêt ou d'une procédure similaire.

[…]

568. **Transition.** En plus de la qualité des données et du contrôle humain, la dernière catégorie d'exigences de conformité dont le contenu est ressorti comme irréaliste dans l'étude menée dans le cadre de la thèse concerne la robustesse et la cybersécurité des systèmes d'IA à haut risque.

C/ Des exigences relatives à la robustesse et à la cybersécurité

569. **Des exigences pour garantir un niveau élevé d'exactitude, de robustesse et de cybersécurité des systèmes d'IA à haut risque.** Le premier paragraphe de l'article 15 du projet de règlement sur l'IA impose aux fournisseurs de concevoir leurs systèmes de façon à atteindre *« un niveau approprié d'exactitude, de robustesse et de cybersécurité »*[1200]. Les paragraphes qui suivent précisent ce qu'il faut entendre par « niveau approprié ». Les fournisseurs doivent par exemple justifier, dans les instructions d'utilisation, des métriques pertinentes sur la performance et l'exactitude du système[1201]. Ils doivent également garantir que les systèmes sont résilients face aux erreurs, aux défaillances ou aux tentatives d'attaques informatiques[1202]. Ces obligations sont plutôt bien accueillies par la doctrine[1203] et correspondent dans l'ensemble aux bonnes pratiques développées par les entreprises du

[1200] *AI Act*, article 15, (1).
[1201] *AI Act*, article 15, (2).
[1202] *AI Act*, article 15, (3) à (7).
[1203] M. Hildebrandt, *op. cit.*, p. 6.

numérique[1204]. Toutefois, si l'on rentre dans le détail, leur contenu apparaît inadapté sous deux aspects : le premier tient au fait que le texte se concentre exclusivement sur la phase de développement des systèmes sans prendre en compte suffisamment les opérations de maintenance au cours du cycle de vie, le deuxième au fait que la Commission ait adopté des formulations impératives conduisant à des obligations de résultats pour les fournisseurs malgré le faible niveau de maturité de la technologie.

570. **La regrettable focalisation des exigences sur la phase de conception des systèmes d'IA.** Dès son premier paragraphe, l'article 15 vise exclusivement *« la conception et le développement des systèmes d'IA à haut risque »*. Pourtant, dans le domaine de la sécurité informatique, il est de notoriété publique que la maintenance tout au long du cycle de vie des systèmes joue un rôle primordial dans leur robustesse et leur résilience face aux potentielles attaques[1205]. Il serait pertinent d'intégrer les opérations de maintenance dans les moyens pour parvenir à un niveau approprié de cybersécurité plutôt que de faire peser cette charge uniquement sur la phase de conception.

571. **Le nécessaire basculement d'obligations de résultat vers des obligations de moyens.** Dans cet article, la Commission européenne a fait le choix de formulations impératives, créant de fait des obligations de résultat à plusieurs reprises. Ainsi, par exemple, les systèmes d'IA *« font preuve de résilience en cas d'erreurs »*[1206] et *« résistent aux tentatives de tiers non autorisés »*[1207]. Les experts techniques sollicités dans le cadre de notre étude sont formels sur leur incapacité à garantir avec un degré de certitude absolu la résilience des systèmes développés. Bien souvent, les failles de sécurité sont découvertes au fil du temps comme en témoigne l'expérience dans le domaine des logiciels. Les fournisseurs ne peuvent en pratique que s'engager à prendre toutes les mesures techniques pour assurer la robustesse des systèmes conformément à l'état de l'art ou à suivre des normes techniques harmonisées. Pour être

[1204] Voir par exemple : MICROSOFT, *Microsoft Responsible AI Standard v2*, document officiel, juin 2022, p. 21, « Reliability & Safety Goals », disponible en ligne : <https://query.prod.cms.rt.microsoft.com/cms/api/am/binary/RE4ZPmV>, consulté le 7 juillet 2022.
[1205] Voir par exemple : ANSSI, *Maîtriser la SSI pour les systèmes industriels : la cybersécurité des systèmes industriels*, guide pratique, 2012, p.22, disponible en ligne : <https://www.ssi.gouv.fr/uploads/IMG/pdf/2012-06-19_CP-Guide_SCADA.pdf>, consulté le 7 juillet 2022.
[1206] *AI Act*, article 15, (3).
[1207] *AI Act*, article 15, (6).

pleinement applicable dans la pratique, l'article 15 devrait être amendé pour passer d'une obligation de résultat à une obligation de moyens visant à atteindre un niveau approprié d'exactitude, de robustesse et de cybersécurité.

572. **Proposition d'amendements.** Pour pallier les lacunes résultant des formulations retenues par la Commission européenne dans l'article 15, ses paragraphes 1, 3 et 6 pourraient être amendés comme suit :

Article 15
Exactitude, robustesse et cybersécurité

1. La conception et le développement des systèmes d'IA à haut risque sont tels qu'ils leur permettent, compte tenu de leur destination, d'atteindre un niveau approprié d'exactitude, de robustesse et de cybersécurité, et de fonctionner de manière cohérente à cet égard tout au long de leur cycle de vie, ***y compris grâce à la mise en œuvre des opérations de maintenance appropriées.***

[…]

3. ***Les systèmes d'IA à haut risque sont conçus, développés et maintenus de manière à faire preuve d'un niveau approprié de résilience*** ~~***font preuve de résilience***~~ *en cas d'erreurs, de défaillances ou d'incohérences pouvant survenir au sein des systèmes eux-mêmes ou de l'environnement dans lequel ils fonctionnent, notamment en raison de leur interaction avec des personnes physiques ou d'autres systèmes.*

[…]

6. ***Les systèmes d'IA à haut risque sont conçus, développés et maintenus, conformément à l'état de l'art, de façon à résister*** ~~***résistent***~~ *aux tentatives de tiers non autorisés visant à modifier leur utilisation ou leurs performances en exploitant les vulnérabilités du système.*

[…]

573. **Conclusion de la Section 2 relative à l'inadaptation du contenu du projet de règlement européen sur l'IA.** La présente Section a permis une étude du contenu des exigences de conformité encadrant la conception, l'utilisation et la fourniture de systèmes d'IA à haut risque. De manière générale, ces exigences sont très contraignantes mais elles correspondent pour la majorité aux bonnes pratiques suivies par les acteurs du secteur de l'électricité. Toutefois, lorsque l'on s'intéresse en détail aux formulations retenues par le législateur, on peut constater plusieurs lacunes.

D'une part, certaines exigences nécessiteraient la mise en œuvre de mesures très coûteuses et pourraient détourner les entreprises des applications considérées comme à haut risque. Il s'agit principalement des exigences relatives à la rédaction d'une documentation technique exhaustive conduisant à la bureaucratisation du développement de systèmes d'IA à haut risque,

ainsi qu'à l'enregistrement des données de fonctionnement, nécessitant en l'état le stockage sans limite dans le temps d'importantes quantités de données sans réelle nécessité.

D'autre part, certaines dispositions applicables aux systèmes d'IA à haut risque s'avèrent tout simplement irréalisables telles qu'elles sont actuellement formulées. L'*AI Act* exige ainsi l'absence absolue d'erreurs dans les données utilisées pour l'entraînement des modèles d'IA, une connaissance « totale » des capacités et limites des systèmes de la part des personnes physiques en charge de leur contrôle ou encore la garantie que l'IA ne peut faire l'objet d'aucune attaque par un tiers.

Ces constats ont permis la formulation de plusieurs propositions d'amendements dont la plupart ont été portés devant les rapporteurs des différentes commissions du Parlement européen compétentes sur l'*AI Act*. Leur prise en compte est importante puisque sans ces amendements la mise en conformité des entreprises risque d'être démesurément coûteuse, voire impossible.

574. **Conclusion du Chapitre 1 relatif à la création d'un cadre juridique européen dédié à l'IA.** La création d'un cadre juridique spécifique pour répondre aux risques nouveaux qui ne seraient pas couverts par le corpus législatif existant s'avère nécessaire. En ce sens, la proposition de la Commission européenne de créer un règlement européen sur l'IA était, sur le papier, bienvenue et prometteuse. Néanmoins, le présent Chapitre a permis de mettre en lumière un grand nombre de lacunes au regard de son application aux systèmes d'IA destinés à être utilisés dans le secteur de l'électricité. Les critiques, non exhaustives, adressées à l'*AI Act* concernent à la fois sa portée trop confuse, en partie du fait des définitions larges retenues par la Commission, et son contenu inadapté à la réalité technique. S'il venait à être publié en l'état, ce qui n'est pas exclu, il risque de produire un effet contreproductif sur le développement de l'IA et d'envoyer un très mauvais signal aux entreprises souhaitant commercialiser des systèmes d'IA considérés comme à haut risque. Dans le secteur de l'électricité où l'innovation n'est déjà pas la priorité des acteurs historiques, l'*AI Act* viendrait désinciter les entreprises à investir dans le développement de l'IA en leur assurant des coûts de mise en conformité prohibitifs. Il est important de contrebalancer ce signal négatif en amendant le texte avant son entrée en vigueur afin de montrer aux différents secteurs d'activité que leurs préoccupations et leurs retours sont pris en compte.

De manière plus générale, on peut légitimement se demander si l'*AI Act* ne relève pas plus du marqueur politique que d'une réelle nécessité juridique pour l'Union européenne. Il semble s'agir avant tout d'un effet d'affichage consistant à vouloir règlementer à tout prix cette technologie afin de se poser en leader sur la scène internationale[1208].

Quoi qu'il en soit, les caractéristiques spécifiques des systèmes d'IA et leur développement rapide dans tous les secteurs d'activité appellent à une vigilance accrue. Il est important d'harmoniser les pratiques dans le domaine de l'IA pour que son développement soit compatible avec la règle de droit. Si la proposition de la Commission n'est pas parfaite, elle a le mérite de poser une première pierre à l'édifice de l'objectif ambitieux visant à promouvoir une IA « digne de confiance » sur le sol européen et peut-être à terme dans le monde entier.

575. **Transition.** Toutefois, les lignes directrices pour une IA « digne de confiance » mentionnent un risque qui n'est pas adressé dans l'*AI Act*. En effet, pour le groupe d'experts de haut niveau en IA établi par la Commission européenne en 2018, l'IA doit être au service du *« bien-être environnemental et du développement durable »*[1209]. Les experts estiment qu'il convient à cet effet *« d'encourager les mesures permettant de garantir que l'ensemble de la chaîne d'approvisionnement du système d'IA respecte l'environnement »*[1210]. Pourtant, la problématique de l'empreinte environnementale de l'IA est absente du projet de règlement européen, comme l'a mis en évidence la commission de l'environnement du Parlement européen[1211]. L'importance de ce risque appelle à une régulation, mais la complexité de la réponse à apporter justifie son traitement de façon indépendante dans la présente thèse. À cet effet, le Chapitre suivant propose notamment un corpus de grands principes directeurs pouvant guider la régulation environnementale de l'IA.

[1208] N. Martial-Braz, « L'intelligence artificielle bientôt régulée. L'incidence du AI Act dans le secteur financier », *Revue de Droit bancaire et financier*, mai 2021, n°3, comm. 81.
[1209] GROUPE D'EXPERTS DE HAUT NIVEAU EN IA, *Lignes directrices du 8 avril 2019 pour une IA digne de confiance*, 8 avril 2019, p. 6 et p. 24.
[1210] *Ibid.*, p. 24.
[1211] COMMISSION DE L'ENVIRONNEMENT, DE LA SANTÉ PUBLIQUE ET DE LA SÉCURITÉ ALIMENTAIRE (ENVI), *Avis sur la proposition de règlement du Parlement européen et du Conseil établissant des règles harmonisées concernant l'intelligence artificielle (législation sur l'intelligence artificielle) et modifiant certains actes législatifs de l'Union (COM(2021)0206 – C9-0146/2021 – 2021/0106(COD))*, p. 3 : *« La rapporteure pour avis craint que la législation sur l'intelligence artificielle ne protège pas suffisamment l'environnement »*.

Chapitre 2 : L'indispensable régulation environnementale du développement de l'IA

577. **La prise de conscience sur les conséquences écologiques du développement de l'IA.** Si les initiatives de régulation portant sur l'éthique du développement de l'IA se multiplient, la question de son empreinte environnementale est bien souvent reléguée au second plan. Les différentes chartes éthiques ou prises de position politiques témoignent pourtant d'une prise de conscience sur cet enjeu.

578. **Une priorité affichée à l'échelle internationale.** Au niveau international d'abord, l'enjeu environnemental du développement de l'IA est présent à la fois dans la Recommandation du Conseil de l'OCDE sur l'Intelligence Artificielle du 22 mai 2019[1212], adoptée par 42 pays[1213], et dans le projet de Recommandation sur l'éthique de l'intelligence artificielle proposé par l'UNESCO en 2020[1214]. Ce dernier contient un principe invitant à mettre les services d'IA au service de la prospérité de l'environnement et des écosystèmes[1215]. À date, aucun instrument normatif contraignant n'a suivi l'adoption de ces principes.

579. **Un enjeu identifié à l'échelle européenne mais absent des initiatives de régulation.** Au niveau européen, ensuite, le livre blanc sur l'IA de la Commission européenne du 19 février 2020[1216] ou la résolution du Parlement européen relative aux aspects éthiques de l'intelligence

---

[1212] OCDE, *Recommandation du Conseil de l'OCDE sur l'Intelligence Artificielle*, 22 mai 2019, OECD/LEGAL/0449, p. 7.

[1213] OCDE, « Quarante-deux pays adoptent les nouveaux Principes de l'OCDE sur l'intelligence artificielle », *Site institutionnel de l'OCDE*, 22 mai 2019, disponible en ligne : <https://www.oecd.org/fr/presse/quarante-deux-pays-adoptent-les-nouveaux-principes-de-l-ocde-sur-l-intelligence-artificielle.htm>, consulté le 29 avril 2021.

[1214] UNESCO, *Avant-projet de Recommandation sur l'éthique de l'intelligence artificielle*, *Bibliothèque numérique de l'UNESCO*, 7 septembre 2020, p. 2. Disponible en ligne : <https://unesdoc.unesco.org/ark:/48223/pf0000373434_fre>, consulté le 01/04/2021.

[1215] UNESCO, *op. cit.*, p. 10 : « *[les acteurs de l'IA] devraient réduire l'impact environnemental des systèmes d'IA, ce qui inclut, sans s'y limiter, leur empreinte carbone, afin de réduire autant que possible les facteurs de risque associés au changement climatique et aux changements environnementaux, et d'empêcher l'exploitation, l'utilisation et la transformation non durables des ressources naturelles, qui contribuent à la détérioration de l'environnement et à la dégradation des écosystèmes.* »

[1216] COMMISSION EUROPÉENNE, *Livre blanc du 19 février 2020 sur l'Intelligence Artificielle – Une approche européenne axée sur l'excellence et la confiance*, 19 février 2020, COM(2020) 65, p. 4 : « *L'importance de l'IA ne cessant de croître, il faut dûment tenir compte de l'incidence environnementale des systèmes d'IA tout au long de leur cycle de vie et sur l'ensemble de la chaîne d'approvisionnement, c'est-à-dire en ce qui concerne l'utilisation des ressources pour l'entraînement des algorithmes et le stockage des données.* »

artificielle, la robotique et autres technologies du 20 octobre 2020[1217] soulignent également cet enjeu écologique. Pourtant, alors même que le « bien-être environnemental » figure bien au rang des sept principes éthiques en matière d'IA auxquels adhèrent les institutions européennes[1218], le cadre règlementaire proposé par la Commission européenne en 2021[1219] ne comporte aucune disposition visant à limiter les conséquences environnementales néfastes du développement de cette technologie.

580. **Des propositions non concrétisées à l'échelle nationale.** À l'échelle nationale, enfin, le constat est identique. Le dilemme écologique du développement de l'IA a été évoqué dès les travaux du député Cédric Villani dans son rapport de 2018[1220], mais, depuis, aucune mesure concrète n'a été prise. On notera tout de même la publication en juillet 2020 de la Feuille de route sur l'environnement et le numérique, coconstruite par le Conseil national du numérique et le Haut conseil pour le climat[1221]. Y figurent de nombreuses mesures visant à la construction d'un numérique sobre, au service de la transition écologique et solidaire. Ces mesures s'articulent autour de trois grands axes, particulièrement pertinents au regard de notre sujet : le premier appelle à réduire l'empreinte environnementale du numérique, le deuxième à mobiliser le potentiel du numérique au service de la transition écologique, et le dernier à accompagner l'ensemble de la société vers un numérique responsable[1222]. Non contraignante, cette publication n'a pour le moment donné lieu à aucune suite. Pourtant, bon nombre des

---

[1217] *Résolution 2020/2012(INL) du Parlement européen du 20 octobre 2020 portant recommandations à la Commission sur un cadre aux aspects éthiques de l'intelligence artificielle, la robotique et autres technologies*, p. 14, N°51, notamment : « *the development, deployment and use of these technologies should contribute to the green transition, preserve the environment, and minimise and remedy any harm caused to the environment during their lifecycle and across their entire supply chain.* »

[1218] GROUPE D'EXPERTS DE HAUT NIVEAU EN IA, *Lignes directrices du 8 avril 2019 pour une IA digne de confiance*, 8 avril 2019, p. 16.

[1219] COMMISSION EUROPÉENNE, *Proposition de règlement du Parlement européen et du Conseil établissant des règles harmonisées concernant l'intelligence artificielle (législation sur l'intelligence artificielle) et modifiant certains actes législatifs de l'union*, 21 avril 2021, 2021/0106 (COD).

[1220] C. Villani, *Donner un sens à l'intelligence artificielle*, Rapport dans le cadre d'une mission parlementaire du 8 septembre 2017 au 8 mars 2018 confiée par le Premier Ministre Edouard Philippe, *La Documentation Française*, 8 mars 2018, p. 20.

[1221] CONSEIL NATIONAL DU NUMÉRIQUE, *Feuille de route sur l'environnement et le numérique - 50 mesures pour un agenda national et européen sur un numérique responsable c'est-à-dire sobre et au service de la transition écologique et solidaire et des objectifs de développement durable*, Rapport remis à la ministre de la Transition écologique et solidaire et au secrétaire d'État chargé du Numérique, juillet 2020, disponible en ligne : <https://cnnumerique.fr/environnement_numerique>, consulté le 01/04/2021.

[1222] *Ibid*, p.12 et s. ; voir aussi : M. Vivant, « Publication des travaux du CNNum sur l'environnement et le numérique », *Revue Lamy Droit de l'Immatériel*, 1er août 2020, n°173.

propositions qu'elle contient sont essentielles pour la construction d'un cadre durable au développement de l'IA, objet de notre propos.

581. **Les raisons de l'immobilisme sur la question de l'empreinte écologique du développement de l'IA.** L'absence d'actions concrètes pour limiter l'empreinte environnementale de l'IA s'explique en partie par l'absence de données précises sur cette empreinte[1223]. Lorsque le risque n'est pas immédiat et avéré, les responsables politiques et les entreprises du secteur privé tardent à agir. Pourtant, le droit a déjà été confronté à pareille situation. En effet, les évolutions technologiques qu'ont connues la France, l'Europe ou le monde en général depuis le début du XXème siècle ont généré autant de bénéfices que de risques pour les populations. Les exemples sont nombreux : développement de nouvelles sources de production d'énergie, des télécommunications, des biotechnologies et bien d'autres. Chacun de ces exemples est marqué à la fois par l'amélioration de la santé et de la qualité de vie des sociétés mais également par les risques et incertitudes sur leurs potentielles conséquences néfastes : exposition aux ondes des antennes de télécommunications, aux rayonnements ionisants générés par un réacteur nucléaire[1224]… Comme le veut l'adage, le Droit s'est adapté au Fait et une nouvelle logique juridique d'anticipation des risques futurs et incertains s'est développée en conséquence, à travers un principe à valeur constitutionnelle en France : le principe de précaution[1225].

582. **Plan.** Le développement de l'IA place à nouveau la société, et par conséquent son droit, face à une situation d'incertitude au regard des conséquences néfastes que pourrait avoir l'adoption massive d'une technologie. À ce titre, il semble que les conditions soient réunies pour appliquer le principe de précaution au développement de l'IA, en ce qu'il promeut une logique de réponse immédiate en présence d'incertitudes scientifiques (**Section 1**). Si la logique juridique à appliquer dans la régulation environnementale de l'IA doit être celle de la précaution, un cadre juridique pour une IA écologique, contenant des obligations concrètes pour les entreprises, doit en découler. En effet, le principe de précaution préconise, en présence

[1223] P. Dhar, « The Carbon Impact of Artificial Intelligence », *Nature Machine Intelligence*, 2020, vol. 2, n°8, 423.
[1224] M. Martuzzi, J.A. Tickner, *The Precautionary Principle: Protecting Public Health, the Environment and the Future of Our Children*, World Health Organization Regional Office for Europe, 2004, p.17, disponible en ligne : < https://www.euro.who.int/__data/assets/pdf_file/0003/91173/E83079.pdf>, consulté le 30 avril 2021.
[1225] *Ibid.*

d'incertitudes scientifiques sur les conséquences du développement d'une nouvelle activité, l'adoption immédiate de mesures destinées à limiter la survenance du risque, même s'il est incertain[1226]. Ces mesures doivent viser, d'une part, à accélérer la recherche sur les conséquences environnementales de l'adoption de l'IA (pour réduire l'incertitude) et, d'autre part, à imposer une logique de sobriété numérique dans le développement de l'IA (pour en diminuer l'empreinte environnementale sans délai) (**Section 2**).

583. **Portée des propositions formulées dans le Chapitre.** Pour l'ensemble de ce Chapitre, le secteur de l'électricité sera uniquement pris en exemple. En effet, les propositions développées pour la création d'un cadre durable au développement de l'IA, que ce soit à travers l'application du principe de précaution ou de la reconnaissance d'obligations de transparence et de sobriété numérique, devraient être adoptées de façon transversale, tout secteur d'activité confondu.

Section 1 : L'application du principe de précaution au développement de l'IA

584. **L'incertitude scientifique sur les conséquences écologiques du développement de l'IA.** La conception, le déploiement et l'utilisation des systèmes d'IA ainsi que de toutes les infrastructures sous-jacentes génèrent une empreinte environnementale conséquente. Si cette affirmation fait consensus, l'ampleur et les conséquences de cette empreinte sont aujourd'hui inconnues, faute de données suffisantes. Cette incertitude ne devrait pas justifier une inaction de la part des responsables politiques et législateurs.

585. **Le principe de précaution face à l'incertitude scientifique.** Face à une telle situation, le principe de précaution peut apporter des réponses sur la manière dont la régulation pourrait s'emparer du problème. L'une de ses expressions peut être trouvée dans le quinzième Principe de la Déclaration de Rio de 1992 :

[1226] Y. Petit, « Environnement », *Répertoire de droit international*, Dalloz, Janvier 2010, 89-92.

> « *En cas de risque de dommages graves ou irréversibles, l'absence de certitude scientifique absolue ne doit pas servir de prétexte pour remettre à plus tard l'adoption de mesures effectives visant à prévenir la dégradation de l'environnement.* »[1227]

586. **Plan.** Le principe de précaution a déjà été appliqué lors du développement d'autres technologies par le passé (**§1**), laissant à penser qu'il aurait toute sa place pour fonder « l'esprit » d'une régulation écologique du développement de l'IA (**§2**).

§1 : La pertinence de l'application du principe de précaution à la régulation de l'IA

587. **Plan.** Le principe de précaution est un concept protéiforme, fruit d'une longue histoire d'adaptation du Droit au Fait, et qui a fait l'objet de nombreuses définitions. Hérité des principes ALARA (« *As Low As Reasonably Achievable* ») anglais[1228] et du « *Vorsorgeprinzip* » allemand (littéralement « *principe de précaution* »)[1229], le principe de précaution a rejoint le rang des principes fondamentaux à l'échelle internationale. En France, il fait aujourd'hui partie du bloc de constitutionnalité et se situe donc au plus haut niveau dans la hiérarchie des normes juridiques. Au cœur du principe de précaution se trouve l'idée que les décideurs doivent agir pour protéger l'environnement (et avec lui les intérêts des générations futures) contre tout dommage, même en l'absence de certitude scientifique sur les conséquences d'une situation particulière. L'étude de son origine, ses sources et ses définitions (**A**), ainsi que de son contenu et son application (**B**) nous laissent penser que son application au développement de l'IA, en tant que nouvelle technologie de rupture, est possible.

[1227] *Déclaration de Rio sur l'environnement et le développement adoptée par la conférence des Nations unies sur l'environnement et le développement*, Rio de Janeiro, 1992, Principe 15. Disponible en ligne : <https://www.un.org/french/events/rio92/rio-fp.htm>, consulté le 30 avril 2021.

[1228] Pour une histoire du « principe de précaution anglais » : A.W.K. Yeung, « The "As Low As Reasonably Achievable" (ALARA) principle: a brief historical overview and a bibliometric analysis of the most cited publications », *Radioprotection*, juin 2019, vol. 54, n°2, p. 103-109.

[1229] Pour une étude comparée du développement du principe de précaution : J. Cameron, J. Abouchar, « The Precautionary Principle: A Fundamental Principle of Law and Policy for the Protection of the Global Environment », *Boston College International and Comparative Law Review*, 1991, vol. 14, p.1.

A/ La pertinence de l'esprit du principe de précaution

588. **L'origine du principe de précaution.** Le principe de précaution est le fruit d'une longue histoire débutant dans la deuxième moitié du XXème siècle. Le terme est apparu pour la première fois dans les années 80 en Allemagne. D'autres pays incorporaient à la même période des éléments de l'approche de précaution dans leur politique environnementale. C'est le cas du Royaume-Uni avec son principe de gestion des risques « ALARA » pour « *As Low As Reasonably Achievable* », visant à réduire les risques pour la santé humaine et l'environnement à un niveau « *aussi faible que raisonnablement possible* », appliqué notamment en matière de radioprotection[1230]. L'Allemagne est allée plus loin dans sa démarche en considérant que les autorités de régulation et les gouvernements avaient pour rôle de minimiser les risques en anticipant tout danger potentiel et, si possible, en le prévenant[1231]. Sa première application par le gouvernement de la République Fédérale d'Allemagne visait à justifier l'adoption de mesures très contraignantes pour lutter contre les pluies acides, le réchauffement climatique et la pollution de la Mer du Nord[1232]. À ces problèmes, le nouveau *Vorsorgeprinzip* allemand imposait d'utiliser les technologies disponibles les plus avancées pour minimiser la pollution et les émissions à leur source. Cet exemple illustre déjà comment ce concept peut être appliqué à notre sujet : si les conditions d'incertitudes scientifiques et de risque environnemental sont réunies, alors il est nécessaire d'adopter, immédiatement, toutes les mesures pour le limiter, y compris en imposant le recours à un certain type de technologies ou de bonnes pratiques. Sous la pression de l'Allemagne, le principe de précaution a par la suite été adopté par les douze États membre de la Communauté économique européenne. En effet, il sera inscrit au sein même du Traité fondateur de Maastricht en 1992 et figure aujourd'hui à l'article 191 du Traité sur le fonctionnement de l'Union européenne (TFUE) qui énumère les principes devant fonder la politique de l'Union dans le domaine de l'environnement[1233]. Dans le même temps, le principe

[1230] INTERNATIONAL COMMISSION ON RADIOLOGICAL PROTECTION, « Recommendations of the International Commission on Radiological Protection », *British Journal of Radiology*, Supplement n°6, 1955.
[1231] M. Martuzzi, J.A. Tickner, *op. cit.*, p.43.
[1232] SERVICE DE RECHERCHE DU PARLEMENT EUROPEEN, *Le principe de précaution – définitions, applications et gouvernance*, décembre 2015, PE 573.876, notamment p. 7, disponible en ligne : <https://www.europarl.europa.eu/RegData/etudes/IDAN/2015/573876/EPRS_IDA%282015%29573876_FR.pdf >, consulté le 30 avril 2021.
[1233] O. Sutterlin, « Synthèse – Principe de précaution », *LexisNexis*, Encyclopédies, 2019, 15.

de précaution s'est développé à l'international et a été reconnu officiellement dans la Déclaration de Rio sur l'environnement et le développement de 1992 établissant un partenariat mondial en la matière[1234]. C'est donc logiquement que la France l'a reconnu par la suite en droit interne en 1995 à travers la loi Barnier[1235]. Son adoption a fait l'objet de nombreux débats partout dans le monde mais plusieurs objectifs communs ont fait consensus[1236] :

- La volonté d'agir avant d'attendre la formalisation d'une preuve formelle du dommage ;
- La volonté d'apporter une réponse proportionnée pour ne pas entraver le développement ;
- La volonté de permettre des marges d'erreur, compte tenu des grandes incertitudes entourant les risques environnementaux ;
- La volonté de reconnaître l'intérêt de la protection d'entités non humaines (environnement, espèces animales…), ainsi que de prendre en compte les conséquences des actions actuelles sur les générations futures ;
- L'intérêt de renverser la charge de la preuve à l'égard des acteurs ayant des activités potentiellement dangereuses, notamment pour l'environnement.

589. **La pertinence des objectifs du principe de précaution dans la régulation de l'IA.** L'ensemble de ces objectifs auxquels le principe de précaution tente de répondre entrent chacun en résonnance avec l'objet de notre propos. En effet, en l'absence de preuve formelle des conséquences néfastes pour l'environnement du développement de l'IA, une réponse proportionnée s'avère nécessaire afin d'anticiper de potentiels effets irréversibles pour les générations futures, notamment en termes d'émissions de CO2 et d'extraction de métaux rares.

[1234] *Déclaration de Rio sur l'environnement et le développement adoptée par la conférence des Nations unies sur l'environnement et le développement*, Rio de Janeiro, *op. cit.*
[1235] *Loi n°95-101 du 2 février 1995 relative au renforcement de la protection de l'environnement*, publiée au JORF le 3 février 1995 ; voir aussi C. Cans, « Grande et petite histoire des principes généraux du droit de l'environnement dans la loi du 2 février 1995 », *Revue juridique de l'environnement*, 1995, p. 193.
[1236] M. Martuzzi, J.A. Tickner, *op. cit.*, p.42.

590. **Les définitions du principe de précaution.** Il n'existe pas une définition unique de ce principe, ce qui a, par ailleurs, fait l'objet de nombreuses critiques[1237]. Certains auteurs considèrent qu'il a donné lieu à une multitude de définitions mais également à de nombreux cas d'espèce où il a été mis en œuvre[1238]. D'autres qu'il est devenu un principe fondamental pour les politiques de protection de l'environnement à tous les niveaux (international, régional, national)[1239]. Les sceptiques, enfin, défendent que le principe de précaution tire sa popularité de son imprécision, qu'il n'existe pas de principe unique mais plutôt la conjonction de plusieurs concepts déjà établis, conduisant à son inefficacité pour répondre aux grands enjeux environnementaux[1240]. La pluralité de ses définitions, voire son imprécision, font donc l'unanimité dans la doctrine. Pour l'objet de notre étude et dans un souci de clarté, nous retiendrons les deux définitions suivantes : la première est celle contenue dans la Déclaration de Rio de 1992 citée précédemment[1241], la seconde est celle contenue dans le Code de l'Environnement : « *l'absence de certitudes, compte tenu des connaissances scientifiques et techniques du moment, ne doit pas retarder l'adoption de mesures effectives et proportionnées visant à prévenir un risque de dommages graves et irréversibles à l'environnement à un coût économiquement acceptable* »[1242]. Notons que nous ne présentons pas ici de définition européenne, puisque qu'elle est absente des traités et que le principe de précaution est en réalité décliné dans certains textes de droit dérivé[1243] ou dans la jurisprudence de la CJUE[1244].

[1237] E. Gaillard, « Principe de précaution – Droit interne », in *JurisClasseur Environnement et Développement durable*, 2014, Fasc. 2410, 16.
[1238] E. Fisher, « Precaution, precaution everywhere: developing a "common understanding" of the precautionary principle in the EC », *Maastricht Journal of European and Comparative Law*, 2002, vol. 9, 7–28 ; voir aussi, sur l'application du principe de précaution par l'organe de règlement de différend de l'OMC : H. Ruiz Fabri, « La prise en compte du principe de précaution par l'OMC », *Revue juridique de l'environnement*, 2000, numéro spécial, p. 55.
[1239] J. Cameron, J. Abouchar, « The precautionary principle », *Boston College International and Comparative Law Review*, 1991, vol. 14, pp. 1–27.
[1240] C. Stone, « Is there a precautionary principle ? », *Environmental Law Reporter*, 2001, vol. 7, p. 107 et s.
[1241] *Déclaration de Rio sur l'environnement et le développement adoptée par la conférence des Nations unies sur l'environnement et le développement*, Rio de Janeiro, *op. cit.*
[1242] Code de l'environnement, article L110-1 II 1°.
[1243] Par ex. *Directive 2001/95/CE du 3 décembre 2001 relative à la sécurité générale des produits*, ou *Règlement (UE) n°1380/2013 du Parlement européen et du Conseil du 11 décembre 2013 relatif à la politique commune de la pêche*.
[1244] Par ex. CJUE, 28 janvier 2010, aff. C-333/08, Commission c/ France, pts 91-93.

591. **Les sources du principe de précaution en droit interne.** Comme présenté plus haut, le principe de précaution figure à l'article L110-1 II 1° du Code de l'environnement, au rang des principes dont doivent s'inspirer toutes les mesures prises en matière de protection de l'environnement « dans le cadre des lois qui en définissent la portée ». Toutefois, il est appliqué en France en l'absence de tout encadrement procédural de nature législative ou réglementaire. Le juge administratif s'y est référé explicitement pour la première fois en 1998[1245], le consacrant comme source de légalité interne[1246]. Le principe de précaution figure également dans la Charte de l'environnement, à valeur constitutionnelle depuis la révision constitutionnelle du 1er mars 2005[1247], sous la formulation suivante :

> « *Lorsque la réalisation d'un dommage, bien qu'incertaine en l'état des connaissances scientifiques, pourrait affecter de manière grave et irréversible l'environnement, les autorités publiques veillent, par application du principe de précaution, et dans leur domaine d'attribution, à la mise en œuvre des procédures d'évaluation des risques et à l'adoption de mesures provisoires et proportionnées afin de parer à la réalisation d'un dommage.* »[1248]

592. **Une pluralité de définitions en droit interne.** Il convient de constater que la définition contenue dans la Charte de l'environnement diffère de celle présente dans le Code de l'environnement, notamment en ce qu'elle s'adresse uniquement aux autorités publiques, qu'elle impose la mise en œuvre de procédures d'évaluation des risques et de mesures nécessairement provisoires. Etant donné que l'article en question est d'application directe, le juge national doit composer avec une source constitutionnelle et une source légale dans l'application du principe de précaution. Concernant son champ d'application, le Conseil d'État a eu l'occasion de se prononcer en faveur d'une conception étroite, circonscrite aux décisions affectant l'environnement[1249] et refusant ainsi d'en faire un principe général de gestion des risques[1250]. Toutefois, comme le droit européen s'impose aux États membres, il convient

---

[1245] CE, 25 septembre 1998, *Association Greenpeace France*, Rec. CE 1998, p. 343 ; JurisData n° 1998-973428.
[1246] O. Sutterlin, *op. cit.*, 18.
[1247] *Loi constitutionnelle n° 2005-205 du 1er mars 2005*, publiée au JORF du 2 mars 2005.
[1248] Charte de l'environnement, article 5.
[1249] CE, 2 septembre 2009, n° 318584 ; AJDA, 2009, 1522.
[1250] F.G. Trébulle, « Droit de l'environnement », *Recueil Dalloz*, 2010, n°2468.

d'identifier les sources européennes du principe de précaution avant de pouvoir analyser comment le juge peut l'appliquer en droit interne.

593. **Les sources du principe de précaution en droit européen.** Principe fondateur de l'UE, le principe de précaution figure à l'article 191 du TFUE qui énumère les principes devant fonder la politique de l'Union dans le domaine de l'environnement. Dans la pratique, le champ d'application du principe est beaucoup plus large et s'étend également, par exemple, à la politique des consommateurs. Il pourra donc tout à fait trouver à s'appliquer en matière de régulation du numérique qui relève des politiques de concurrence, de protection des consommateurs et, dans une moindre mesure, d'environnement. S'il n'est pas défini dans le TFUE, sa mise en œuvre est toutefois encadrée par le droit dérivé de l'UE et la jurisprudence de la CJUE qui s'impose aux États membres.

594. **Des précisions apportées par la CJUE.** Concernant la jurisprudence de la CJUE, on pourra citer en particulier l'arrêt *Afton Chemical* du 8 juillet 2010[1251] dans lequel la Cour considère que dans un cadre technique complexe et évolutif (correspondant à notre objet d'étude, le numérique et l'IA), « *l'application correcte du principe de précaution présuppose, en premier lieu, l'identification des conséquences potentiellement négatives pour la santé de l'utilisation proposée* » et « *en second lieu, une évaluation complète du risque pour la santé fondée sur les données scientifiques disponibles les plus fiables et les résultats les plus récents de la recherche internationale* »[1252]. La Cour retient également que « *lorsqu'il s'avère impossible de déterminer avec certitude l'existence ou la portée du risque allégué en raison de la nature insuffisante, non concluante ou imprécise des résultats des études menées, mais que la probabilité d'un dommage réel pour la santé publique persiste dans l'hypothèse où le risque se réaliserait, le principe de précaution justifie l'adoption de mesures restrictives, sous réserve qu'elles soient non discriminatoires et objectives* »[1253]. Cet arrêt définit donc précisément les conditions dans lesquelles le principe de précaution trouve à s'appliquer. Il constitue l'une des

[1251] CJUE 8 juillet 2010, n° C-343/09, Afton Chemicals.
[1252] *Ibid.*, §60.
[1253] *Ibid.*, §61.

principales sources auxquelles il conviendra de se référer pour en comprendre le contenu, étudié ci-dessous.

B/ La pertinence du contenu du principe de précaution

595. **Critères d'application du principe de précaution.** D'abord, lorsque la mise en œuvre du principe de précaution relève du champ d'application d'un acte de droit dérivé de l'UE, c'est ce dernier qui encadrera son application[1254]. Si un acte de droit dérivé (règlement ou directive européenne) venait à être adopté pour réguler l'IA, son contenu pourrait très bien prévoir et encadrer l'application d'un principe de précaution dans son champ d'application. Dans ce cas, ces dispositions s'imposeraient aux États membres. Si la mise en œuvre du principe de précaution n'est en revanche pas encadrée par un texte européen, c'est à la jurisprudence de la CJUE qu'il convient de se référer[1255]. En effet, la CJUE a précisé les contours du principe de précaution en prévoyant notamment deux prérequis :

- l'identification préalable des conséquences potentiellement négatives pour la santé publique ou l'environnement découlant de l'exercice d'une activité ; et
- la réalisation d'une évaluation complète du risque, à partir des données scientifiques disponibles les plus fiables et les résultats les plus récents de la recherche internationale.

596. **L'application par défaut du principe de précaution en cas d'évaluation du risque incomplète.** L'évaluation du risque ne peut pas se fonder sur des considérations purement hypothétiques. Si ces deux étapes ont été réalisées mais qu'il s'avère impossible de déterminer avec certitude l'existence ou la portée du risque allégué en raison de la nature insuffisante, non concluante ou imprécise des résultats des études menées, mais que la probabilité d'un dommage réel persiste dans l'hypothèse où le risque se réaliserait, le principe de précaution justifie

[1254] J. Molinier, « Primauté du droit de l'Union européenne », *Répertoire de droit européen*, Dalloz, septembre 2011, 2–20.
[1255] *Ibid.*, 21–36.

l'adoption de mesures restrictives, sous réserve qu'elles soient non discriminatoires, objectives et proportionnées[1256].

597. **Application du principe de précaution à la régulation de l'IA.** Au regard de notre sujet, la régulation de l'IA, il convient de distinguer deux hypothèses : soit l'IA fait l'objet d'une régulation à l'échelle européenne (ce qui est la piste poursuivie par les institutions européennes), soit l'IA fait l'objet d'une régulation à l'échelle nationale. Ces deux hypothèses ne sont pas exclusives puisque, même si la régulation européenne était d'application directe (un règlement), il n'est pas exclu que les États membres la complètent par des mesures d'ordre interne. De la même manière, si la régulation européenne se fait par la voie d'une directive, alors des actes de transposition devront être adoptés par les États membres. Dans un cas comme dans l'autre, le principe de précaution pourrait être invoqué tant devant la CJUE que devant les juridictions administratives nationales en tant que source de légalité interne. Son invocation pourrait permettre de contrôler, si ses critères sont remplis, l'adoption effective de mesures destinées à prévenir les conséquences environnementales néfastes que pourrait avoir le développement de l'IA. À cet égard, il convient d'analyser les facteurs déclencheurs[1257] du principe de précaution au regard de notre sujet :

- L'existence d'une incertitude scientifique : l'incertitude doit résulter de l'insuffisance des connaissances scientifiques et techniques, et non au caractère aléatoire du phénomène concerné (l'incertitude ne doit pas porter uniquement sur la date de survenance du dommage)[1258]. Elle doit faire l'objet d'un début de preuve, permettant de fonder le risque sur une hypothèse scientifiquement crédible[1259]. Il nous semble que ce critère soit rempli dans le cas de l'empreinte environnementale de l'IA. En effet, si son existence est réelle et fait consensus au niveau mondial, son ampleur et ses conséquences potentielles font l'objet d'une incertitude marquée par l'insuffisance des données et connaissances en la matière. Les données manquent pour évaluer la consommation énergétique des systèmes d'IA, et les connaissances manquent pour en

[1256] CJUE, 28 janvier 2010, aff. C-333/08 , Commission c/ France, pts 91-93 ; CJUE, 29 avril 2010, aff. C-446/08, Solgar Vitamin's France ; CJUE, 8 juillet 2010, aff. C- 343/09, Afton, voir aussi O. Sutterlin, *op. cit.*, N°17.
[1257] O. Sutterlin, *op. cit.*, 35 et s.
[1258] *Ibid.*
[1259] G. Viney, P. Kourilsky, *Le principe de précaution*, Rapport au Premier Ministre, 15 octobre 1999, p. 65.

déterminer les conséquences environnementales à plus ou moins long terme. Et cela vaut également pour les besoins en métaux rares de ces systèmes.

- L'existence d'un risque de dommage grave ou irréversible : L'autorité publique qui met en œuvre le principe, ou la personne qui réclame sa mise en œuvre ou allègue sa violation, doivent fournir des éléments suffisamment précis et circonstanciés quant aux effets potentiellement négatifs qu'ils suspectent (nature et ampleur du dommage)[1260]. Sa gravité, si elle ne doit pas être parfaitement démontrée, doit être au moins suspectée[1261]. Sur le risque en lui-même, il doit être également suspecté. Il se distingue du risque avéré, pour lequel la probabilité d'occurrence du dommage est déterminée avec certitude : on sait que le risque va se réaliser mais il peut persister une incertitude sur l'échéance de la réalisation du dommage[1262]. En ce qui concerne l'impact écologique du développement de l'IA, il ne ressort pas de la littérature scientifique que le dommage environnemental soit déterminé avec certitude[1263], même si de nombreux éléments permettent de le suspecter[1264]. En effet, des incertitudes persistent sur la réalité du danger que l'empreinte environnementale de l'IA représente, ainsi que sur sa cible (qui sera impacté ?). La gravité du dommage potentiel, elle, peut être en grande partie démontrée par analogie avec d'autres activités fortement émettrices de CO2 et consommatrices de métaux rares.

- La réalisation d'une évaluation scientifique des risques : Elle constitue une garantie procédurale importante afin d'éviter la prise de mesures arbitraires[1265]. Elle doit être réalisée par des experts, soumis à une exigence d'excellence, de transparence et d'indépendance. L'étude doit traiter successivement de l'identification du danger, de sa caractérisation, de l'évaluation de l'exposition au risque et de la caractérisation du

[1260] O. Sutterlin, *op. cit.*, 35 et s.
[1261] CJCE, 9 sept. 2003, aff. C236/01, Monsanto e.a., pt 111 ; CJCE, 26 mai 2005, aff. C132/03, Codacons et Federconsumatori, pt 61 ; CJCE, 12 janv. 2006, aff. C504/04, Agrarproduktion Staebelow, pt 39.
[1262] O. Sutterlin, *op. cit.*, 36.
[1263] B. D'Amico, R.J. Myers, J. Sykes, *et al.*, « Machine Learning for Sustainable Structures: A Call for Data », *Structures*, juin 2019, vol. 19, n°1, 4 ; A.L. Stein, « Artificial Intelligence and Climate Change », *Yale Journal on Regulation*, 2020, 37, n°890, p. 31 et s. ; C. Villani, *op. cit.*, p. 123 et s.
[1264] *Ibid.* ; voir aussi P. Dhar, *op. cit.*
[1265] TPICE, 30 juin 1999, aff. T-13/99 R, pt 172.

risque (probabilité, fréquence…)[1266]. Jamais réalisée pour les conséquences, notamment environnementales, du développement de l'IA, une telle évaluation, commanditée par la France ou par une institution de l'Union européenne, permettrait de mieux cerner l'ampleur de l'enjeu. En France, l'Agence de la Transition énergétique, ex-ADEME, par la nature de ses missions, semble être la mieux placée pour réaliser cette évaluation. Toutefois, l'octroi de moyens supplémentaires à l'Arcep, proposé dans un amendement au projet de loi « Climat et Résilience », pourrait lui permettre d'y contribuer fortement[1267].

- L'existence d'une politique de gestion des risques : Eminemment politique, la gestion des risques nécessite de déterminer le niveau de risque acceptable au regard d'une activité potentiellement dangereuse[1268]. On retrouve ici un concept bien connu des autorités britanniques, appliqué dans leur principe ALARA[1269], celui du « risque acceptable ». En matière de radioprotection, il s'agit de quantifier le taux maximum d'exposition à des rayonnements ionisants auquel on accepte d'exposer les ressortissants d'un État, au vu des risques associés. Dans le sujet qui nous intéresse, la régulation environnementale de l'IA, la gestion des risques devrait, selon nous, se fonder sur les objectifs de neutralité carbone pris par l'Union européenne, et par la France dans le cadre de l'Accord de Paris[1270] et de la Loi relative à la transition énergétique pour la croissance verte[1271]. Ainsi, ne devrait être acceptable qu'un taux

[1266] COMMISSION EUROPÉENNE, *Communication sur le recours au principe de précaution*, COM(2000) 1 final, 2 février 2000, pt 5.1.2 ; O. Sutterlin, *op. cit.*, 37.
[1267] ARCEP, *L'empreinte environnementale des réseaux*, dossier thématique sur le site officiel de l'Arcep, 19 mars 2021, disponible en ligne : <https://www.arcep.fr/la-regulation/grands-dossiers-thematiques-transverses/lempreinte-environnementale-des-reseaux.html>, consulté le 1er mai 2021 ; R. Balenieri, « L'Arcep, en passe d'obtenir de nouveaux pouvoirs en matière d'environnement », *Les Echos*, 2 avril 2021, disponible en ligne : <https://www.lesechos.fr/tech-medias/hightech/larcep-en-passe-dobtenir-de-nouveaux-pouvoirs-en-matiere-denvironnement-1303980>, consulté le 1er mai 2021.
[1268] M.-A. Hermitte, V. David, « Évaluation des risques et principe de précaution », *LPA*, 30 nov. 2000, p. 13.
[1269] Voir Supra, 587 et s.
[1270] *Accord de Paris sur le Climat*, adopté le 12 décembre 2015, lors de la 21ème session de la Conférence des Parties à la Convention-cadre des Nations Unies sur les changements climatiques, visant en particulier à contenir « *l'élévation de la température moyenne de la planète nettement en dessous de 2°C par rapport aux niveaux préindustriels et en poursuivant l'action menée pour limiter l'élévation de la température à 1,5°C par rapport aux niveaux préindustriels, étant entendu que cela réduirait sensiblement les risques et les effets des changements climatiques* » (article 2).
[1271] *Loi n° 2015-992 du 17 août 2015 relative à la transition énergétique pour la croissance verte*, publiée au JORF n°0189 du 18 août 2015.

d'émission de CO2 compatible avec l'objectif fixé de réduire de 40% nos émissions d'ici 2030 par rapport au niveau de 1990 et de 75% d'ici 2050[1272]. La régulation environnementale de l'IA devra aussi contribuer à l'objectif de réduire la consommation énergétique de 50% d'ici 2020[1273]. La France ayant déjà été condamnée par le Conseil d'État pour son défaut d'action suffisante au regard de ces objectifs[1274], le risque acceptable lié à l'empreinte environnementale causée par un déploiement massif de systèmes d'IA pourrait être particulièrement bas.

Le développement des technologies d'IA et les risques environnementaux associés semblent ainsi remplir les critères d'application du principe de précaution. Reste à savoir quels seraient les acteurs concernés par une telle application.

598. **Acteurs concernés par le principe de précaution.** Au vu de ses sources et des exemples cités précédemment, il ressort que le principe de précaution s'adresse avant tout aux autorités publiques, qui le mettent en œuvre dans leur domaine de compétence[1275]. Il s'agit donc plus particulièrement des législateurs nationaux d'États membres de l'Union européenne qui souhaiteraient réguler l'impact environnemental des systèmes d'IA, d'une part, et du système législatif européen en charge de la régulation de l'IA, d'autre part. Toutefois, il convient de noter que les obligations liées à la précaution seraient répercutées par les décideurs publics sur les personnes privées et les entreprises[1276]. C'est pourquoi l'industrie ne doit pas occulter ces questions et contribuer activement au débat public et ce, afin de ne pas subir une règlementation trop contraignante au regard de l'impact véritable de la technologie. D'autant plus que c'est bien le secteur privé, et en particulier les industriels, qui détiennent le plus d'information sur l'empreinte environnementale des systèmes d'IA. La co-construction de la régulation est donc ici primordiale.

[1272] GOUVERNEMENT FRANÇAIS, « COP 21 : Les engagements de la France », *Site officiel du gouvernement*, 1er décembre 2015, disponible en ligne : < https://www.gouvernement.fr/cop21-les-engagements-nationaux-de-la-france-3403>, consulté le 1er mai 2021.
[1273] *Ibid.*
[1274] CE, 19 novembre 2020, *Commune de Grande Synthe*, Recueil Lebon, n° 2020:427301.
[1275] O. Sutterlin, *op. cit.*, 17.
[1276] P. Kromarek, « Le principe de précaution vu par l'industrie », *Droit de l'environnement*, n° 7-8, 2001, p. 189.

599. **Une logique d'action pertinente dans la régulation de l'IA.** Si les conditions de l'application du principe de précaution semblent pouvoir être remplies par la situation causée par le développement exponentiel des systèmes d'IA, c'est avant tout la logique d'action en découlant qui nous apparaît pertinente au regard de notre sujet. En effet, dans un contexte d'incertitude scientifique et d'absence de consensus, c'est bien à l'adoption immédiate de mesures visant à la fois à quantifier les conséquences environnementales potentielles du développement de l'IA ainsi qu'à la minimisation de son empreinte, que nous appelons et ce, dans une logique de précaution. L'application du principe de précaution à la régulation de l'IA, notamment sur son volet environnemental, pousserait à l'adoption de telles mesures.

§2 : Des conséquences de l'application du principe de précaution à la régulation de l'IA

600. **L'empreinte environnementale liée à la consommation énergétique des centres de données.** Les besoins en énergie (notamment électrique) des systèmes d'IA sont particulièrement importants. Pour rappel, les centres de données (ou *« data centers »*) représentent aujourd'hui plus de 2% de la consommation mondiale d'électricité[1277]. Selon des études récentes, ce chiffre devrait être porté, d'ici 2025, à 8% dans le meilleur des cas, ou à 21% si rien n'est fait pour entraver la croissance de ce secteur[1278].

601. **L'empreinte environnementale liée à la fabrication des infrastructures et terminaux électroniques.** À cela, doivent s'ajouter les besoins en métaux rares pour la construction des infrastructures, centres de données, objets connectés pour la collecte de données et autres matériels nécessaires au bon fonctionnement de ces systèmes. Une étude du *Massachussets Institute of Technology* réalisée en 2019 met en lumière le manque de données pour lever les

[1277] F. Pierce, « Energy Hogs: Can World's Huge Data Centers Be Made More Efficient ? », *Yale Environment*, 3 avril 2018, 360, disponible en ligne : <https://e360.yale.edu/features/energy-hogs-can-huge-data-centers-be-made-more-efficient>, consulté le 5 avril 2021.
[1278] M. Giles, « Is Al the Next Big Climate-Change Threat? We Haven't a Clue », *MIT Technology Review*, 29 juillet 2019, disponible en ligne : <https://www.technologyreview.com/2019/07/29/663/ai-computing-cloud-computing-microchips>, consulté le 5 avril 2021, voir aussi A.S.G. Andrae, T. Edler, « On Global Electricity Usage of Communication Technology: Trends to 2030 », *Challenges*, 2015, vol. 6,n °1, pp. 117-138.

incertitudes qui pèsent sur l'ampleur des conséquences de l'impact environnemental du développement de l'IA[1279].

602. **Plan.** C'est dans ce contexte d'incertitude que nous plaidons pour la mise en œuvre de la logique d'anticipation prônée par le principe de précaution. En effet, même sans données fiables, une gouvernance écologique de l'IA est possible. Sans patienter en quête d'une certitude scientifique, le principe de précaution doit nous pousser à agir dès aujourd'hui[1280]. Ce dernier, adapté à la situation actuelle, doit constituer la philosophie, l'esprit ou l'objectif de tout cadre de régulation de l'IA (**A**), dont l'impact environnemental doit être pris en compte. La précaution dans la gouvernance du développement de l'IA, tous secteurs confondus, doit d'abord porter sur son empreinte environnementale (**B**). Toutefois, l'IA peut également être utilisée de telle manière qu'elle accroît la pollution ou les émissions d'une activité, par exemple en participant à l'optimisation de processus d'extraction de matières premières. Une précaution particulière devrait donc aussi être portée à la finalité poursuivie par l'usage de l'IA (**C**), en plus de sa propre empreinte environnementale.

A/ La précaution comme philosophie de la régulation écologique de l'IA

603. **Inclure la logique de précaution dans la régulation de l'IA.** La situation induite par le développement de l'IA, et en particulier l'incertitude scientifique planant sur ses conséquences environnementales, pourrait remplir les conditions jurisprudentielles d'application du principe de précaution[1281]. C'est ici à sa logique d'anticipation et d'action immédiate que nous faisons appel. En l'absence de consensus scientifique sur l'ampleur de l'empreinte environnementale de l'IA et sur les conséquences de son utilisation dans des secteurs émetteurs de CO2 tels que

[1279] M. Giles, *op. cit.* ; voir aussi A.L. Stein, *op. cit.*, p. 29 et s.
[1280] Pour un argumentaire en faveur de l'application du principe de précaution à la gouvernance de l'IA, voir : M. Kuziemski, « Un principe de précaution face à l'intelligence artificielle », *Media24 (blog)*, 2018, disponible en ligne : <ttps://www.medias24.com/chro18268403052018Un-principe-de-precaution-face-a-l-intelligence-artificielle.html>, consulté le 14 avril 2021.
[1281] Voir Supra, 597.

les industries pétrolières et gazières, il semble urgent de tenter de limiter le risque autant que possible.

604. **Une logique d'action contenue dans la Charte des Nations unies.** L'article 11 de la Charte des Nations unies présente une logique d'action concrète dérivée du principe de précaution qui peut être appliquée à la régulation de l'IA :

> « ***Les activités pouvant avoir un impact sur la nature seront contrôlées*** *et* ***les meilleures techniques disponibles, susceptibles de diminuer l'importance des risques ou d'autres effets nuisibles sur la nature, seront employées*** *en particulier :*

a) les activités qui risquent de causer des dommages irréversibles à la nature seront évitées ;

b) les activités comportant un degré élevé de risques pour la nature seront précédées d'un examen approfondi et leurs promoteurs devront prouver que les bénéfices escomptés l'emportent sur les dommages éventuels sur la nature et, lorsque les effets nuisibles éventuels de ces activités ne sont qu'imparfaitement connus, ces dernières ne devraient pas être entreprises ;

c) *les activités pouvant perturber la nature seront précédées d'une évaluation de leurs conséquences et des études concernant l'impact sur la nature des projets de développements seront menées suffisamment à l'avance ; au cas où elles seraient entreprises, elles devraient être planifiées et exécutées de façon à réduire au minimum les effets nuisibles qui pourraient en résulter.* »[1282]

605. **Application à la régulation de l'IA.** À la lecture de l'article 11 de la Charte des Nations unies, le déploiement des systèmes d'IA devrait faire l'objet d'un contrôle particulier et des mesures destinées à diminuer l'importance du risque devraient être employées. En effet, l'impact du développement de l'IA sur la nature est certain, bien que son ampleur fasse encore l'objet d'incertitudes. Les mesures à mettre en œuvre devraient d'abord conduire à éviter les applications d'IA qui causeraient un dommage irréversible sur la nature, notamment celles qui *permettent* une atteinte à l'environnement qui n'aurait pas été possible sans elle (par exemple,

[1282] Charte des Nations unies, signée le 26 juin 1945 à la Conférence des Nations Unies pour l'Organisation internationale, San Francisco, article 11.

des systèmes capables de déterminer des zones propices à l'exploitation de ressources naturelles, qui causerait une destruction de l'écosystème alentours). Elles devraient ensuite imposer la réalisation d'une évaluation des risques et d'un calcul coût-avantages dans tous les cas où un système d'IA mobilise de grandes ressources énergétiques (liées notamment à l'utilisation de supercalculateurs ou de techniques d'apprentissage automatique). Les mesures de précaution, si l'on suit la logique de la Charte des Nations unies, devraient enfin prévoir la minimisation systématique des effets nuisibles du déploiement de systèmes d'IA, autrement dit imposer une obligation de minimiser l'empreinte environnementale des systèmes systématiquement dans leur conception et leur opération.

B/ Une précaution à porter à l'empreinte environnementale des systèmes d'IA

606. **Le risque lié à l'empreinte environnementale des systèmes d'IA.** Si l'on souhaite mettre le potentiel des technologies d'IA au service de la transition énergétique et de la lutte contre le réchauffement climatique, en particulier dans le secteur de l'électricité, il faut s'assurer que ses conséquences négatives sur l'environnement soient contrebalancées par ses impacts positifs. En effet, le déploiement et l'opération de systèmes d'IA requièrent de grandes capacités de calcul et de stockage de données, auxquelles s'ajoutent le cas-échéant, les infrastructures nécessaires à la collecte des données traitées. Ces besoins matériels, souvent oubliés[1283], ont des conséquences environnementales non négligeables. Tout d'abord, les capacités de calcul et de stockage nécessaires à l'opération des techniques algorithmiques dites d'IA présentent une consommation énergétique très importante[1284]. Certaines techniques d'IA, notamment l'apprentissage automatique, sont par ailleurs plus consommatrices que d'autres, qui ne nécessitent pas le traitement de millions voire de milliards de données[1285]. S'ils s'opèrent

[1283] A. Borning, « What Pushes Back Considering Materiality in IT ? », *Limits*, 2018, disponible en ligne : <https://computingwithinlimits.org/2018/papers/limitsl8-borning.pdf>, consulté le 7 avril 2021.
[1284] D. Amodei, D. Hernandez, « Al and Compute », *Open AI (blog)*, 16 mai 2018, disponible en ligne : <https://openai.com/blog/ai-and-compute/#fn2>, consulté le 7 avril 2021.
[1285] Les émissions de CO2 liées à l'apprentissage de certains modèles de langage naturel peuvent représenter jusqu'à cinq fois les émissions d'une voiture pendant toute sa durée de vie : E. Strubel, « Energy and Policy Considerations for Deep Learning in NLP », *57th Annual meeting of the Association for Computational Linguistics*, 5 juin 2019, p. 1 et p. 4.

« dans le *Cloud* » dans l'imaginaire collectif, ces calculs sont en réalité localisés sur des serveurs ou supercalculateurs, souvent situés eux-mêmes dans des centres de données. Ce sont donc bien des infrastructures matérielles qui sont derrière l'opération des systèmes d'IA. Ces matériels, ainsi que ceux destinés à la collecte des données, nécessitent pour leur conception des matériaux rares tels que le tantale, l'indium ou le cobalt[1286]. L'extraction de matières premières pour la conception de ces matériels électroniques, qui vont se multiplier au même rythme que le recours aux systèmes d'IA, peut avoir des conséquences environnementales irréversibles[1287]. Pour que les conséquences environnementales négatives du déploiement de systèmes d'IA soient contrebalancées par ses impacts potentiellement positifs, il faut donc s'attacher, en premier lieu, à limiter sa propre empreinte environnementale, à savoir sa consommation énergétique, ainsi que ses besoins en métaux rares pour la conception des matériels nécessaires à son utilisation. C'est à ces deux risques que doit s'appliquer une logique de précaution.

607. **La précaution environnementale dans la conception et l'opération des systèmes basés sur l'IA.** Pour rappel, le principe de précaution consiste en l'adoption de mesures immédiates, proportionnées et provisoires lorsqu'une situation risque de causer un dommage grave et irréversible à l'environnement, mais que ce dommage fait l'objet d'une incertitude scientifique[1288]. Comme la situation causée par le développement de l'IA semble le justifier, il est nécessaire de mettre en place un cadre de régulation pour limiter son empreinte environnementale même en l'absence de données précises. L'incertitude qui pèse sur les conséquences de cette empreinte, ainsi que la nécessité d'agir dès maintenant pour la limiter font l'objet d'un consensus à l'international[1289]. Certains domaines de l'IA, tels que le

[1286] ADEME, *La face cachée du numérique – Réduire les impacts du numérique sur l'environnement*, janvier 2021, p .6, rubrique « Des objets qui pèsent lourds dans notre quotidien ».
[1287] P. Dhar, *op. cit.*
[1288] Voir Supra, 597.
[1289] L. Calandri, « Pollution numérique et intelligence artificielle : variations autour du rapport de Monsieur le Député C. Villani, "Donner un sens à l'intelligence artificielle" », *Energie - Environnement – Infrastructures*, novembre 2018, n°11, 15 ; voir aussi P. Rejcek, « AI Is an Energy-Guzzler. We Need to Re-Think Its Design, and Soon », *Singularity Hub*, 28 février 2020, disponible en ligne : <https://singularityhub.com/2020/02/28/ai-is-an-energy-guzzler-we-need-to-re-think-its-design-and-soon/>, consulté le 7 avril 2021 ; M. Whittaker, R. Dobbe, « AI and Climate Change: How they're connected, and what we can do about it », *Medium (blog)*, 17 octobre 2019, disponible en ligne : <https://medium.com/@AINowInstitute/ai-and-climate-change-how-theyre-connected-and-what-we-can-do-about-it-6aa8d0f5b32c>, consulté le 7 avril 2021.

traitement du langage, font déjà l'objet d'initiatives pour limiter leurs impacts[1290]. Ces mesures de précaution devraient viser à limiter l'empreinte environnementale par tous moyens, dans l'attente d'avoir plus de données pour identifier plus précisément ses conséquences. Elles pourraient consister en des obligations incombant aux développeurs ou opérateurs de systèmes d'IA.

608. **Obligations de sobriété pour les développeurs de systèmes d'IA.** Pour les développeurs, ces obligations devraient encadrer la conception des systèmes matériels et logiciels, le choix des méthodes d'apprentissage et des données d'entraînement pour s'assurer que la consommation énergétique liée à la conception du système global et l'impact environnemental du choix des matériaux soient minimisés.

609. **Obligations de sobriété pour les opérateurs de systèmes d'IA.** Pour les opérateurs, des mesures de précaution pourraient être adoptées pour encadrer la performance énergétique des serveurs sur lesquels les algorithmes fonctionnent, ainsi que sur leur source d'énergie, qui devrait être la plus décarbonée possible.

610. **Des obligations visant à la minimisation de l'empreinte écologique de l'IA.** De telles mesures, si elles sont contraignantes, devraient permettre de limiter autant que possible les conséquences environnementales du développement de systèmes d'IA et ce, dans l'attente de certitude scientifique sur ces conséquences et d'avancées techniques en matière de réduction des besoins énergétiques de ces systèmes.

611. **Intégrer les mesures de précaution environnementale à la régulation de l'IA.** L'ensemble des mesures présentées ci-dessus devraient, selon nous, être adoptées dans une logique de co-régulation, à l'instar de la régulation de l'IA proposée dans le reste de la thèse. D'une part, elles doivent être contraignantes pour ne pas créer de déséquilibre compétitif entre les entreprises vertueuses et les autres. De l'autre, leur mise en œuvre doit être assez souple pour ne pas créer un fardeau règlementaire trop lourd, qui freinerait l'innovation et, *in fine*, ralentirait la recherche pour le développement d'une technologie écoresponsable. Comme l'objectif et l'esprit sont les mêmes que ceux de la régulation proposée dans le reste de la thèse,

[1290] E. Strubel, *op. cit.*, p. 8.

il semble opportun d'intégrer ces mesures au cadre règlementaire en construction au niveau européen. De plus, cette approche écologique de l'adoption de l'IA semble en phase avec les objectifs environnementaux de l'UE, ses principes fondateurs[1291] et son récent *Green Deal*[1292]. Les modalités de l'intégration de ces mesures dans le cadre règlementaire en construction ainsi que leur détail seront abordés dans la Section 2 du présent Chapitre.

C/ Une précaution à porter à la finalité de l'usage des systèmes d'IA

612. **La nécessité de détourner l'IA de ses applications à fort impact environnemental.** Les mesures de précaution destinées à construire un cadre écologique au développement de l'IA doivent participer à détourner cette technologie d'applications polluantes ou facilitant la conduite d'activités polluantes. Une piste pourrait être l'intégration d'un critère relatif à la finalité écologique de l'IA, condition à son utilisation *via* un processus de certification *ex-ante*. La mise en œuvre d'un tel critère pourrait conduire les fournisseurs de solutions d'IA à délaisser les secteurs à forte empreinte environnementale. Les secteurs pétrolier[1293] et gazier[1294] peuvent ici être pris en exemple, puisque l'Agence Internationale de l'Energie a reconnu la nécessité de

[1291] UNION EUROPEENNE, *Synthèse de la législation européenne en matière d'environnement et de lutte contre le changement climatique*, disponible en ligne : <https://eur-lex.europa.eu/summary/chapter/environment.html?root_default=SUM_1_CODED=20&locale=fr>, consulté le 6 mai 2021 : « *La politique environnementale de l'UE s'appuie sur l'article 11, ainsi que sur les articles 191 à 193 du traité sur le fonctionnement de l'Union européenne. En vertu de l'article 191, la lutte contre le changement climatique est un objectif explicite de la politique environnementale de l'Union. Le développement durable est un objectif primordial de l'UE, qui s'engage à assurer «un niveau élevé de protection et d'amélioration de la qualité de l'environnement» (article 3 du traité sur l'Union européenne).* »

[1292] COMMISSION EUROPÉENNE, *Le pacte vert européen*, Communication de la Commission au Parlement européen, au Conseil européen, au Conseil, au Comité économique et social européen et au Comité des régions, 11 décembre 2019, COM(2019) 640 final, notamment p. 10 : « *Les technologies numériques s'avèrent d'une importance cruciale pour atteindre les objectifs fixés par le pacte vert en matière de développement durable, et ce dans une grande variété de secteurs. La Commission étudiera des mesures visant à faire en sorte que les technologies numériques, telles que l'intelligence artificielle […] puissent accélérer et optimiser l'impact des politiques de lutte contre le changement climatique et de protection de l'environnement* ».

[1293] S.D. Mohaghegh, « Recent Developments in Application of Artificial Intelligence in Petroleum Engineering », *Petroleum Tech*, 2005, 57, 86.

[1294] J. Williams, « Does Al Have the Power to Refine Oil and Gas Efficiency? », *EY (blog)*, 4 juin 2019, disponible en ligne : <https://www.ey.com/en us/oil-gas/does-ai-have-the-power-to-refine-oil-and-gas-efficience>, consulté le 8 avril 2021.

désinvestir ces secteurs pour atteindre l'objectif de neutralité carbone d'ici 2050[1295]. De plus, un rapport de Greenpeace de 2020[1296] a mis en évidence les liens entre les acteurs majeurs de ces secteurs tels que Shell, ExxonMobil ou Total, et les géants de l'intelligence artificielle tels que Microsoft, Google et Amazon[1297] et ce, malgré leurs engagements à lutter contre le réchauffement climatique. Dans ces secteurs, l'IA est principalement utilisée afin d'extraire une plus grande quantité de pétrole et de gaz des sols, et afin de diminuer le coût de l'extraction, plutôt que pour réduire leurs émissions. À titre d'exemple, Amazon propose depuis 2018 une gamme de modèles d'apprentissage automatique destinés à *« Prédire le prochain puit de pétrole en quelques secondes grâce au machine learning »*[1298]. Si cette pratique n'est pas illégale, elle reste difficilement compatible avec les engagements de lutte contre le changement climatique que se sont imposés ces entreprises. C'est pourquoi, à la suite de ces révélations, Google a fait le choix de renoncer à développer des systèmes d'IA destinés à faciliter l'extraction de pétrole ou de gaz[1299]. Ses deux principaux concurrents, Microsoft et Amazon, vont en revanche continuer de fournir ce secteur d'activité, bien que polluant, par « souci d'égalité d'accès à la technologie »[1300]. L'adoption de mesures contraignantes encadrant la finalité de l'usage de l'IA permettrait d'éviter ce genre de situation où la technologie est mise au service d'activités particulièrement polluantes.

613. **La nécessité de mettre l'IA au service de la précaution environnementale.** Plutôt que simplement interdire l'usage de l'IA au service d'activités polluantes, il pourrait être intéressant de promouvoir l'utilisation de l'IA au service de l'environnement. Rejoignant une proposition de régulation formulée par l'*AI Now Institute*[1301], une telle incitation pourrait contrebalancer les conséquences environnementales négatives des systèmes d'IA. Il faut remarquer que cette

---

[1295] IEA, *Net Zero by 2050 : A Roadmap for the Global Energy Sector*, mai 2021, disponible en ligne : <https://iea.li/nzeroadmap>, consulté le 22 mai 2021.
[1296] GREENPEACE, *Greenpeace Report : Oil in the Cloud*, Rapport, 19 mai 2020, disponible en ligne : <https://www.greenpeace.org/usa/reports/oil-in-the-cloud/>, consulté le 8 avril 2021.
[1297] *Ibid.*
[1298] AMAZON WEB SERVICES, « Your Guide to AI and machine learning at re:Invent 2018", *Amazon (blog)*, 27 septembre 2018, disponible en ligne : < https://aws.amazon.com/fr/blogs/machine-learning/your-guide-to-ai-and-machine-learning-at-reinvent-2018/>, consulté le 8 avril 2021.
[1299] R. Sandler, « Google Halts AI Tools For Oil Extraction », *Forbes (blog)*, 19 mai 2020, disponible en ligne : <https://www.forbes.com/sites/rachelsandler/2020/05/19/google-halts-ai-tools-for-oil-industry-after-greenpeace-report/>, consulté le 8 avril 2021.
[1300] *Ibid.*
[1301] M. Whittaker, R. Dobbe, *op. cit.*

technologie est déjà utilisée à des fins écologiques ou pour préserver l'environnement[1302]. Elle est ainsi utilisée pour analyser les fonds marins et les coraux afin de mesurer les effets du changement climatique[1303], ou pour aider à la protection des forêts[1304]. Avec plus d'un milliard de personnes sans électricité dans le monde[1305], l'IA peut également aider à l'électrification en la rendant accessible financièrement[1306] et en facilitant le développement de réseaux locaux d'électricité décarbonée (des « *microgrids* »)[1307]. Enfin, les systèmes d'IA peuvent, en prédisant l'origine des émissions de gaz à effet de serre, aider les décideurs politiques à déterminer comment réguler la production d'énergie[1308]. Au-delà de ces exemples, l'idée de mettre les meilleures technologies disponibles au service de la lutte contre le réchauffement climatique est en phase avec le contenu du principe de précaution environnementale. Le fait que les systèmes d'IA peuvent contribuer à la poursuite des objectifs environnementaux et de développement durable fait l'objet d'un consensus scientifique à l'échelle internationale[1309]. Dans le cadre de la régulation de l'IA, le Parlement européen a d'ailleurs reconnu que « *le développement de l'intelligence artificielle, la robotique et les technologies associées, peut contribuer à atteindre les objectifs de développement durable contenus dans le Green Deal européen dans de nombreux secteurs* » ainsi que « *[ces technologies] peuvent également*

[1302] A.L. Stein, *op. cit.*, p. 29.
[1303] J. Jeffrey, « 8 Companies Using AI to Tackle Climate Change », *Entrepreneur*, 27 Septembre 2019, disponible en ligne : <https://www.entrepreneur.com/article/340002>, consulté le 5 avril 2021.
[1304] J. Snow, « How Artificial Intelligence Can Tackle Climate Change », *National Geographic*, 18 juillet 2019, disponible en ligne : <https://www.nationalgeographic.com/environment/2019/07/artificial-intelligence-climate-change>, consulté le 5 avril 2021.
[1305] H. Ritchie, « Number of People in the World Without Electricity Falls Below One Billion », *Our World Data*, 18 janvier 2019, disponible en ligne : <https://ourworldindata.org/number-of-people-in-the-world-without-electricity-access-falls-below-one-billion>, consulté le 5 avril 2021.
[1306] Y. Landa, « How Artificial Intelligence Will Incredibly Lower Your Energy Bill », *Medium (blog)*, 29 avril 2019, disponible en ligne : <https://medium.com/datadriveninvestor/how-artificial-intelligence-will-incredibly-lower-your-energy-bill-33914791eala> , consulté le 5 avril 2021.
[1307] M. Nichols, « How Will AI Improve Microgrid Energy Efficiency ? », *Schooled By Science*, 25 avril 2019, disponible en ligne : <https://schooledbyscience.com/how-will-ai-improve-microgrid-energy-efficiency>, consulté le 5 avril 2021.
[1308] J. Snow, *op. cit.*
[1309] R. Vinuesa, H. Azizpour, I. Leite, *et al.*, « The Role of Artificial Intelligence in Achieving the Sustainable Development Goals », *Nature Communications*, décembre 2020, 11, n°1, 233 ; pour une liste exhaustive des applications de l'apprentissage automatique au service de la lutte contre le réchauffement climatique, secteur par secteur, voir le manifeste de data scientists « Tackling Climate Change with Machine Learning » : D. Rolnick, P.L. Donti, L.H. Kaack, *et al.*, « Tackling Climate Change with Machine Learning », *ArXiv*, 5 novembre 2019, 1906.05433, disponible en ligne : <http://arxiv.org/abs/1906.05433>, consulté le 8 avril 2021.

contribuer à réduire grandement les émissions de gaz à effet de serre »[1310]. Au vu du consensus scientifique existant sur l'intérêt de mettre l'IA au service de l'environnement et du besoin de contrebalancer sa propre empreinte, inclure un critère de finalité écologique dans la régulation de l'IA semble donc particulièrement pertinent et serait en phase avec les objectifs poursuivis par les institutions européennes.

614. **Un exemple à travers le secteur de l'énergie électrique.** L'IA peut contribuer à réduire les émissions d'un des secteurs les plus polluants dans le monde[1311] : le secteur de l'électricité. À l'échelle internationale, l'électricité représente près de 25% des émissions de gaz à effet de serre, principalement de CO2[1312]. En parallèle, la demande en électricité devrait augmenter dans les années et décennies à venir en raison de l'électrification des usages, avec notamment la transition du secteur des transports vers les véhicules électriques[1313]. À moins que des mesures ne soient prises pour décarboner la production d'électricité (ce qui est déjà le cas en France, mais pas dans tous les États, notamment européens), une croissance de la demande en électricité conduira à une hausse des émissions du secteur. Le secteur de l'énergie électrique, comprenant à la fois la production, le transport, la distribution ainsi que la vente d'électricité et les services associés, présente de nombreuses opportunités pour l'IA. Elle peut accélérer le développement de modes de production d'énergie bas carbone, améliorer les prévisions de demande d'électricité, faciliter la conduite du réseau et renforcer sa sécurité[1314]. L'IA permettrait donc de réduire les émissions de gaz à effet de serre émanant de ce secteur, principalement en permettant d'optimiser les infrastructures de réseau pour éviter toute perte d'énergie,

---

[1310] *Résolution 2020/2012(INL) du Parlement européen du 20 octobre 2020 portant recommandations à la Commission sur un cadre aux aspects éthiques de l'intelligence artificielle, la robotique et autres technologies*, p.4, D.

[1311] « Dans le monde » puisque le mix énergétique français, majoritairement décarboné, rend le secteur bien moins émetteur de gaz à effet de serre que d'autres pays qui utilisent, plutôt que le nucléaire, plus de charbon ou de gaz : RTE, *Synthèse du bilan électrique 2020*, disponible en ligne : <https://bilan-electrique-2020.rte-france.com/synthese-les-faits-marquants-de-2020/#>, consulté le 6 mai 2021.

[1312] IEA, *Data and Statistics : CO2 emissions by energy source*, 2018, disponible en ligne : <https://www.iea.org/data-and-statistics/data-browser/?country=WORLD&fuel=CO2%20emissions&indicator=CO2BySector>, consulté le 6 mai 2021.

[1313] IEA, *Global EV Outlook 2019*, 2019.

[1314] A.L. Stein, *op. cit.*, p. 12 ; pour un détail des applications potentielles dans le secteur de l'énergie, voir également : Eurelectric, *AI Insights : The Power Sector in a Post-Digital Age*, Rapport, 26 novembre 2020, disponible en ligne : < https://www.eurelectric.org/media/5016/ai-insights-final-report-26112020.pdf>, consulté le 6 mai 2021.

d'optimiser la performance énergétique des bâtiments, ou enfin d'améliorer la résilience et la fiabilité du système électrique. On voit ici que l'on peut appliquer de façon très concrète la logique de précaution à un secteur donné, en identifiant les applications d'IA permettant de réduire les émissions de gaz à effet de serre. Un déploiement à finalité écologique de l'IA est donc possible dans le secteur de l'électricité.

615. **Transition.** Une application du principe de précaution au développement de l'IA est possible et pourrait être déclinée secteur par secteur. Toutefois, il apparaît que la question des conséquences environnementales du développement de l'IA n'est pas spécifique à cette technologie. Elle est en réalité inhérente à l'utilisation des technologies de l'information et de la communication (TIC) d'une manière générale[1315]. Bien que la thèse ne se concentre que sur la régulation de l'IA, il n'est pas pertinent de raisonner sur ce point uniquement à l'échelle d'un secteur, ni d'une seule technologie. La démarche doit, selon nous, être étendue à l'ensemble du secteur numérique.

§3 : Un indispensable élargissement de la démarche aux autres technologies numériques

616. **Plan.** Les conséquences environnementales du déploiement massif d'une technologie numérique ne sont pas spécifiques aux systèmes d'IA. S'il est certain que le fonctionnement même de l'IA en fait une des technologies les plus énergivores, il ne faut pas perdre de vue qu'en termes de proportion l'empreinte environnementale de l'ensemble des TIC dépasse largement celui de l'IA[1316]. Pourtant, les technologies numériques de rupture (IA, blockchain, internet de objets ou informatique quantique) ont des points communs, au regard des composants électroniques nécessaires à leur opération, de leurs besoins énergétiques ou encore de leur utilisation de données. C'est pourquoi il semble que les principes de précaution ou, à défaut, de prévention ne doivent pas être appliqués uniquement aux systèmes d'IA mais que la réflexion doit être élargie à l'ensemble des technologies numériques (**A**). Les bénéfices apportés

[1315] ADEME, *op. cit.*
[1316] *Ibid.*

par les TIC à la société étant indiscutables, cette application doit être raisonnée et favoriser l'innovation vertueuse (**B**).

A/ L'insuffisance de l'application du principe de précaution aux seuls systèmes d'IA

617. **Un enjeu pour tout le secteur du numérique.** L'utilisation des TIC a entraîné une augmentation considérable de la consommation d'énergie[1317] : le problème n'est pas spécifique à l'IA. L'augmentation du nombre de clients de services numériques, la multiplication des services à la demande grâce au *Cloud computing*, le déploiement de la 4G puis de la 5G, l'essor du streaming en ligne sont autant de raisons à cette augmentation[1318]. Le numérique représente aujourd'hui 4% des émissions de gaz à effet de serre à l'échelle mondiale[1319] et son empreinte ne cesse d'augmenter d'année en année. Pourtant, le problème n'est pas nouveau et la revue *Science* publiait déjà en 2013 une analyse détaillée de l'impact environnemental du développement d'Internet[1320]. En France, c'est uniquement depuis 2018 que le sujet est abordé tant dans des rapports institutionnels[1321] que dans des prises de position de collectifs militants[1322] ou dans la presse grand public[1323]. Face à un consensus sur l'existence et l'importance du problème, plusieurs outils juridiques ont été ou sont en train d'être mis en place afin d'y répondre.

[1317] INSEE, « Enjeux du numérique », *INSEE Références*, édition 2019, Fiche 4.2.
[1318] D. Reforgiato Recupero, « Toward a Green Internet », *Science*, 29 mars 2013, 339, n°6127, 1533, p.8.
[1319] The Shift Project, *Pour une sobriété numérique*, octobre 2018, disponible en ligne : <https://theshiftproject.org/wp-content/uploads/2018/11/Rapport-final-v8-WEB.pdf>, consulté le 10 avril 2021.
[1320] D. Reforgiato Recupero, *op. cit.*
[1321] Voir notamment le rapport sénatorial SENAT, *Rapport d'information sur l'empreinte environnementale du numérique*, rapport sénatorial, 24 juin 2020.
[1322] Voir notamment le rapport de WWF France : Iddri, GreenIT, FING et WWF France, *Livre blanc Numérique et environnement*, 2018, disponible en ligne : <https://www.wwf.fr/sites/default/files/doc-2018-03/180319_livre_blanc_numerique_environnement.pdf>, consulté le 11 mars 2020 ; ou celui du Shift Project : The Shift Project, *Déployer la sobriété numérique*, rapport, disponible en ligne : <https://theshiftproject.org/wp-content/uploads/2020/01/2020-01.pdf>, consulté le 11 mars 2020.
[1323] B. Goldet, « Il est urgent de mesurer les impacts énergétiques et environnementaux de nos solutions numériques », *Le monde de l'énergie*, disponible en ligne : <https://www.lemondedelenergie.com/impacts-energetiques-environnementaux-numerique/2019/10/22/>, consulté le 16 juin 2020.

618. **Les premières réflexions juridiques sur la sobriété numérique.** Il n'existe pas de dispositions encadrant spécifiquement la fabrication ou l'utilisation des TIC pour limiter leur empreinte environnementale. Pour autant, plusieurs initiatives législatives récentes laissent à penser que le sujet est en train d'être pris au sérieux[1324]. Il faut ici distinguer entre le régime réservé aux équipements électroniques, supports matériels au numérique, et celui réservé au volet immatériel du numérique, notamment les données et le streaming.

619. **La sobriété numérique à travers la pratique de l'écoconception.** Concernant le volet matériel du numérique, il s'agit principalement d'encadrer la fabrication et le traitement des déchets liés aux équipements électroniques. L'encadrement de la fabrication de ces équipements peut se faire *via* la notion « d'écoconception » introduite et encouragée par la loi du 10 février 2020 relative à la lutte contre le gaspillage et à l'économie circulaire (dite loi AGEC)[1325]. Cette notion est définie notamment par la norme ISO 14006/2011 comme « *l'intégration des aspects environnementaux dès la conception d'un bien ou d'un service et a pour objectif de réduire les impacts environnementaux tout au long de son cycle de vie : de l'extraction des matières premières, à la production, la distribution, l'utilisation et jusqu'à la fin de vie du produit ou du service* ». Plusieurs travaux sont menés, notamment par l'Agence de la Transition écologique, afin de proposer des normes d'écoconception destinées aux TIC[1326]. L'écoconception doit également prendre en compte des éléments relatifs à la fin de vie du produit. À ce titre, plusieurs textes sont venus encadrer les Déchets d'Equipements Electriques et Electroniques (les « DEEE »), notamment des directives européennes[1327]. Ces directives ont amené la France à développer une filière de gestion spécifique des déchets, règlementée aux articles L541-10-2 et suivants du Code de l'environnement. La loi AGEC du 10 février 2020 a

[1324] V. Piguet, « Sobriété numérique : les premiers outils juridiques mis en place », *Village de la Justice*, 15 février 2021, disponible en ligne : <https://www.village-justice.com/articles/numerique-pollution,38079.html>, consulté le 9 avril 2021.

[1325] *Loi n° 2020-105 du 10 février 2020 relative à la lutte contre le gaspillage et à l'économie circulaire*, publiée au JORF n°0035 du 11 février 2020, dite loi « AGEC », notamment son article 62 disposant « *En application du principe de responsabilité élargie du producteur, il peut être fait obligation à toute personne physique ou morale […] d'adopter une démarche d'écoconception des produits […]* ».

[1326] ADEME, *TIC et impacts environnementaux*, dossier thématique, disponible en ligne : < https://communication-responsable.ademe.fr/digital-eco-responsable/tic-et-impacts-environnementaux/tic-axes-daction-pour-reduire-les-impacts>, consulté le 6 mai 2021.

[1327] Notamment : *Directive 2002/95/CE du 27 janvier 2003, relative aux substances dangereuses dans ces équipements* ; et *Directive 2002/96/CE, relative aux déchets d'équipements électriques et électroniques*.

complété ces dispositions avec l'ajout des articles L541-9-1 relatif à l'information du consommateur et L541-9-2 instaurant un indice de réparabilité pour favoriser l'allongement de la durée de vie des produits. L'ensemble de ces dispositions ne sont pas encore pleinement applicables mais elles participeront à la prise en compte des impacts environnementaux des supports matériels du numérique, en informant mieux le consommateur et en favorisant la réparation des équipements[1328]. Enfin, s'agissant des règles déjà en vigueur, on pourra citer l'application du régime des installations classées pour la protection de l'environnement (ICPE) aux data center[1329] ou le principe de la responsabilité élargie des producteurs (« REP ») dans certaines filières[1330] qui pourrait être élargi au numérique.

620. **La nécessaire sobriété dans les usages des données.** Concernant le volet immatériel du numérique, à savoir principalement les flux de données, plusieurs rapports convergent sur l'insoutenabilité de l'usage actuel de la donnée, dont l'empreinte environnementale ne cesse de croître[1331]. Toutefois, aucun texte n'encadre pour le moment ces usages pour en limiter l'empreinte. On notera tout de même que la Commission européenne semble vouloir adresser ce problème dans sa stratégie sur les données[1332] tout en souhaitant encourager la collecte et le partage de données[1333], ce qui peut paraître paradoxal. En France, si le sujet était déjà présent dans le Rapport Villani en 2018[1334], une loi en date du 15 novembre 2021 contient des propositions pour réduire l'impact environnemental du numérique[1335]. Le gouvernement se saisit du problème et a présenté, début 2021, un plan sans contraintes pour limiter l'impact

---

[1328] V. Piguet, *op. cit.*

[1329] Code de l'Environnement, articles L. 511-1 et s.

[1330] MINISTÈRE DE LA TRANSITION ÉCOLOGIQUE, *Cadre général des filières à responsabilité élargie des producteurs*, 6 avril 2021, disponible en ligne : <https://www.ecologie.gouv.fr/cadre-general-des-filieres-responsabilite-elargie-des-producteurs>, consulté le 6 mai 2021.

[1331] SENAT, *op. cit.* ; voir aussi, à travers l'exemple du streaming en ligne : The Shift Project, *Climat : l'insoutenable usage de la vidéo en ligne*, rapport, juillet 2019, disponible en ligne : <https://theshiftproject.org/wp-content/uploads/2019/07/2019-01.pdf>, consulté le 17 mars 2020.

[1332] COMMISSION EUROPÉENNE, *Une stratégie européenne pour les données*, communication de la Commission au Parlement européen, au Conseil, au Comité économique et social européen et au Comité des régions, 19 février 2020, COM(2020) 66 final.

[1333] COMMISSION EUROPÉENNE, « Commission proposes measures to boost data sharing and support European data spaces », *Communiqué de presse accompagnant la proposition du Data Governance Act*, 25 novembre 2020.

[1334] C. Villani, *op. cit.*, p. 127 et s.

[1335] *Loi n° 2021-1485 du 15 novembre 2021 visant à réduire l'empreinte environnementale du numérique en France*, publiée au JORF n°0266 du 16 novembre 2021.

écologique du numérique[1336]. Il semble bien que le sujet commence à être pris au sérieux par les régulateurs, même si les mesures concrètes se font toujours attendre.

621. **La réticence des entreprises du secteur privé.** Les entreprises du secteur privé se mobilisent également face à l'enjeu environnemental que soulève le numérique et expriment leurs craintes. À titre d'exemple, le MEDEF reconnaît « *l'intérêt partagé de la protection de l'environnement et de la compétitivité des entreprises* » et a déjà proposé, dès 2017, 40 propositions pour moderniser et simplifier le droit de l'environnement[1337]. Si ces propositions sont pour la plupart dépassées ou non applicables au numérique, elles témoignent néanmoins de la crainte du secteur privé de subir des contraintes trop importantes, malgré une prise de conscience de la nécessité de protéger l'environnement.

622. **L'écoconception déjà explorée dans les grandes entreprises.** De grandes entreprises appliquent déjà l'écoconception à leurs systèmes numériques. C'est le cas d'EDF S.A. qui travaille au déploiement d'un programme « Numérique responsable ». L'écoconception dans le numérique consiste en la prise en compte des impacts environnementaux à toutes les étapes de la conception d'un système, d'une application, d'un service ou d'un produit numérique. Ces étapes peuvent concerner des choix matériels (fabrication des composants électroniques), des choix relatifs à la programmation informatique (choix de l'algorithme ou des données à utiliser) ou encore des choix relevant de l'utilisation finale du produit ou service (plus il y aura d'utilisateurs finaux, plus la capacité de calcul ou l'espace de stockage nécessaire sera important). L'écoconception consiste à analyser ces étapes et à identifier pour chacune d'entre elles ou les choix les plus respectueux de l'environnement, tout en préservant la performance et la qualité du produit ou service concerné. Des exemples d'application de cette démarche, souvent promue sous le nom de « *Green IT* », peuvent être trouvés dans plusieurs grandes entreprises[1338].

[1336] S. Mandard, « Le gouvernement présente un plan sans contrainte pour limiter l'impact écologique du numérique », *Le Monde*, 24 février 2021.

[1337] E. Choux, O. Viano, « Livre blanc du MEDEF : 40 propositions pour moderniser et simplifier le droit de l'environnement dans l'intérêt partagé de la protection de l'environnement et de la compétitivité des entreprises », *Bulletin du Droit de l'Environnement Industriel*, 1er juin 2017, n°69.

[1338] H. D'Agrain, « Du Green IT au Green by IT : exemples d'applications dans les grandes entreprises », *CIGREF*, janvier 2017, disponible en ligne : <https://www.cigref.fr/wp/wp-content/uploads/2017/01/CIGREF-Du-Green-IT-au-Green-by-IT-2017.pdf>, consulté le 11 mars 2020.

623. **Transition.** S'il est selon nous pertinent de réfléchir à l'application d'une logique de précaution au développement du numérique pour limiter ses risques environnementaux, elle doit être raisonnée afin de ne pas avoir pour effet de freiner l'innovation, mais plutôt pour promouvoir un développement technologique vertueux sur le plan environnemental.

B/ Une nécessaire conciliation entre précaution et innovation

624. **Les critiques du principe de précaution.** Le principe de précaution fait l'objet de nombreuses critiques. Le sens même du mot « précaution » induit dans l'esprit du public une connotation négative. Ainsi, le principe de précaution serait synonyme de « *frilosité* »[1339] face à l'innovation, voire de « *technophobie* »[1340], car, selon certains, il « *freinerait l'innovation en imposant des contraintes avant même d'être certain des risques qu'elle implique* »[1341]. En effet, dans la logique de ses détracteurs, le principe de précaution constituerait un frein disproportionné à l'innovation puisque, si le risque zéro n'existe pas, alors il impliquerait de se détourner de la plupart des innovations. Toutefois, ce raisonnement repose sur un sens commun donné au terme précaution et ne révèle pas sa réalité. Ses conditions d'application, restrictives, garantissent justement la mesure et la proportion dans les contraintes à imposer face à un risque incertain, qui doit lui-même être qualifié après une évaluation rigoureuse des risques. Dès lors, on peut lui reprocher le fardeau administratif qu'il implique (réalisation des évaluations, études d'impact des mesures…) mais pas le fait de contraindre aveuglément toute innovation issue du développement technologique.

625. **Une proportionnalité apportée par les modalités d'application du principe de précaution.** Le contenu du principe de précaution présente des garanties permettant d'éviter d'en faire un frein systématique au développement technologique.

626. **Une évaluation des risques préalable.** En effet, la logique de précaution implique premièrement la réalisation d'une évaluation des risques préalable et rigoureuse[1342]. Cette étude

[1339] D. Bourg, « L'éco-scepticisme et le refus des limites », *Études*, 2010, vol. 7, n°413, p. 29.
[1340] D. Lecourt, *Technophobie*, Cités, 2000, p. 15.
[1341] E. Gaillard, *op. cit.*, 3.
[1342] Voir Supra, 597.

doit permettre de mieux comprendre la nature des risques liés à l'innovation concernée. Si l'état des connaissances scientifiques ne permet pas de qualifier la réalité ou la mesure du risque avec certitude, alors le principe de précaution trouvera à s'appliquer. Si, en revanche, cette évaluation permet d'affirmer que le risque est avéré mais que seul persiste un doute sur le moment de sa réalisation, alors le principe de prévention prendra le relais et aboutira également à la prise de mesure pour le limiter[1343]. Si cette évaluation ne permet pas de suspecter un risque ou d'en établir la réalité, ni le principe de précaution, ni le principe de prévention ne sauraient être appliqués. Cette première étape constitue déjà une première garantie.

627. **L'organisation d'un débat public.** La logique de précaution implique également, dans un deuxième temps, l'organisation d'un débat public sur la gestion du risque concerné. En l'occurrence, il s'agirait d'un débat sur la gestion du risque environnemental induit par le développement de l'IA. Or, au vu de la prise de conscience de l'enjeu à la fois par les institutions, le secteur public ou le grand public, un tel débat aurait de grande chance d'aboutir à un consensus sur la conduite à tenir dans la gestion du risque. De plus, les intérêts parfois contradictoires des parties prenantes doivent être conciliés, ce qui ne peut être fait que par le dialogue.

628. **L'adoption de mesures proportionnées et provisoires.** Après avoir défini la politique de gestion du risque face à l'activité dangereuse concernée, le principe de précaution impose l'adoption de mesures proportionnées et provisoires. La proportion des mesures vise précisément à ne pas constituer un frein disproportionné au développement technologique et permettrait la prise en compte des intérêts des acteurs de la recherche en IA. Le caractère éphémère des mesures à prendre dans une logique de précaution semble également approprié au développement de l'IA. Des contraintes pour limiter les conséquences environnementales ne devraient être adoptées que dans l'attente de disposer de plus de données sur l'ampleur de ces conséquences, ou d'avancées sur la compréhension des systèmes d'IA. De telles mesures ne seraient plus justifiées si les connaissances techniques permettaient à l'avenir de quantifier avec certitude les émissions liées à la conception et l'opération de ces systèmes et de développer des

[1343] Sur l'application du principe de prévention, voir : Y. Petit, *op. cit.*, 86-89.

systèmes frugaux, dont l'empreinte environnementale passerait sous le seuil d'acceptabilité défini lors dans la politique de gestion du risque définie à l'étape précédente.

629. **Transition.** Plus qu'à une application juridique *stricto sensu*, c'est surtout à l'esprit ou la philosophie du principe de précaution qu'il faut avoir recours pour bâtir une régulation écologique de l'IA. En effet, quand bien même les conditions du principe de précaution ne seraient pas remplies, des mesures visant à limiter l'impact environnemental de cette technologie doivent tout de même être prises : le principe de « prévention » face aux risques avérés prendrait alors le relais de la « précaution » face aux risques suspectés, mais les mesures proposées ci-après nous semblent indispensables quelle que soit l'approche retenue. Si cette application est possible, il n'en reste pas moins qu'elle doit être concrétisée dans les textes : Quelles mesures concrètes peuvent être prises pour limiter l'empreinte environnementale de l'IA ? Où et par qui ces mesures devraient être adoptées ? Peut-on les intégrer au cadre règlementaire en cours de construction à l'échelle européenne ? C'est à ces questions que se consacre la Section suivante.

Section 2 : La création d'un cadre juridique ambitieux en faveur d'une IA écologique

630. **L'absence de mesures concrètes pour assurer la sobriété écologique du développement de l'IA.** Peu d'auteurs ont été amenés à proposer des mesures pour limiter l'empreinte environnementale du numérique. Le risque environnemental que soulève le développement de l'IA nécessite à la fois d'accélérer la recherche et la transparence sur l'impact sur l'environnement de cette technologie, et d'assurer sa frugalité *via* une logique de sobriété numérique.

631. **La proposition de régulation environnementale dans le rapport Villani sur l'IA en 2018.** Si le sujet est déjà évoqué dans des rapports institutionnels antérieurs[1344], le rapport Villani de mars 2018 nous donne des premières pistes pour une régulation environnementale de l'IA, lesquelles ont été commentées et complétées en doctrine[1345], fondant une théorie de régulation de « l'écologie digitale »[1346]. Elle se fonde sur un tryptique célèbre en droit de l'environnement sous le nom des « principes aarhusiens »[1347] : l'information du public, la participation du public et l'accès à la justice.

632. **Le contenu des principes aarhusiens comme source d'inspiration.** Le premier de ces principes impose d'une part de garantir un droit d'accès du public à l'information en matière d'environnement et, d'autre part, une obligation pour les autorités publiques de collecter et diffuser des informations en matière d'environnement[1348]. Le deuxième vise à impliquer plus en profondeur le public dans les prises de décisions environnementales, notamment en prévoyant l'intervention d'une audition publique préalable à la prise de décision ou l'extension du processus participatif à l'élaboration instruments règlementaires[1349]. Enfin, le dernier de ces principes fondamentaux en droit de l'environnement vise à garantir l'effectivité du contentieux

---

[1344] ADEME, *La face cachée du numérique*, *op. cit.*
[1345] L. Calandri, *op. cit.*
[1346] C. Villani, *op. cit.*, p. 124.
[1347] Principes contenus dans la *Convention d'Aarhus*, 25 juin 1998, ratifiée par la France le 8 juillet 2002 par la *Loi n° 2002-285 du 28 février 2002 autorisant l'approbation de la convention d'Aarhus*, publiée au JORF le 1er mars 2002, p. 3904.
[1348] B. Drobenko, « La Convention d'Aarhus et le droit français », *Revue Juridique de l'Environnement*, numéro spécial, 1999, p. 37.
[1349] *Ibid*, p. 42.

de l'environnement. Cette effectivité passe par un large accès du public à la justice avec des possibilités de recours suffisants, des procédures devant être « objectives, équitables et rapides »[1350], et la levée des obstacles, notamment financiers, que le public pourrait rencontrer pour accéder à la justice[1351].

633. **La transposition des principes aarhusiens à la régulation de l'IA.** Ces trois principes semblent pouvoir constituer une base de départ pour un cadre de régulation écologique à l'IA. Toutefois, il en ressort plusieurs grandes idées devant être traduite dans la règlementation à venir, notamment celle de la transparence (l'information du public, condition à sa participation) et la responsabilité des parties prenantes (derrière « l'accès à la justice »). À ces grandes idées s'ajoute celle découlant de l'application du principe de précaution : la minimisation de l'empreinte environnementale. Les propositions présentées dans cette Section s'inspirent en partie des pistes contenues dans le rapport Villani. Pour autant, elles sont complétées et refondues dans un nouveau cadre plus adaptable à la règlementation en cours de construction à l'échelle européenne.

634. **L'absence de mesures de régulation environnementale dans le projet d'*AI Act*.** L'épineuse question de l'empreinte environnementale des systèmes d'IA figure déjà dans les nombreux rapports ou textes institutionnels antérieurs à la proposition de l'*Artificial Intelligence Act*. Pourtant, aucune mesure concrète pour répondre à ce problème ne figure dans les premières propositions de régulation par la Commission européenne[1352]. Cette dimension environnementale doit selon nous être intégrée au cadre règlementaire en cours de construction à l'échelle européenne. Ce dernier prévoit la mise en place de processus de conformité *ex ante* pour les systèmes d'IA « à haut risque », comprenant notamment une documentation exhaustive ou la conservation des données d'apprentissage[1353]. Il est regrettable qu'aucune de ces obligations ne vise à minimiser l'impact environnemental du système concerné. Les mesures proposées dans la présente Section pourraient être intégrées à ces obligations de conformité *ex*

[1350] Convention d'Aarhus, *op. cit.*, article 9.
[1351] B. Drobenko, *op. cit.*, p. 53.
[1352] COMMISSION EUROPÉENNE, *Proposition de règlement du Parlement européen et du Conseil établissant des règles harmonisées concernant l'intelligence artificielle (législation sur l'intelligence artificielle) et modifiant certains actes législatifs de l'Union*, 21 avril 2021, 2021/0106 (COD).
[1353] *Ibid.*, articles 10 et suivants.

ante, autoévaluées par les destinataires de la norme. En revanche, ces mesures ne devraient pas être restreintes aux seuls systèmes d'IA « à haut risque » tels que définis dans la proposition de règlement[1354] mais étendues à tout système d'IA mobilisant une capacité de calcul, ou de stockage, suffisamment importante pour être considérée comme « à risque » d'un point de vue environnemental. L'idée étant de ne pas pénaliser de petits acteurs qui développeraient des applications peu énergivores (applications mobiles à public restreint…), un seuil pourrait être défini par un groupe d'experts. Ainsi, un système d'IA pourrait être qualifié d'à « haut risque environnemental » si la consommation énergétique générée par l'ensemble de son cycle de vie dépassait le seuil de l'acceptable. On pourrait également envisager que le recours à certains métaux rares dans la conception du support matériel à l'IA fasse basculer d'office le système dans cette nouvelle catégorie. Enfin, pour être effectives, les mesures proposées doivent faire l'objet d'un contrôle *a posteriori*, déjà prévu dans le cadre règlementaire européen pour l'IA avec des sanctions en cas de non-conformité, et la possibilité d'un contrôle externe par une autorité de supervision[1355].

635. **Plan.** La régulation environnementale de l'IA pourrait donc passer par l'intégration de mesures dans la règlementation européenne en cours de construction. Cette régulation peut être fondée, selon nous, sur quatre principes directeurs, découlant des constats effectués dans les parties précédentes.

Premièrement, la régulation proposée se fonde sur l'application d'un **principe de précaution** visant à l'adoption immédiate de mesures pour limiter le risque environnemental de l'IA, sans attendre de certitudes scientifiques sur son ampleur. Si la qualification du risque n'est pas incertaine mais avérée, le principe de prévention prendrait le relais du premier pilier de notre proposition. Ces principes incluent, dans leur application, un principe de proportionnalité, lequel obligera à toujours rechercher le compromis conciliant limitation du risque environnemental et promotion de la dynamique d'innovation. Ce premier pilier a déjà été présenté dans la Section précédente.

[1354] *Ibid.*, article 6 et Annexe III.
[1355] *Ibid.*, Titre VIII.

Deuxièmement, le cadre proposé devrait inclure un **principe de transparence** visant au partage (aux autorités de régulation et au public) de toute l'information pertinente relative aux conséquences environnementales liées au développement des systèmes d'IA par les entreprises (**§1**).

Troisièmement, un **principe de minimisation** de l'empreinte environnementale, visant à imposer des obligations de frugalité ou de sobriété, devrait être imposé aux développeurs ou opérateurs d'IA (**§2**).

Enfin, quatrièmement, il conviendrait de garantir un **principe de responsabilité** des acteurs à travers un contrôle effectif du respect de la norme et la garantie de l'accès à la justice pour le public (**§3**).

§1 : La nécessité d'assurer la transparence de l'impact environnemental des systèmes d'IA

636. **Le constat du défaut d'information du public sur l'empreinte environnementale de l'IA.** Le rapport Villani constate, dès 2018, « *une carence en matière d'information du public car il n'existe pas aujourd'hui de politique claire en faveur de l'évaluation écologique des solutions numériques* »[1356]. Le constat est similaire outre-Atlantique[1357] : les données manquent pour quantifier l'impact environnemental des systèmes d'IA, et accélérer la recherche sur l'écoconception et le développement de systèmes d'IA « écologiques ».

637. **Plan.** Les entreprises pourraient être incitées, voire contraintes, à quantifier l'impact environnemental des systèmes d'IA qu'elles déploient (**A**). Cela permettrait aux régulateurs de concentrer leurs efforts sur les systèmes les plus énergivores, à la recherche de progresser dans la sobriété numérique et au grand public de prendre conscience du risque invisible du développement de l'IA pour l'environnement, encore faut-il que cette information soit transparente (**B**).

[1356] C. Villani, *op. cit.*, p. 127.
[1357] A.L. Stein, *op. cit. ;* P. Dhar, *op. cit.*

A/ La quantification de l'empreinte environnementale des systèmes d'IA comme prérequis à la régulation

638. **Pourquoi quantifier ?** Le manque de données concernant l'impact environnemental des systèmes d'IA a de nombreuses conséquences indésirables. Tout d'abord, cette lacune rend invisibles les conséquences de l'IA sur l'environnement qui, pourtant sont réelles comme le révèlent plusieurs études[1358]. Ensuite, il rend impossible toute comparaison entre différents systèmes, voire entre différentes entreprises : chacun peut se dire le plus écologique étant donné que personne ne sait réellement quoi mesurer. Cette carence génère également une réticence des politiques et des investisseurs à encourager la recherche dans le domaine de l'IA. Les systèmes basés sur de l'apprentissage automatique sont extrêmement énergivores mais personne n'est en mesure de garantir que leurs conséquences environnementales sont mineures, alors les responsables politiques et les investisseurs ne peuvent prendre le risque de les financer ou les subventionner massivement. Enfin, ce manque de données ralentit la recherche pour développer l'écoconception dans le secteur du numérique, en particulier dans le traitement massif de données ou l'entraînement de modèles d'IA. C'est pourquoi les entreprises doivent être incitées, voire contraintes pour les applications les plus énergivores, à quantifier cette empreinte.

639. **Comment quantifier ?** Si la nécessité de quantifier l'empreinte environnementale des systèmes d'IA fait consensus et justifie l'adoption de mesures pour l'imposer, cette obligation ne devrait être envisagée que si elle s'accompagne de normes à suivre. En effet, les outils visant à estimer les émissions liées à un modèle d'IA, son entraînement ou son opération existent mais ils manquent de standardisation. La construction de ces normes devrait, conformément aux principes du droit de l'environnement, faire l'objet d'une participation des parties prenantes et du public. Il sera également important d'estimer les éventuelles émissions de gaz à effet de serre, mais également les besoins matériels en métaux rares ou autres composants dont l'extraction ou l'utilisation peuvent avoir des conséquences sur l'environnement. Un tel

[1358] A.L. Stein, *op. cit.*

« baromètre »[1359] standardisé et construit à partir de consensus scientifique et industriel permettrait de faciliter la mise en conformité des entreprises et rendrait possible la comparaison de différents systèmes (le cas-échéant pour choisir le plus écologique). Ce travail de standardisation et la coordination du débat sur sa construction pourraient se faire par une autorité indépendante telle que l'Autorité de la transition écologique en France (ex-ADEME). On ajoutera enfin l'importance de la prise en compte de tout le cycle de vie et de l'entière chaîne d'approvisionnement dans la quantification de l'empreinte d'un système, suivant la méthode d'Analyse du Cycle de Vie, ou « *ACV* »[1360], bien connue des écologues.

640. **L'existence d'outils pour aider à la quantification de l'empreinte environnementale de l'IA.** Il existe déjà de nombreuses initiatives isolées, visant à proposer des outils et standards de quantification de l'empreinte environnementale des systèmes d'IA. On pourra ici citer le développement de la librairie Python CodeCarbon[1361], les guides proposés par Green Concept[1362] ou encore les travaux du collectif « Conception Numérique Responsable »[1363]. Capitaliser sur ces démarches existantes est essentiel. Cela permettrait d'accélérer le déploiement de l'écoconception et de diffuser rapidement des référentiels d'évaluation des solutions numériques. Nécessitant la coopération des acteurs et leur transparence sur les composants, les matériaux et autres données, une telle démarche n'est envisageable que si elle accompagne la mise en œuvre d'obligations contraignantes pour les entreprises.

641. **La quantification de l'empreinte environnementale du numérique sur le modèle du bilan « GES » de la norme ISO 14 064-1.** Des travaux de standardisation similaires à ceux

[1359] Reprenant en partie la proposition du Rapport Villani, C. Villani, *op. cit.*

[1360] ADEME, *L'analyse du cycle de vie*, 18 juin 2018, disponible en ligne : <https://www.ademe.fr/expertises/consommer-autrement/passer-a-laction/dossier/lanalyse-cycle-vie/quest-lacv>, consulté le 7 mai 2021.

[1361] Boston Consulting Group, « Des experts de haut niveau en IA lancent CodeCarbon, un outil révolutionnaire de réduction des émissions de CO2 liées à l'informatique », *BCG (blog)*, 1 décembre 2020, disponible en ligne : <https://www.bcg.com/fr-fr/press/1december2020-top-ai-experts-launch-codecarbon>, consulté le 7 mai 2021 ; K. Goyal, « AI Computing Emits CO_2. We Started Measuring How Much », *Medium (blog)*, 30 novembre 2020, disponible en ligne : <https://medium.com/bcggamma/ai-computing-emits-co%E2%82%82-we-started-measuring-how-much-807dec8c35e3>, consulté le 7 mai 2021.

[1362] Green Concept, *Livre blanc sur l'écoconception numérique*, 21 février 2020, disponible en ligne : <http://www.greenconcept-innovation.fr/wp-content/uploads/2020/02/greenconcept_21022020.pdf>, consulté le 7 mai 2021, présentant une liste de 45 actions préconisées aux entreprises pour réduire les impacts environnementaux de leur service numérique.

[1363] Collectif Numérique Responsable, site officiel, disponible en ligne : <https://collectif.greenit.fr/>, consulté le 7 mai 2021.

que l'on propose ont déjà été réalisés dans le cadre des bilans d'émissions de gaz à effet de serre (« GES »). Ces travaux ont abouti à la publication de la norme ISO 14 064-1[1364], spécifiant les principes et les exigences applicables aux entreprises pour la quantification et la rédaction de rapports sur les émissions et suppressions de gaz à effet de serre (GES). Les principes contenus dans cette norme pourraient servir de modèle pour la création de standards pour la quantification de l'empreinte environnementale du numérique. En effet, aux termes de cette norme, la réalisation d'un bilan « GES » consiste en six étapes[1365] :

- La préparation vise à définir les modalités de réalisation du bilan des émissions de gaz à effet de serre, son périmètre organisationnel et opérationnel, l'année de référence ou encore les méthodes de calcul qui seront employées. Le point le plus important dans cette étape est la définition du périmètre du bilan. La norme ISO 14 064-1 impose en effet de définir dans un premier temps le périmètre « organisationnel » et de répondre à la question « Quelles sont les installations concernées par l'étude ? »[1366]. Dans le cadre du numérique et de l'IA, cette étape consisterait à définir quels systèmes rentrent dans le champ du bilan environnemental. Logiquement, tous les systèmes développés, vendus, utilisés ou maintenus par l'entreprise réalisant le bilan devraient rentrer dans le champ de l'analyse. La norme impose ensuite de définir le périmètre « opérationnel » du bilan « GES », correspondant aux catégories et postes d'émissions liées aux activités du périmètre organisationnel (dans notre cas, les systèmes rentrant dans le champ de l'étude). Il existe trois catégories d'émissions qu'il est absolument nécessaire de prendre en compte dans la quantification de l'empreinte environnementale du numérique : les émissions directes de GES (Scope 1), les émissions indirectes liées à l'énergie (Scope 2) associées à la production d'électricité ou d'énergie nécessaires au fonctionnement de l'activité (ou des systèmes) concernés, et les autres émissions indirectes (Scope 3) produites par les activités, telles que l'achat des matières premières nécessaires pour la

[1364] ISO, *ISO 14064-1:2018, « Gaz à effet de serre — Partie 1: Spécifications et lignes directrices, au niveau des organismes, pour la quantification et la déclaration des émissions et des suppressions des gaz à effet de serre »*.
[1365] ADEME, « Etapes d'un bilan GES », *Bilan GES – Site de l'ADEME (blog)*, disponible en ligne : <https://www.bilans-ges.ademe.fr/fr/accueil/contenu/index/page/Etapes%2Bbilan%2BGES/siGras/0>, consulté le 8 octobre 2021.
[1366] ADEME, « Bilan GES – Organisation », *Site de l'ADEME (blog)*, disponible en ligne : <https://www.bilans-ges.ademe.fr/fr/accueil/contenu/index/page/bilan%2Bges%2Borganisation/siGras/1>, consulté le 8 octobre 2021.

fabrication des produits ou encore les émissions liées à l'utilisation qui est faite des produits ou services vendus[1367]. Dans le cadre du numérique, les émissions directes sont généralement limitées (pas de combustion générant du gaz,…) ce qui explique « l'invisibilisation » de son empreinte environnementale. C'est pourquoi il est primordial de prendre en compte les autres scopes d'émissions, recouvrant à la fois les émissions liées à la consommation énergétique et son origine, ainsi que les émissions générées par l'utilisation finale des systèmes numériques ou d'IA.

- La collecte des données est l'étape essentielle du processus puisqu'elle consiste en le recensement des données pertinentes, disponibles en interne (les « données primaires ») ou qu'il faut solliciter auprès de fournisseurs, partenaires ou clients (les « données secondaires »). Cela peut être des données de consommation électrique, sur son origine, sur les matériaux utilisés pour la fabrication des terminaux matériels, sur l'utilisation des systèmes par les clients,… Ces données peuvent ensuite être entrées dans la Base Carbone de l'Ademe pour déterminer leur équivalent carbone[1368]. Cette base est une source de données centralisée et la base de référence de l'article L229-25 du Code de l'environnement qui prévoit l'exigence de réalisation de bilan GES pour certaines entreprises et les modalités de son contrôle. Il pourrait s'avérer pertinent d'enrichir cette base avec des postes d'émissions plus précis en lien avec le numérique, dont certains matériaux rares utilisés pour la fabrication des terminaux physiques.

- Le calcul du bilan GES, sur la base des données collectées, de leur équivalent carbone et, le cas-échéant, des données récoltées sur l'utilisation des systèmes par les clients. Ainsi, un système d'IA utilisé aux fins d'identifier de nouveaux puits de pétrole contribuera à la réalisation de nouvelles émissions indirectes, tandis qu'un système utilisé pour réaliser des économies d'énergie en économisera autant.

[1367] Les différentes catégories d'émissions sont découpées en 23 « postes d'émission », qui correspondent aux sources et causes précises des émissions (AFNOR, *ISO-TR 14069, « Guide d'application de la norme 14064-1 WD3 »*, Mars 2011). Pour une présentation simplifiée du contenu du périmètre opérationnel, voir ADEME, « Bilan GES Organisation », *Bilan GES – Site de l'ADEME (blog)*, *op. cit.*

[1368] ADEME, « Base Carbone », *Bilan GES – Site de l'ADEME (blog)*, disponible en ligne : <https://www.bilans-ges.ademe.fr/fr/accueil/contenu/index/page/presentation/siGras/0>, consulté le 8 octobre 2021.

- La présentation du bilan GES aux décideurs, partenaires et clients pour les interpeler, les sensibiliser et engager la discussion sur de potentielles actions de réduction.
- La planification des actions de réduction.
- La publication du bilan et du plan d'action, obligatoire pour certaines organisations au titre de l'article L229-25 du Code de l'environnement, dans une logique de transparence. Au vu des enjeux de l'empreinte environnementale du numérique et des risques liés à son « invisibilité » pour les individus, il semble approprié d'imposer un certain degré de transparence aux acteurs du numérique pour permettre un contrôle effectif des technologies numériques à la fois par les pouvoirs publics et par la société dans son ensemble.

642. **Conclusion et transition.** La méthodologie du bilan GES présentée dans la norme ISO 14064-1 et dont la mise en œuvre est accompagnée par l'ADEME nous semble particulièrement pertinente pour quantifier l'empreinte environnementale du numérique ou de l'IA. En effet, la prise en compte des émissions indirectes de scope 2 et 3 est essentielle pour pallier l'absence d'émissions directes et ne plus occulter la « face cachée du numérique »[1369]. La quantification de cette empreinte est une première étape indispensable pour pouvoir gérer les risques environnementaux générés par le développement de l'IA et des technologies numérique. Une deuxième étape consiste en une démarche de transparence, indispensable si l'on souhaite pouvoir comparer et identifier des applications créant un risque environnemental trop important.

[1369] ADEME, *La face cachée du numérique – Réduire les impacts du numérique sur l'environnement*, janvier 2021, p. 3.

B/ La mise en œuvre d'un principe de transparence environnementale dans la régulation de l'IA

643. **Plan.** La reconnaissance d'un principe de transparence nous apparaît justifiée dans le cadre de la régulation de l'IA (**1**). Les moyens de sa mise en œuvre (**2**) et son contenu (**3**) devront être précisés.

1. L'utilité de la transparence de l'empreinte environnementale de l'IA

644. **La nécessaire divulgation d'informations requise pour évaluer l'empreinte écologique des systèmes d'IA.** L'empreinte environnementale, si elle est quantifiée suivant un standard commun, peut devenir un critère d'évaluation des systèmes d'IA, au même titre que leur précision ou leur intelligibilité. Aujourd'hui, la plupart des publications scientifiques dans le domaine de l'apprentissage automatique ne comprennent aucune information concernant la consommation énergétique nécessaire à l'entraînement du modèle alors que cette technique figure parmi les plus énergivores comme cela a été démontré dans plusieurs études[1370].

645. **L'utilité de la transparence pour les autorités de régulation.** Afficher l'empreinte environnementale de ces systèmes aurait d'abord un intérêt vis-à-vis des autorités de régulation. Ces dernières pourraient plus facilement identifier et encadrer les systèmes « à haut risque environnemental ». Le choix du seuil d'émissions à partir duquel un système passerait dans cette catégorie devrait faire l'objet d'un choix politique éclairé par des études scientifiques rigoureuses. Les systèmes qui entreraient dans cette qualification pourraient par exemple se voir imposer de plus grandes contraintes ou une supervision accrue de la part d'une autorité compétente, mais cela n'est possible que si l'empreinte environnementale est à la fois quantifiée et transparente.

646. **L'utilité de la transparence pour le milieu de la recherche.** Ensuite, la transparence environnementale des systèmes d'IA pourrait faciliter la recherche et le développement de

[1370] P. Dhar, *op. cit.* ; E. Strubel, « Energy and Policy Considerations for Deep Learning in NLP », *op. cit.*, p. 4.

méthode d'écoconception. Les chercheurs auraient à leur disposition de nombreuses données de consommation liées à chaque technique employée dans le domaine de l'IA, et pourraient donc travailler à la minimisation de leur empreinte environnementale.

647. **L'utilité de la transparence pour le public.** Enfin, l'affichage de l'empreinte des systèmes d'IA aurait un intérêt pour le grand public. De la même façon que l'industrie agroalimentaire a adopté des labels sur la qualité nutritionnelle des produits, ou que la performance énergétique des bâtiments doit faire l'objet d'une évaluation transparente, un baromètre environnemental des systèmes d'IA permettrait au consommateur de prendre conscience de l'impact écologique de la solution qu'il va utiliser. Si le consommateur est le premier concerné, la démarche aurait en réalité de l'intérêt pour tout futur utilisateur de logiciels basés sur de l'IA, y compris les entreprises, qui pourrait choisir entre différentes solutions en fonction de leur impact environnemental.

2. Les moyens de la transparence de l'empreinte environnementale de l'IA

648. **La transparence à l'égard d'une autorité de contrôle.** D'abord, la transparence environnementale doit permettre aux pouvoirs publics de disposer d'un plus grand nombre d'information sur l'empreinte écologique de l'IA. L'autorité de contrôle compétente en matière d'IA pourrait ainsi organiser une remontée d'informations en imposant la publication de rapports dont le contenu serait standardisé. Il est aussi possible d'imposer une obligation de coopération aux entreprises développant des systèmes d'IA visant à les contraindre à répondre à toutes les requêtes de l'autorité dans le cadre de ses missions de quantification de l'empreinte environnementale de l'IA.

649. **La transparence dans les communications scientifiques.** L'empreinte environnementale des systèmes d'IA pourrait ensuite être affichée dans les publications et communications scientifiques, pour sensibiliser le milieu académique à la problématique et accélérer la recherche en la matière. Chaque publication relative à un algorithme d'IA pourrait comprendre des informations relatives à son empreinte écologique. Ces informations pourraient

également devenir un critère vérifié par les organisateurs de manifestations scientifiques ou les éditeurs de revue[1371].

650. **La transparence dans les relations commerciales.** L'empreinte environnementale de l'IA pourrait également figurer au titre de l'obligation d'information précontractuelle pour la fourniture de biens ou de services[1372], et ce afin d'éclairer le choix d'un consommateur lorsque ce dernier souhaiterait avoir recours à un service numérique basé sur de l'IA ou acheter un produit intégrant de l'IA.

651. **La transparence centralisée dans une base de données librement accessible.** De plus, comme le rapport Villani le proposait en 2018, une base de données pourrait être mise à disposition sur un site Internet, national ou européen, permettant « *de comparer l'impact écologique des différents produits et services, logiciels, hardware, impliqués dans la chaîne de valeur du numérique. Ce site devra s'appuyer sur une base de données permettant d'évaluer l'impact environnemental de tous les aspects de la dématérialisation à l'œuvre via le numérique, tant pour les particuliers (impact de la recommandation personnalisée, des chatbots, des techniques de reconnaissance d'image...) que pour les entreprises, afin de leur permettre d'évaluer leurs fournisseurs numériques.* »[1373]. La base de données des systèmes « à haut risque » qui sera créé dans le cadre de la règlementation européenne sur l'IA pourrait parfaitement jouer ce rôle[1374], à condition qu'elle soit accessible par le grand public.

652. **La transparence dans la documentation obligatoire au titre de la future règlementation européenne de l'IA.** L'évaluation de l'empreinte environnementale pourrait enfin être imposée dans le cadre de la procédure de gestion des risques que doivent mettre en place les fournisseurs d'IA « à haut risque » conformément à l'article 9 du projet de règlement

[1371] La célèbre conférence scientifique sur l'IA « NeurIPS » a d'ailleurs intégré dans le formulaire de proposition d'articles des questions relatives à l'impact écologique des travaux de recherche présenté : https://neurips.cc/public/guides/PaperChecklist.
[1372] Code de la Consommation, article L221-5 et s.
[1373] C. Villani, *op. cit.*, p.127.
[1374] COMMISSION EUROPÉENNE, *Proposition de règlement du Parlement européen et du Conseil établissant des règles harmonisées concernant l'intelligence artificielle (législation sur l'intelligence artificielle) et modifiant certains actes législatifs de l'Union*, 21 avril 2021, 2021/0106 (COD), article 60.

européen sur l'IA[1375]. Enfin, l'empreinte environnementale ainsi quantifiée pourrait utilement figurer dans la documentation technique accompagnant le système[1376].

3. Le contenu de la transparence de l'empreinte environnementale de l'IA

653. **Un contenu fondé sur la quantification de l'empreinte environnementale de l'IA.** Les émissions de gaz à effet de serre sont le principal risque environnemental lié au développement massif des systèmes d'IA. Dans le cadre de la transparence proposée, ces émissions devraient être quantifiées tout au long du cycle de vie du système concerné. Cela inclut les émissions liées à l'extraction des matières premières nécessaires à la construction des supports matériels du système, les émissions liées à la consommation énergétique pour les phases d'apprentissage et d'inférence, ainsi que les émissions liées au traitement du système en fin de vie (éventuel recyclage…). Concernant la consommation énergétique, la source de l'énergie (souvent de l'électricité) utilisée devrait faire l'objet d'une attention particulière et pourrait être inclue dans les informations soumises à transparence. À cet égard, on pourra citer les préconisations de Greenpeace sur les performances énergétiques du secteur informatique : l'ONG a demandé aux plus grandes entreprises d'Internet de s'engager pour un approvisionnement basé à 100 % sur des énergies renouvelables. Les géants du net – Facebook, Apple et Google – ont ainsi été les premiers à s'y être engagés. Aujourd'hui Greenpeace demande à d'autres entreprises du secteur (Amazon, Twitter, Netflix) de se joindre au mouvement, mais aussi aux géants asiatiques, notamment Alibaba[1377]. La transparence ne saurait être complète si elle n'inclut pas la source de l'énergie utilisée pour approvisionner les systèmes d'IA.

654. **Transition.** Cette transparence, couplée avec une obligation de minimisation de l'impact environnemental des systèmes d'IA, responsabiliserait les entreprises et les contraindrait à se détourner des modèles trop gourmands en énergie.

[1375] *Ibid.*, article 9
[1376] *Ibid.*, article 10 et Annexe IV « Technical documentation ».
[1377] GREENPEACE, *Clicking Clean*, Rapport, 2017, disponible en ligne : <http://www.clickclean.org/france/fr/>, consulté le 8 avril 2021, p. 8-13.

§2 : La nécessité de minimiser l'empreinte environnementale de l'IA

655. **Des contraintes justifiées par la logique de précaution.** Pour rappel, l'instauration d'un cadre juridique contraignant pour promouvoir le développement écologique de l'IA résulte de l'application d'une logique de précaution, visant à la prise de mesures immédiates pour limiter un risque aux conséquences mal connues. La première de ces mesures est la transparence sur l'empreinte environnementale des systèmes d'IA. Cette dernière ne peut suffire à limiter le risque écologique né des émissions liées à la conception et au déploiement de systèmes basés sur de l'IA.

656. **Plan.** C'est pourquoi nous proposons dans ce Paragraphe la reconnaissance d'une véritable obligation de sobriété numérique, à la charge des développeurs de systèmes d'IA (**A**). Cette obligation de moyens imposerait aux entreprises de prendre des mesures techniques et organisationnelles pour limiter l'impact environnemental des systèmes qu'elles développent ou utilisent. Si elle s'avère nécessaire au vu de l'ampleur potentielle du risque environnemental causé par l'IA, cette obligation n'en reste pas moins contraignante et pourrait par conséquent être crainte par les éditeurs d'IA. C'est pourquoi les entreprises devraient être dûment accompagnées et la mise en conformité dûment supervisée par une autorité compétente et dotée des moyens humains et techniques nécessaires (**B**).

A/ Une obligation de minimisation de l'impact environnemental des systèmes d'IA

657. **Les leviers pour limiter l'impact environnemental des systèmes d'IA.** Le but de la régulation environnementale de l'IA est de limiter son empreinte écologique. Pour ce faire, les entreprises pourraient se voir imposer un principe de minimisation, reprenant l'idée générale présente dans le RGPD. Ce dernier impose, lors du traitement de données à caractère personnel, que le responsable de traitement traite uniquement les données adéquates, pertinentes et limitées

à ce qui est nécessaire au regard des finalités pour lesquelles elles sont traitées[1378]. De la même manière, un principe de minimisation environnementale pourrait être consacré dans la régulation de l'IA. Il imposerait aux entreprises de prendre des mesures pour limiter l'impact environnemental de leurs systèmes à ce qui est strictement nécessaire et incompressible au regard de leur finalité. Sous la forme d'une obligation de moyens, dont la complétion serait interprétée par le juge mais ferait l'objet de lignes directrices non contraignantes, cette minimisation pourrait utiliser plusieurs leviers :

- **Le choix des méthodes de programmation** : il existe des langages de programmation et des méthodes de codage permettant de diminuer le besoin en capacité de calcul, et par conséquent la consommation d'énergie y relative.

- **La minimisation de la quantité de données utilisées** (qu'il s'agisse de données à caractère personnel ou non) : le stockage et le traitement de grandes quantités de données font augmenter la consommation d'énergie des systèmes d'IA. De plus, bien souvent, des quantités trop importantes de données sont traitées par rapport à la finalité poursuivie, aboutissant à un phénomène de surentraînement et à une baisse de performance. C'est pourquoi il serait pertinent que les entreprises soient attentives à n'utiliser que les données « adéquates, pertinentes et limitées à ce qui est nécessaire au regard des finalités pour lesquelles elles sont traitées », reprenant la formule présente dans le RGPD (qui ne s'applique qu'aux systèmes d'IA traitant des données personnelles). Cela allongera sans doute la phase de recherche en amont de l'expérimentation et de l'apprentissage sur les données, mais permettra de réaliser d'importantes économies en termes de consommation énergétique.

- **La limitation du nombre d'entraînements des modèles** : la phase d'apprentissage est particulièrement consommatrice en énergie, comme en témoignent l'exemple du plus grand modèle de compréhension du langage naturel GPT-3[1379] et l'estimation réalisée par le

[1378] *Règlement (UE) 2016/679 du Parlement européen et du Conseil du 27 avril 2016, relatif à la protection des personnes physiques à l'égard du traitement des données à caractère personnel et à la libre circulation de ces données, et abrogeant la directive 95/46/CE*, dit « RGPD », article 5.
[1379] R. Toews, « Deep Learning's Carbon Emissions Problem », *Forbes*, 17 juin 2020, disponible en ligne : <https://www.forbes.com/sites/robtoews/2020/06/17/deep-learnings-climate-change-problem/>, consulté le 7 mai 2021.

Massachussets Institute of Technology[1380]. Une piste prometteuse en termes de sobriété numérique est la limitation du nombre de ces entraînements, d'une part, en s'assurant qu'un même modèle ne soit pas entraîné plusieurs fois de la même manière et, d'autre part, en garantissant un partage de données afin que plusieurs acteurs puissent bénéficier des calculs déjà réalisés, sans avoir à renouveler l'apprentissage[1381].

- **La source de l'électricité utilisée pour alimenter les systèmes lors des phases d'apprentissage et d'inférence** : malgré les chiffres impressionnants des émissions de CO2 liés à la phase d'apprentissage d'un modèle d'apprentissage automatique, plusieurs acteurs du secteur convergent pour dire que cette phase ne génère que 10 à 20% de l'empreinte totale d'un système (les 80% d'émissions restantes émanant principalement de la phase d'inférence). D'où l'intérêt de décarboner l'intégralité de l'approvisionnement en électricité. Les entreprises développant des systèmes d'IA à fort impact environnemental (par exemple dans le domaine du traitement du langage ou *NLP*[1382]) devraient être incitées à avoir recours à des sources d'électricité bas carbone. Le choix des sources à prioriser pourrait se baser sur la taxonomie des énergies « vertes » à l'échelle européenne[1383]. De la même manière que cette taxonomie flèche les investissements financiers vers des énergies considérées comme souhaitables pour atteindre les objectifs environnementaux de l'Union, elle pourrait également servir pour flécher le choix de la source d'énergie dans des secteurs à fort impact environnemental. Une telle incitation pourrait également inciter les acteurs de l'IA à relocaliser leurs serveurs dans des pays au mix énergétique décarboné et donc potentiellement attirer des investissements au sein de l'Union européenne.

- **La performance énergétique des serveurs ou centres de données hébergeant données et modèles** : les entreprises pourraient enfin être incitées à héberger leurs systèmes sur des serveurs ou dans des centres de données ayant une performance énergétique optimale. Des normes existent déjà en la matière, certifiant que le centre de données concerné utilise les

[1380] E. Strubel, « Energy and Policy Considerations for Deep Learning in NLP », *op. cit.*, p. 4.
[1381] Le partage de données est d'ailleurs l'un des objectifs de la stratégie européenne sur les données.
[1382] E. Strubel, *op. cit.*
[1383] *Règlement (UE) 2020/852 du Parlement européen et du Conseil, du 18 juin 2020 sur l'établissement d'un cadre visant à favoriser les investissements durables et modifiant le règlement (UE) 2019/2088*, publié au JOUE n°L198/13 le 22 juin 2020.

meilleures technologies pour limiter son impact environnemental, comme la norme ISO 5001[1384]. Le recours à cette norme pourrait être favorisé et la preuve de cette certification permettrait aux entreprises de prouver leurs efforts pour minimiser l'empreinte environnementale de leurs systèmes informatiques.

658. **Pour une obligation de moyens à destination des développeurs de systèmes d'IA.** L'obligation de moyens imposerait aux entreprises d'être en capacité de démontrer qu'elles ont mis en œuvre des moyens techniques et organisationnels pour limiter l'impact environnemental de leurs systèmes d'IA. Ces moyens pourraient être de la même nature que les leviers évoqués ci-dessus et devraient quoi qu'il arrive faire l'objet de recommandations précises pour guider les entreprises. Ces mesures pourraient faire l'objet d'un critère supplémentaire dans le mécanisme de conformité *ex ante* proposé dans le règlement européen sur l'IA[1385]. Toutefois, comme il ne s'applique qu'aux systèmes « à haut risque pour les droits fondamentaux ou la sécurité », cet examen de conformité ne concernerait pas les applications « à haut risque environnemental » mais sans conséquences pour la sécurité ou les libertés individuelles. C'est pourquoi la réflexion sur le seuil à retenir pour qualifier le « haut risque environnemental » devra être utilisée ici pour définir quelles applications d'IA devront être concernées par ces obligations. Au-delà du champ d'application matériel, le champ d'application territorial doit être, à l'instar de la règlementation européenne sur l'IA, le plus étendu possible et concerner tous systèmes mis sur le marché européen, utilisés ou importés par des entreprises européennes. Enfin, pour ne pas constituer des contraintes disproportionnées à l'innovation, ces obligations doivent faire l'objet d'un fort accompagnement et d'une supervision rigoureuse pour en assurer l'effectivité, possiblement par la nomination ou la création d'une autorité dédiée.

[1384] ISO, *Norme ISO 5001:2018 « Systèmes de management de l'énergie » ;* AFNOR, « EDF obtient la certification ISO 50001 pour son parc de datacenters », *Communiqué de presse*, 19 février 2016, disponible en ligne : <https://www.afnor.org/presse_fev2016/edf-obtient-la-certification-iso-50001-pour-son-parc-de-datacenters/>, consulté le 7 mai 2021.

[1385] COMMISSION EUROPÉENNE, *Proposition de règlement du Parlement européen et du Conseil établissant des règles harmonisées concernant l'intelligence artificielle (législation sur l'intelligence artificielle) et modifiant certains actes législatifs de l'Union*, 21 avril 2021, 2021/0106 (COD), article 16.

B/ Le rôle des autorités publiques dans la régulation environnementale de l'IA

659. **Accompagner la conformité.** Les mesures à prendre pour limiter le risque environnemental né du développement de l'IA, dans une logique de précaution, doivent être proportionnées et provisoires[1386]. Pourtant, les mesures que nous proposons dans cette Section peuvent sembler particulièrement contraignantes, d'autant plus qu'elles s'ajouteraient à toutes les obligations déjà contenues dans la règlementation européenne de l'IA. C'est pourquoi, afin d'éviter la création de freins disproportionnés à l'innovation, une réflexion doit être menée pour faire en sorte que la mise en conformité avec ces obligations environnementales soit la plus simple possible pour les entreprises. Elles devraient être accompagnées dans leur mise en conformité par une autorité compétente et dotée des moyens techniques et humains nécessaires. À l'instar de la CNIL, cette dernière pourrait avoir la charge d'émettre des recommandations, de valider des standards ou des normes, ou de répondre aux interrogations des entreprises pendant un délai de mise en conformité. En France, l'Agence pour la transition écologique (ex-ADEME) apparaît comme une candidate idéale. Cette mission pourrait également être attribuée à la future autorité nommée pour assurer la mise en œuvre de la régulation de l'IA, avec l'appui de l'ADEME pour les missions nécessitant une expertise environnementale (notamment la quantification de l'empreinte environnementale et la publication de standards). Les entreprises devront, en effet, être guidées dans leur mise en conformité en leur permettant de se référer à des normes simples et accessibles. Ces normes pourront concerner les méthodes pour quantifier l'impact environnemental de leurs systèmes ou pour le limiter. De nombreux standards existent déjà tels que la norme ISO 14062 sur l'écoconception[1387], la méthode « ERC » (pour Eviter –

[1386] Voir Supra, 628.

[1387] ISO, *Norme ISO/TR 14062:2002 « Management environnemental — Intégration des aspects environnementaux dans la conception et le développement de produit »*.

Réduire – Compenser)[1388] ou l'analyse du cycle de vie (méthode ACV)[1389] permettant de construire des processus et produits respectueux de l'environnement. L'autorité nommée pourrait avoir la charge de décliner les standards existants aux spécificités du secteur numérique et des technologies d'IA. Cet accompagnement par une autorité externe devrait réduire le poids que constitueraient les obligations environnementales décrites précédemment.

660. **Superviser la conformité.** Pour assurer l'effectivité de l'obligation de minimisation, il est essentiel que l'autorité en charge de l'accompagnement de la conformité des entreprises dispose également de moyens de contrôle et de sanctions, sur le modèle de la CNIL et comme proposé dans la proposition de règlement européen sur l'IA[1390]. Il ne s'agit pas ici de proposer la création d'une nouvelle autorité mais d'ajouter la mission relative à la régulation environnementale à celle qui sera nommée au titre de la régulation de l'IA. Cette autorité aurait alors la possibilité de mener des investigations, auditer des systèmes, demander l'accès à des informations ou des données relatives à la consommation énergétique de systèmes d'IA considérés comme « à haut risque environnemental ». Ces enquêtes devraient, pour que les mesures de régulation soient effectives, pouvoir aboutir à des sanctions financières mais également réputationnelles, *via* par exemple l'affichage obligatoire de la sanction et de l'impact environnemental réel du système. Ces sanctions possibles doivent produire un effet dissuasif à l'égard des entreprises qui seraient tentées de dissimuler leur réel impact environnemental ou de pratiquer le *green-washing* en présentant une solution d'IA comme écologique alors que toutes les mesures de minimisations n'auraient pas été prises. À travers cette supervision et ces

[1388] MINISTÈRE DE LA TRANSITION ÉCOLOGIQUE, « Éviter, réduire et compenser les impacts sur l'environnement », *Site gouvernemental*, 14 juin 2021, disponible en ligne : <https://www.ecologie.gouv.fr/eviter-reduire-et-compenser-impacts-sur-lenvironnement>, consulté le 22 août 2021 ; A. Mechin, S. Pioch, « Séquence ERC : comment améliorer l'utilisation des méthodes de dimensionnement de la compensation écologique ? », *Revue électronique en sciences de l'environnement*, décembre 2014, vol. 14, n°3 ; V. aussi l'équivalent anglo-saxon, la méthode des 4 « R » C. Radclyffe, « The Four 'R's Of Sustainable Tech », *Forbes*, 17 août 2021, disponible en ligne : <https://www.forbes.com/sites/charlesradclyffe/2021/08/17/the-four-rs-of-sustainable-tech/?sh=6811c604243c>, consulté le 22 août 2021.

[1389] ISO, *Norme ISO 14044:2006 « Management environnemental — Analyse du cycle de vie — Principes et cadre »* ; et ISO, *Norme ISO ISO 14044:2006 « Management environnemental — Analyse du cycle de vie — Exigences et lignes directrices »*.

[1390] COMMISSION EUROPÉENNE, *Proposition de règlement du Parlement européen et du Conseil établissant des règles harmonisées concernant l'intelligence artificielle (législation sur l'intelligence artificielle) et modifiant certains actes législatifs de l'Union*, 21 avril 2021, 2021/0106 (COD), article 60.

pouvoirs de sanction, c'est à une réelle responsabilité environnementale des développeurs et opérateurs de systèmes d'IA que nous appelons.

§3 : La nécessité de garantir la responsabilité environnementale des acteurs de l'IA

661. **De la responsabilisation à la responsabilité des acteurs du développement de l'IA.** Les mesures de régulation visant à la prise en compte du risque environnemental dans la conception et l'opération des systèmes d'IA poursuivent un objectif de responsabilisation des acteurs de leur développement. Il semble pourtant nécessaire, qu'au-delà de leur responsabilisation, les acteurs qui développeraient des systèmes à fort impact environnemental au mépris de leurs obligations devraient être tenus responsables de leurs actes. Un premier élément de réponse est apporté par les pouvoirs de contrôle et de sanction qui pourraient être accordés à l'autorité indépendante nommée dans le cadre de la régulation de l'IA. Une deuxième réponse pourrait être le juge, grâce au contentieux climatique, ce qui requerrait de faciliter l'accès aux tribunaux à la société civile.

662. **La responsabilité environnementale des développeurs d'IA devant les tribunaux.** C'est ici le principe de « l'accès à la justice », fondamental en droit de l'environnement depuis la Convention d'Aarhus[1391], qui doit trouver à s'appliquer. En effet, plusieurs auteurs estiment que la responsabilité environnementale des entreprises passe par l'accès de la société civile aux tribunaux[1392]. Nombreux sont ceux qui considèrent que « *la capacité d'évaluation et d'audit des IA ne peut être réservée à un organe public, mais doit aussi provenir de la société civile* »[1393]. Le phénomène de militantisme numérique, le *Datactivism* aux États-Unis mais également en France[1394], témoigne de cette implication croissante d'associations dans le débat public autour de la régulation du numérique. Le rôle des pouvoirs publics est de les soutenir en anticipant « *la question des barrières financières d'accès aux moyens permettant aux organes de défense*

[1391] Convention d'Aarhus, *op. cit.*
[1392] L. Calandri, *op. cit.* ; C. Huglo, « Comment adapter le droit de l'environnement aux immenses défis lancés à l'humanité par le changement climatique ? », Bulletin du Droit de l'Environnement Industriel, 1 mars 2019, n°80.
[1393] C. Villani, *op. cit.*, p. 136.
[1394] I. Bruno, E. Didier, J. Prévieux, *Statactivisme : Comment lutter avec des nombres*, La Découverte, Paris, 2014.

d'intérêts civils et au journalisme de continuer à jouer efficacement leur rôle de vigie dans une époque numérisée [...] »[1395] et *« a minima de rendre les courroies de transmission plus fluides entre les autorités, la recherche, et la société civile »*[1396], notamment en facilitant l'accès aux données. Cette facilitation de l'accès à la justice pour des associations permettrait tenir les entreprises responsables pour les conséquences environnementales du développement de systèmes trop énergivores. Toutefois, cet accès aux tribunaux se confrontera aux mêmes limites que celles inhérentes à tout contentieux climatique.

663. **Les limites du contentieux climatique.** Si le principe du contentieux climatique est attractif, sa mise en œuvre n'en est pas moins difficile. Elle se heurte à de nombreuses barrières qui ont donné lieu à une littérature abondante[1397]. Etant un domaine de recherches en pleine ébullition, notre thèse se contentera d'identifier les quelques problématiques majeures qui se posent à l'égard de notre sujet, le numérique et l'IA, ainsi que les pistes d'évolution qui nous semblent prometteuses.

664. **La question de l'intérêt à agir.** Le premier obstacle à surmonter est la qualification de l'intérêt à agir. Elle est d'autant plus pertinente en ce qui concerne l'IA car généralement les systèmes touchent plusieurs milliers voire dizaines de milliers de personnes. De plus, les conséquences environnementales des émissions liées au système vont toucher la société au sens large, sans lien avec les utilisateurs de la solution. Définir la ou les personnes disposant d'un intérêt à agir contre un développeur d'IA « rouge »[1398] n'est donc pas chose aisée. Les pistes consistant à favoriser l'action d'associations[1399] ou la réunion de consommateurs dans des actions de groupe nous semblent être prometteuses à cet égard. Sur cette dernière piste, il est à noter que plusieurs lois vont dans le sens de l'accès aux procédures pour des groupes avec la création d'actions de groupe qu'il conviendra de bien articuler. En effet, faut-il en reconnaître de nouvelles formes alors qu'il existe déjà, en droit français, les actions de groupe sui

[1395] C. Villani, *op. cit.*, p. 129.
[1396] *Ibid.*
[1397] F.G. Trébulle, « Droit de l'environnement », *Recueil Dalloz*, 2010, p. 2468.
[1398] Pour reprendre l'expression « Red AI » (P. Dhar, *op. cit.*), entendre un système d'IA ayant des conséquences particulièrement néfastes pour l'environnement, que ce soit par sa consommation énergétique, les émissions de CO2 au cours de son cycle de vie, l'épuisement de ressources naturelles qu'il permet par son utilisation.
[1399] C. Huglo, *op. cit.*, p. 5.

generis[1400], l'action de groupe « données personnelles »[1401] et l'action de groupe environnementale[1402] ?

665. **La question du préjudice.** La deuxième limite du contentieux climatique appliqué à l'IA et celle de la qualification du préjudice. Sur ce point, la jurisprudence est en train d'évoluer pour adopter une vision plus large de la notion de préjudice[1403], ce qui pourrait, à terme, aboutir à la reconnaissance d'un « *préjudice numérique environnemental* »[1404].

666. **La question de la preuve.** Enfin, une dernière limite que nous pouvons identifier est celle de la preuve[1405]. Comment la société civile pourrait-elle apporter la preuve qu'une entreprise cause un dommage environnemental du fait des émissions liées aux systèmes qu'elle développe ? Comment accéder aux données nécessaires pour être en mesure d'apporter cette preuve ? C'est la raison pour laquelle les obligations environnementales et la supervision d'un tiers de confiance telles que proposées précédemment sont primordiales : elles fixent les règles pour les entreprises et donnent les armes juridiques à la société civile (grâce à la transparence) pour les faire respecter devant les tribunaux.

667. **Conclusion du Chapitre 2 relatif à la régulation environnementale du développement de l'IA.** Le niveau colossal des émissions de CO2 liées au cycle de vie d'un système d'IA, de l'extraction des matières premières à son utilisation finale, en passant par son apprentissage, et le silence du droit en la matière ont été présenté dans la première Partie de notre thèse. Le constat est unanime mais l'invisibilité du danger, l'incertitude pesant sur les conséquences concrètes de ces émissions et leur évolution dans le temps font que les pouvoirs publics ont préféré éluder le sujet des premières réflexions pour réguler l'IA. Pourtant, cette réaction est parfaitement contraire à l'esprit du principe de précaution, présent dans les traités internationaux, le traité fondateur de l'Union européenne et même dans le bloc de

[1400] Voir notamment l'article L. 423-1 du Code de la consommation.
[1401] Loi Informatique et Libertés, article 43 ter III, modifié par la loi n° 2018-493 du 20 juin 2018 relative à la protection des données personnelles, article 25.
[1402] Code de l'environnement, article L. 1142-3-1.
[1403] C. Huglo, *op. cit.*
[1404] L. Calandri, *op. cit.*
[1405] C. Huglo, *op. cit.*, p. 6.

constitutionnalité français. Ce principe impose de prendre des mesures visant à limiter le risque environnemental d'une activité dangereuse, même en l'absence de certitudes scientifiques sur la nature ou l'ampleur de ce risque.

La première Section du Chapitre 2 a démontré que la situation créée par le développement de l'IA, et particulièrement son empreinte environnementale « invisible », justifie l'application du principe de précaution. Quand bien même l'incertitude du risque n'était pas qualifiée, le principe de prévention prendrait le relais pour faire face au risque « avéré » que ferait peser l'IA sur l'environnement. Dès lors, des mesures proportionnées doivent être prises pour limiter, sans attendre de certitudes sur leur ampleur et leurs conséquences, les émissions de CO2 générés par le développement de l'IA. Ces mesures de précaution doivent en premier lieu consister à détourner l'IA de ses applications au service d'activité fortement émettrices de CO2. Elles doivent ensuite viser à réduire les émissions liées directement au système d'IA en lui-même, en créant un véritable cadre juridique pour une IA écologique.

La deuxième Section présente ce que pourrait être une cadre juridique en faveur d'une IA écologique. Ce dernier serait fondé sur quatre principes : précaution, quantification et transparence, minimisation et responsabilité. En premier lieu, le principe de précaution vise à l'adoption immédiate de mesures pour limiter le risque environnemental de l'IA, sans attendre de certitudes scientifiques sur son ampleur. Le principe de transparence impliquerait le partage (aux autorités de régulation et au public) de toute l'information pertinente relative aux conséquences environnementales liées au développement des systèmes d'IA par les entreprises. Le principe de minimisation imposerait aux parties prenantes de réduire autant que possible l'empreinte environnementale de l'IA, dans une logique de sobriété numérique. Enfin, la responsabilité des développeurs d'IA serait garantie par un contrôle effectif à la fois par une autorité indépendante et, le cas-échéant, par l'accès à la justice pour la société civile.

L'ensemble des propositions formulées dans ce Chapitre présentent plusieurs particularités. Contrairement à la majeure partie de la thèse, elles ne sont pas spécifiques au secteur de l'électricité. Elles peuvent donc inspirer une règlementation plus générale, transversale, et/ou être déclinée dans d'autres secteurs d'activité. Enfin, notre raisonnement peut être élargi à l'ensemble du secteur du numérique. En effet, les problèmes posés par l'IA ont de nombreux points communs avec d'autres technologies de rupture telles que la blockchain, l'internet des objets, ou l'informatique quantique. Si bien que la problématique environnementale se pose de la même manière pour toutes les technologies numériques. À ce titre, l'action des pouvoirs

publics en faveur d'une écologie numérique devient indispensable et, à de nombreux égards, urgente.

668. **Conclusion du Titre 2 relatif à la concrétisation de la régulation proportionnée de l'IA.** Au-delà des réflexions théoriques sur les meilleurs moyens de réguler la technologie, il est important de se confronter à la réalité. La régulation parfaite, par son contenu et par sa forme, ne peut exister car les processus de création de la norme impliquent de nombreux acteurs aux positions politiques divergentes. Pourtant, les défis posés par le développement de l'IA au regard des droits et libertés des individus sont réels et appellent une réponse. C'est pourquoi il est crucial que nous puissions concrétiser un cadre de régulation et le faire partager le plus largement possible. Cette concrétisation doit, d'une part, reposer sur les initiatives déjà avancées telles que le projet de règlement européen dédié à l'IA. Ce dernier semble néanmoins déséquilibré et créerait une forte contrainte juridique sur de nombreuses applications de l'IA dans le secteur de l'électricité, souvent sans réelle nécessité. Son amendement, avant son entrée en vigueur, est essentiel pour assurer une régulation proportionnée de l'IA préservant la capacité d'innovation des énergéticiens. La concrétisation de la régulation de l'IA doit, d'autre part, être suffisamment ambitieuse pour adresser des risques jusqu'alors occultés. C'est le cas, par exemple, pour la question de l'empreinte environnementale de l'IA. La réponse à y apporter est complexe tant l'IA peut à la fois présenter un danger au vu des émissions qu'elle génère et être un puissant levier pour la transition écologique, notamment dans le secteur de l'électricité. L'application de grands principes inspirés du droit de l'environnement tels que les principes de précaution, de prévention ou de l'information du public constitue une piste intéressante pour bâtir une ambitieuse régulation environnementale de l'IA.

669. **Conclusion de la Partie 2 relative à la nécessaire création d'un droit de l'IA.** L'adaptation du corpus juridique existant ne suffira pas à couvrir l'intégralité des risques générés par l'IA. En effet, il est essentiel d'harmoniser les bonnes pratiques en matière de conception et de maintenance des systèmes d'IA afin d'en garantir un fonctionnement sûr, non-discriminatoire et éco-responsable. À cet égard, le droit est un formidable outil pour parvenir à cet objectif en ce qu'il permet d'assujettir les comportements individuels à des normes contraignantes. La création d'un nouveau cadre juridique spécifique à l'IA apparaît nécessaire pour prévenir les risques spécifiques liés à l'utilisation de l'IA. Ce constat se vérifie dans le

secteur de l'électricité où de nombreuses applications peuvent être particulièrement dangereuses, notamment lorsqu'elles concernent la production d'électricité, le pilotage du réseau électrique ou l'analyse des données de consommation des individus. Outre la prévention des risques, la création d'un cadre de régulation pour l'IA ne devrait pas avoir pour effet d'entraver l'innovation de façon démesurée. En effet, de nombreuses applications de l'IA sont souhaitables telles que les systèmes de prévision de la production des sources de production d'énergies renouvelables, les outils intelligents de performance énergétique, et bien d'autres.

Le premier Titre a permis de présenter à la fois les motifs justifiant la création d'un nouveau cadre juridique et les moyens à mettre en œuvre pour parvenir à une régulation proportionnée. Pour identifier ces moyens, il est utile de s'intéresser à l'expérience de la régulation dans d'autres domaines présentant des similarités avec notre sujet : la protection des données ou les régulations financière et environnementale. Ces expériences nous ont enseigné que l'équilibre peut, d'abord, être atteint en adoptant une approche de co-régulation, laissant aux acteurs régulés la liberté de choisir les moyens mis en œuvre pour se conformer aux exigences fixées par les textes. Ensuite, il convient d'adopter une approche différenciée suivant les risques. En effet, le niveau de risque diffère grandement suivant le cas d'usage concerné. L'évaluation des risques doit faire l'objet d'un consensus entre les pouvoirs publics et les acteurs régulés, qui sont les mieux placés pour réaliser cette tâche. Le contenu des exigences de conformité doit également être réfléchi pour correspondre à la réalité technique et être suffisamment souple pour s'adapter aux évolutions futures de la technologie. Enfin, une régulation proportionnée et acceptable de l'IA passe par l'adoption d'un mode de gouvernance adapté. Au vu de la nature sociétale des risques portés par l'IA, il est primordial que sa régulation se fasse dans un cadre démocratique. L'implication des citoyens en leur reconnaissant un droit d'action ou un droit à l'information sur le modèle du droit de l'environnement s'avère alors incontournable. Sur un sujet aussi complexe que l'IA, la mise en conformité sera nécessairement difficile pour les entreprises. C'est pourquoi la mise en œuvre de la régulation par une autorité indépendante, chargée de contrôler mais aussi d'accompagner les acteurs régulés dans leur mise en conformité comme c'est le cas dans la législation en matière de données personnelles, semble plus que pertinente.

Les réflexions théoriques développées dans le premier Titre ont ensuite été appliquées concrètement dans le second. La proposition de règlement européen sur l'IA est bienvenue mais peine à convaincre de sa capacité à concilier prévention des risques et promotion de l'innovation. C'est la raison pour laquelle la présente thèse contient une dizaine de propositions

d'amendement visant à équilibrer le texte. Pour terminer, nos recherches se sont concentrées sur *« l'angle mort de la doctrine environnementaliste »* pour reprendre l'expression du Professeur Gilles Martin[1406] : l'empreinte écologique de l'IA. La concrétisation de la régulation passe également par une réponse ambitieuse à cette problématique complexe et paradoxale, tant l'IA est à la fois source de risques et de potentiels bénéfices pour l'environnement. Les principes fondateurs du droit de l'environnement offrent à cet égard une base de réflexion très riche dont l'application à l'IA est conditionnée à un prérequis : l'harmonisation des moyens de quantification de l'empreinte environnementale des systèmes d'IA. À cet effet, nos développements ont mis en évidence la pertinence des méthodologies de quantification des émissions de gaz à effet de serre (émissions directes et indirectes), y compris appliquées aux systèmes d'IA. Le chemin vers la création d'un cadre juridique couvrant tous les risques reste encore long et nécessitera de réunir un grand nombre d'acteurs pour parvenir à une régulation proportionnée. À cette fin, la pluridisciplinarité est essentielle. Il est du devoir du juriste de monter en compétence sur ces sujets techniques afin, d'une part, de contribuer à la création d'une norme adaptée en influant sur le travail législatif et, d'autre part, d'accompagner au mieux les développeurs d'IA dans leur mise en conformité. Finalement, le succès de la régulation de l'IA ne résidera qu'en la capacité du juriste et du législateur à comprendre les problématiques techniques et les conséquences que leurs choix peuvent avoir dans la pratique, notamment au regard de la capacité d'innovation des entreprises.

[1406] G. Martin, « Les angles morts de la doctrine juridique environnementaliste », *Revue juridique de l'environnement*, Lavoisier, 2020, 45, 67-80.

CONCLUSION GÉNÉRALE

671. Dans son ouvrage « Comment il faut faire sa thèse de doctorat en droit », le Professeur Henri Capitant estimait qu'à défaut de conclusion l'ouvrage que constitue la thèse était inachevé[1407]. Selon ses mots, *« il est indispensable pour la clarté et l'utilité de la thèse de ramasser en quelques pages les idées essentielles qui se dégagent du travail accompli, et de formuler les conclusions en termes nets et précis »*[1408]. En termes de clarté, un exemple peut être trouvé en la thèse du Professeur Emmanuel Gaillard où l'auteur présente sa conclusion sous la forme d'une liste de positions de thèses[1409]. L'ingéniosité et l'utilité du procédé sont soulignées par le non moins illustre Professeur Gérard Cornu dans sa préface : *« quant à chercher une introduction à l'ouvrage, sautez d'emblée à la conclusion de l'auteur. Les dix-huit positions de thèse émaillent le chemin. C'est par là qu'il faut y entrer, pour s'y accorder, thèse lue »*[1410]. La lecture desdites positions, présentées sous forme de liste, achève de convaincre que ce format correspond parfaitement à la dimension appliquée du travail de recherche effectué dans le cadre de notre étude. En effet, ses conclusions doivent être exprimées suffisamment clairement pour pouvoir être portées à la connaissance des acteurs du secteur de l'électricité et pour que ces derniers puissent s'en saisir. C'est la raison pour laquelle nous proposons de rappeler, sous une forme résumée, les dix-sept positions de thèse défendues dans le manuscrit.

[1407] H. Capitant, *Comment il faut faire sa thèse de doctorat en droit*, Librairie Dalloz, 1926, p. 46.
[1408] *Ibid.*
[1409] E. Gaillard, *Le pouvoir en droit privé*, thèse pour le doctorat en droit privé, Economica, coll. Droit civil, 1985, pref. G. Cornu, pp. 232-235.
[1410] *Ibid.*, p. 4.

Sur la définition de l'IA :

1. En l'absence de consensus scientifique sur la définition technique de l'IA, les systèmes d'IA devraient être qualifiés à partir de leurs spécificités par rapport aux systèmes logiciels classiques. Ces spécificités sont au nombre de quatre : la complexité structurelle, l'opacité du fonctionnement, la potentielle autonomie et la dépendance à la donnée. Elles peuvent gêner l'application des règles de droit existantes, générant une insécurité juridique et donc un frein à l'innovation[1411].

Sur l'application des régimes de responsabilité aux dommages causés par des systèmes d'IA :

2. En droit français, les régimes de responsabilité et ses principes sous-jacents sont en grande partie adaptables aux dommages causés par des systèmes d'IA, en l'état actuel de la technique. Des incertitudes résiduelles pourraient naître en cas de développement de systèmes complètement autonomes et auto-apprenants[1412].
3. L'harmonisation des régimes de responsabilité civile au niveau européen est souhaitable pour pallier les lacunes des régimes nationaux mais ne devrait pas venir troubler les régimes dont les principes sont déjà suffisamment adaptés[1413].

Sur l'application du régime de protection des données à caractère personnel aux traitements réalisés par des systèmes d'IA :

4. Le régime européen de la protection des données à caractère personnel est pleinement applicable aux systèmes d'IA[1414].

[1411] Sur les caractéristiques spécifiques des systèmes d'IA, voir Supra, 11-16.
[1412] Sur l'application des régimes de responsabilité aux dommages causés par des systèmes d'IA, voir Supra, 66-93.
[1413] Sur le projet d'harmonisation européenne des règles de responsabilité applicables aux systèmes d'IA, voir Supra, 97-99.
[1414] Sur l'effectivité des régimes de protection des données appliqués aux traitements réalisés au moyen de systèmes d'IA, voir Supra, 105-117.

5. L'application du RGPD peut néanmoins créer des freins disproportionnés au développement de l'IA en raison des contradictions entre ses principes et les concepts sous-jacents au fonctionnement des systèmes d'IA, du défaut d'explicabilité du fonctionnement de certains systèmes et des contraintes limitant le recours à des solutions d'IA non-européennes[1415].
6. Ces contraintes peuvent être levées par la mise en œuvre de nouvelles méthodes de protection des données, notamment de chiffrement, par les responsables de traitement[1416].

Sur l'application des règlementations du secteur de l'électricité aux systèmes d'IA :

7. Dans le secteur de l'électricité, certaines règlementations sectorielles génèrent des freins à l'utilisation de l'IA.
8. L'intégration de l'IA dans le système d'information des centrales nucléaires de production d'électricité n'est pas compatible avec les régimes de sûreté et de sécurité nucléaire, en particulier au regard des processus de certification[1417].
9. Le processus de certification des systèmes logiciels utilisé dans le cadre de la démonstration de sûreté nucléaire auprès de l'ASN repose sur des principes pertinents mais son application en pratique devra être adaptée aux spécificités des systèmes d'IA et à la façon dont ils sont conçus[1418].
10. Une discussion collégiale réunissant les autorités de régulation, les opérateurs d'infrastructures critiques, des experts indépendants ainsi que des représentants du

[1415] Sur la disproportion de la contrainte résultant de l'application des régimes de protection des données appliqués aux traitements réalisés au moyen de systèmes d'IA, voir Supra, 120-143.
[1416] Sur la nécessaire standardisation des mesures de protection des données applicables aux systèmes d'IA, voir Supra, 291.
[1417] Sur l'incompatibilité des caractéristiques des systèmes d'IA avec la règlementation applicable aux INB, voir Supra, 157-200.
[1418] Sur l'incompatibilité des caractéristiques des systèmes d'IA avec le processus de qualification des logiciels utilisés dans la démonstration de sûreté nucléaire, voir Supra, 190-200.

grand public est nécessaire pour étudier l'opportunité d'assouplir certaines contraintes de sécurité empêchant le recours à l'IA dans le contexte des systèmes critiques[1419].

11. Le régime juridique encadrant la collecte et le traitement de données de consommation électrique apparaît également disproportionné lorsque le traitement par l'IA poursuit une finalité écologique d'intérêt général (réalisation d'économies d'énergie, équilibrage du réseau d'électricité…) et que des mesures techniques de protection de la vie privée peuvent être mises en œuvre[1420].
12. La dynamique d'*open data* est bien établie dans le secteur de l'électricité mais sa portée est encore limitée dans la pratique. Les législations européennes sur les données en cours d'adoption doivent être soutenues en ce qu'elles permettront de promouvoir encore l'ouverture des données[1421].

Sur la nécessité d'une régulation *sui generis* de l'IA :

13. Le déploiement de systèmes d'IA dans le secteur de l'électricité génère des risques qui ne sont pas couverts par le corpus législatif existant (risques éthiques, de cybersécurité et environnementaux)[1422].
14. Ces risques justifient la création d'une régulation *sui generis* pour les prévenir[1423].
15. La régulation de l'IA devrait adopter un modèle de co-régulation, fondé sur une approche par les risques et des mécanismes de conformité, afin de ne pas créer un frein disproportionné à l'innovation[1424].
16. La proposition européenne d'un cadre règlementaire sur l'IA est pertinente en ses principes, mais comporte de nombreux écueils qu'il convient de corriger afin de ne pas

[1419] Sur la nécessité d'une réflexion collégiale questionnant la proportionnalité des règles de sécurité applicables aux systèmes critiques dans le secteur de l'électricité, voir Supra, 205-209.
[1420] Sur les solutions à la disproportion des contraintes au traitement des données de consommation énergétique, voir Supra, 289-292.
[1421] Sur les apports des nouvelles règlementations européennes du numérique pour libérer les données dans le secteur de l'électricité, voir Supra, 534-535.
[1422] Sur l'existence de risques résiduels, non couverts par le corpus juridique existant, voir Supra, 297 ; 310-313.
[1423] Sur les justifications de la création d'un cadre juridique applicable aux systèmes d'IA, voir Supra, 308-313.
[1424] Sur les moyens juridiques à mettre en œuvre pour créer un cadre de régulation proportionné, voir Supra, 342-469.

créer un fardeau administratif disproportionné pour les fournisseurs d'IA et ainsi générer un effet contreproductif sur l'innovation[1425].

17. L'empreinte environnementale du développement de l'IA doit faire l'objet d'un cadre juridique adapté qui peut être fondé sur quatre principes issus des principes directeurs du Droit de l'environnement : précaution et prévention ; quantification de l'impact environnemental et transparence ; minimisation ; responsabilité des acteurs[1426].

[1425] Sur les nécessaires amendements à apporter au projet de règlement européen sur l'IA, voir Supra, 491-574.
[1426] Sur la construction d'un cadre juridique visant à réguler l'empreinte environnementale du développement des systèmes d'IA, voir Supra, 577-667.

BIBLIOGRAPHIE

Plan de la bibliographie :

§I : Ouvrages, traités, manuels et cours

A/ Sources juridiques

Atias C., *Philosophie du droit*, PUF, 4ème éd., 2016, p. 331.

Bensamoun A., Bertrand B., *Le règlement général sur la protection des données : Aspects institutionnels et matériels*, Mare & martin, 2020.

Bensamoun A., Loiseau G., *Droit de l'intelligence artificielle*, LGDJ, coll. Les intégrales, 2019, 444p.

Bensoussan A., Bensoussan J., *Droit des robots*, Larcier, 2015.

Bétaille J., *Le droit d'accès à la justice en matière d'environnement*, Presses de l'Université Toulouse 1 Capitole, LGDJ – Lextenso Editions, 2016.

Boddington P., *Towards a Code of Ethics for Artificial Intelligence*, Springer, 2017, 124 p.

Boutonnet M., *Le principe de précaution en droit de la responsabilité*, LGDJ, 2005.

Bourcier D., *La décision artificielle*, PUF, coll. Les voies du droit, Paris, 1995, 240 p.

Boursier D., Asset P., Roquilly C., *Le droit et l'intelligence artificielle : une Révolution de la connaissance juridique*, Romillat, Paris, 1994, 304 p.

Burg M., *Droit des risques technologiques*, Legitech, 2020, 1ère ed., 216 p.

Capitant H., *Comment il faut faire sa thèse de doctorat en droit*, Librairie Dalloz, 1926.

Castets-Renard C., Ndior V., Rass-Masson L., *Enjeux internationaux des activités numériques : entre logique territoriale des États et puissance des acteurs privés*, Larcier, septembre 2020, 202 p.

Champeil-Desplats V., *Méthodologies du droit et des sciences du droit*, Dalloz, 2016, coll. Méthodes du droit, 2ème ed.

Colson J.P., Idoux P., *Droit public économique*, LGDJ, 4e éd., 2008.

Cornu G., *Vocabulaire juridique*, PUF, 2014, 10ème ed.

Corrales M., Fenwick M., Forgó N., *Robotics, AI and the Future of Law*, Springer, 2018.

De Terwangne C., Rosier K., *Le Règlement général sur la protection des données (RGPD / GDPR), Analyse approfondie*, Coll. du CRIDS, Larcier, 1re édition, 2018.

Debet A., Massot J., Metallinos N., *La protection des données à caratère personnel en droit français et européen*, Lextenso Editions, 2015, n° 855.

Dignum V., *Responsible Artificial Intelligence : how to develop and use AI in a responsible way*, Springer, coll. Artificial Intelligence, 2019, 1st ed., 127p.

Ebers M., Navas Navarro S., *Algorithms and Law*, Cambridge University Press, 2019.

Féral-Schuhl C., *Cyberdroit*, Dalloz, coll. Praxis, 8ème éd., 2020, 1852 p.

Guastini R., *Teoria del diritto. Approccio metodologico*, Mucchi, 2013, coll. Piccole conferenze.

Kelsen H., *General Theory of the Law and the State*, Harvard University Press, 1949, 544 p.

Kershaw D., *Principles of takeover regulation*, Oxford University Press, 2016, 418 p.

Kutz C., *Complicity: Ethics and Law for a Collective Age*, Cambridge University Press, 2000, 1st edition, 344p.

Le Tourneau P., *Droit de la responsabilité et des contrats*, Dalloz action, 2021.

Linant de Bellefonds X., *L'informatique et le Droit*, PUF, coll. Que sais-je ?, Paris, 1981, 127 p.

Loiseau G., *Le droit des personnes*, Ellipses, 2016, p. 74.

Mathieu M.L., *Logique et raisonnement juridique*, PUF, Paris, 2015, coll. Themis 2ème ed., 446 p.

Mayaud Y., *Droit pénal général*, PUF, 2018, 6e éd., n°232.

Meneceur Y., *L'intelligence artificielle en procès : plaidoyer pour une règlementation internationale et européenne*, Bruylant, coll. Macro Droit – Micro Droit, 2020, 209 p.

Perelman C., Obrechts-Tyteca L., *Traité de l'argumentation*, Editions de l'université de Bruxelles, 2008, 6ème ed., 740 p.

Portney P.R., Stavins R.N., *Public Policies for Environmental Protection*, Resources for the Future, 2000, 2ème ed.

Quéméner M., *Le droit face à la disruption numérique*, Gualino, Lextenson, 2018.

Salzman J., Thompson B.H., *Environmental law policy*, Foundation Press, 3rd ed., 2010, 397p.

Savatier S., *Les métamorphoses économiques et sociales du droit civil d'aujourd'hui*, Panorama des mutations, 3e éd., 1964.

B/ Sources extra-juridiques

Bruno I., Didier E., Prévieux J., *Statactivisme : Comment lutter avec des nombres*, La Découverte, Paris, 2014.

Bush S.F., *Smart Grid: Communication-Enabled Intelligence for the Electric Power Grid*, John Wiley & Sons Ltd, West Sussex, UK, 2014.

Copeland J., *The Essential Turing: Seminal Writings in Computing, Logic, Philosophy, Artificial Intelligence, and And Artificial Life plus The Secret of Enigma*, Oxford University Press, New York, 2004.

D. Gronier, *L'intelligence artificielle en action : Santé, environnement, énergie... ce que l'IA change concrètement*, Eyrolles, 2020, pref. C. Villani, 220 p.

Doern G.B., *Canadian Energy Policy and the Struggle for Sustainable Development*, University of Toronto Press, 2005.

Doern G.B., *Innovation, Science, Environment: Canadian Policies and Performance*, McGill-Queens University Press, 2006.

Dubber M., Pasquale F., *The Oxford Handbook of AI Ethics*, Oxford University Press, 2019, 896 p.

Durkheim E., *De la division du travail social*, PUF, Paris, 1962.

Ganascia J. G., *Le mythe de la singularité. Faut-il craindre l'intelligence artificielle ?*, Seuil, 2017.

Goertzel B., Pennachin C., *Artificial General Intelligence*, Springer, Berlin, 2007.

Grau Ruiz M.A., *Interactive Robotics: Legal, Ethical, Social and Economic Aspects: Selected contributions to the INBOTS conference 2021*, Springer, coll. Biosystems & Biorobotics, 2022, vol 30.

Gronier D., *L'intelligence artificielle en action : Santé, environnement, énergie... ce que l'IA change concrètement*, Eyrolles, 2020, pref. C. Villani, 220 p.

Haëntjens J., *Smart city, ville intelligente : quels modèles pour demain ?*, La Documentation française, 2021, 180 p.

Hood C., Rothstein H., Baldwin R., *The government of risk : understanding risk regulation regimes*, Oxford University Press, 2001, 232 p.

Koenig G., *La fin de l'individu : voyage d'un philosophe au pays de l'intelligence artificielle*, L'Observatoire, coll. De Facto, 189 p.

Kurzweil R., *The Singularity is Near: When Humans Transcend Biology*, Viking Penguin, New York, 2005, 602 p.

Lecourt D., *Technophobie*, Cités, 2000.

Leveson N., *Engineering a Safer World: Systems Thinking Applied to Safety*, MIT Press, 2012, Cambridge, MA, USA.

Leveson N., *Safeware : System Safety and Computers*, Addison-Wesley Professional, 1995, 704p.

Lewandowski C., *Le nucléaire*, Presses Universitaires de France, coll. Que sais-je ?, 2021, 128 p.

Marciano A., Tourrès B., *Regards critiques sur le principe de précaution : le cas des OGM*, Vrin, Paris, 2011.

Martin K., *Ethics of Data and Analytics*, Auerbach Publications, New York, 1ère ed., 2022.

Merad M., Dechy N., Dehouck L., *et al.*, *Risques majeurs, incertitudes et décisions - Approche pluridisciplinaire et multisectorielle*, MA Edition, ESKA, Paris, 2016, 315 p.

Mill J.S., *De la liberté*, 1859, disponible en ligne : <www.utilitarianism.com/ol/five.html>, consulté le 21 janvier 2022.

Oberkampf W., Roy C., *Verification and validation in scientific computing*, Cambridge University Press, 2010, Cambridge, UK.

O'Neil C., *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*, Crown Random House, 2016, 272 p.

Pasquale F., *The Black Box Society: The Secret Algorithms That Control Money and Information*, Harvard University Press, Cambridge, 2015, 260 p.

Roache P. J., *Verification and validation in Computational Science and Engineering*, Hermosa publishers, 1998, 446 p.

Russel S. J., Novig P., *Artificial Intelligence : a modern approach*, Pearson, 2020, 4th ed., 1152 p.

Triclot M., *Le moment cybernétique. La constitution de la notion d'information*, Champ Vallon, Paris, 2008, 384 p.

§II : Thèses et mémoires

Buijze A.W., *The principle of transparency in EU law*, thèse pour le doctorat en droit, Utrecht University, 338 p.

Dupichot P., *La garde de la structure et la garde du comportement dans la responsabilité civile*, thèse pour le doctorat en droit privé, Université Paris XII, 1984.

Gaillard E., *Le pouvoir en droit privé*, thèse pour le doctorat en droit privé, Economica, coll. Droit civil, 1985, pref. G. Cornu.

Guebels G., *Le phénomène des bulles de filtres sur Internet : Le moteur de recherche Google nous oriente-t-il à notre insu à cause de son algorithme de personnalisation ?*, Mémoire de recherche, dir. L. Groetaers, Faculté des sciences économiques, sociales, politiques et de communication, Université catholique de Louvain, 2018, disponible en ligne : <https://dial.uclouvain.be/memoire/ucl/object/thesis:16354>, consulté le 2 août 2021.

Guegan G., *L'élévation des robots à la vie juridique*, Thèse pour le doctorat en droit privé, Université Toulouse 1 Capitole, 2016, 368 p.

Henin C., *Expliquer et justifier les systèmes de décisions algorithmiques*, thèse pour le doctorat en informatique, Université de Lyon, 2021, 176 p.

Leukam Lako F., *Protection des données à caractère personnel pour les services énergétiques*, thèse pour le doctorat en informatique, réseaux et télécommunications, Institut polytechnique de Paris, 2021.

Maxwell W., *A Method to Assess Regulatory Measures Designed to Limit Access to Harmful Content on the Internet*, Thèse pour le doctorat en sciences économiques, Telecom ParisTech, 2016.

Merabet S., *Vers un droit de l'intelligence artificielle*, Thèse de doctorat en droit privé, Dalloz, coll. Nouvelles bibliothèque de thèses, 2020, vol. 197, 1ère ed., 592 p.

Moreau V., *Méthodologie de représentation des impacts environnementaux locaux et planétaires, directs et indirects - Application aux technologies de l'information*, thèse pour le doctorat en sciences et génie de l'environnement, Ecole Nationale Supérieure des Mines de Saint-Etienne, 2012, 361 p.

Paquier Y., *Le principe de transparence des traitements algorithmiques : de l'étude juridique d'un enjeu démocratique*, Thèse pour le doctorat en droit public, 10 novembre 2021, Université de Caen Normandie.

Pouget J., *La réparation du dommage impliquant une intelligence artificielle*, Thèse pour le doctorat en droit privé, Université d'Aix Marseille, 2019, 410 p.

Vallier R., *La conformité environnementale, une politique juridique au service de la performance globale de l'entreprise*, thèse pour le doctorat en Droit, Université Côte d'Azur, 357 p.

§III : Articles, études et conférences

A/ Sources juridiques

Abbott K., Snidal D., « Hard and Soft Law in International Governance », *International Organization*, 2000, vol. 54, n°3, 421-456.

Ackerman B.A., Stewart R.B., « Reforming Environmental Law », *Stanford Law Review, 1985*, vol. 37, n°1333.

Andrieu S., Gourdou J., « Le bac à sable règlementaire dans le secteur de l'énergie », *Energie – Environnement – Infrastructures*, LexisNexis, novembre 2020, n°11, étude 16.

Ashley L., « Why regulators must adapt to fintech », *International Financial Law Review*, 2015, vol. 34, n°6.

Babor D.D., « Environmental Rights in Ontario: Are Participatory Mechanisms Working », *Colorado Journal of International Environmental Law and Policy*, 1999, vol. 10, n°1998, p. 121-135.

Baralle P.J., « Maîtrise de l'urbanisation autour des installations dangereuses », *JurisClasseur Environnement et Développement durable*, LexisNexis, 15 janvier 2009, Fasc. 4035.

Barbin E., « Le contrôle juridictionnel de l'outil numérique d'aide à la décision administrative », *RFDA*, 2021, p. 491.

Barocas S., Hood S., Ziewitz M., « Governing Algorithms: A Provocation Piece », *Proceedings of the « Governing Algorithms » conference*, 16 mai 2013, New York University.

Barocas S., Selbst A., « Big Data's Disparate Impact », *California Law Review*, 2016, vol. 104, 671.

Barret O., Brun P., « Vente : effets – Garanties contre les vices cachés », *Répertoire de droit civil*, 2020, pp. 655-658.

Baumevieille M., « L'industrie du médicament – Autorisation de mise sur le marché – Évolution du cadre juridique – Principes généraux », *Litec Droit pharmaceutique*, version mise à jour le 13 juillet 2015, Fasc. 33.

Beatrix O., « Open data et secteur de l'énergie : le début de l'histoire », *RFDA*, 2018, p. 49.

Beaudouin V., Bloch I., Bounie D., *et al.*, « Flexible and Context-Specific AI Explainability: A Multidisciplinary Approach », *SSRN Electronic Journal*, 2020, disponible en ligne : < https://www.ssrn.com/abstract=3559477>, consulté le 29 avril 2020.

Beaudouin V., Bloch I., Bounie D., *et al.*, « Identifying the Right Level of Explanation in a Given Situation », *Hal archives ouvertes*, 2020, hal-02507316, disponible en ligne : <https://hal.telecom-paristech.fr/hal-02507316>, consulté le 20 avril 2020.

Ben Youssef W.A., « Les cyber risques : nature, étendue et moyens de couverture », *Droit et patrimoine*, 1er janvier 2020, n°298.

Benabou V.-L., « Un droit vivant. Manifeste pour des juristes incarnés et sensibles à l'heure de l'intelligence artificielle », *Mélanges Michel Vivant*, 2020, pp. 715-730.

Bensamoun A., « Des robots et du droit... », *Dalloz IP/IT*, 2016, p. 281.

Bensamoun A., « Artificial Intelligence Act : l'Union européenne invente la pyramide des risques de l'intelligence artificielle », *Le Club des juristes (blog)*, 21 mai 2021, disponible en ligne : <https://blog.leclubdesjuristes.com/artificial-intelligence-act-lunion-europeenne-invente-la-pyramide-des-risques-de-lintelligence-artificielle/>, consulté le 28 février 2022.

Bensamoun A., Groffe J., « Création numérique », *Répertoire de droit civil*, Dalloz, 2013, pp. 34–45.

Bensamoun A., Loiseau G., « L'intégration de l'intelligence artificielle dans l'ordre juridique en droit commun : questions de temps », *Dalloz IP/IT*, 2017, p. 239.

Bensamoun A., Loiseau G., « L'intelligence artificielle : faut-il légiférer », *Recueil Dalloz*, 2017, p. 581.

Bensoussan A., « Droit des robots : science-fiction ou anticipation ? », *Recueil Dalloz*, 2015, p. 1640.

Bergé J.S., Le Métayer D., « Phénomène de masse et droit des données », *Commerce Communication Electronique*, décembre 2018, Etudes, n°20.

Bernatchez S., « La certification en tant que droit de la gouvernance », *Ethique publique*, 2019, vol. 21, n°1.

Bernelin M., « Intelligence artificielle : une proposition de directive sur la responsabilité civile extracontractuelle », *Dalloz actualité*, 22 novembre 2022.

Bertrand A., « L'intelligence artificielle, la robotique, les systèmes experts et le droit », *Expertises*, 1987, n°96, pp. 219-224.

Bertrand B., « Chronique Droit européen du numérique : La nouvelle approche de la cybersécurité européenne », *RTD Eur.*, 2021, p. 155.

Bertrand B., « Chronique Droit européen du numérique - La politique européenne du numérique : une vision politique européenne ? », *RTD Eur.*, 2022, p. 449.

Bertrand B., « Chronique de Droit européen du numérique : La souveraineté technologique européenne », *RTD eur.*, 2021, p. 139.

Bertrand B., « Chronique Droit européen du numérique : Perfectibilité de la protection des données personnelles », *RTD Eur.*, 2021, p. 143.

Bertrand B., « The Twin Digital and Green Transition », *RTD eur.*, 2022, p. 619.

Binet J. R., « Personnalité juridique des robots : une voie à ne pas suivre », *Revue Droit de la famille*, LexisNexis, 2017, n°6.

Blache D., « Règles prudentielles européennes applicables aux établissements de crédit, entreprises d'investissement, établissements de paiement et établissements de monnaie électronique », *JurisClasseur Banque - Crédit - Bourse*, Fasc. 110 Droit bancaire et financier européen.

Bobbio N., « Metodo », in *Contributi ad un dizionario giuridico*, Giapichelli editore, Turino, 1994.

Boiteau C., « L'entreprise régulée », *RFDA*, 2018, p. 469.

Borghetti J.S., « Civil Liability for Artificiel Intelligence : What Should Its Basis Be ? », *Revue des juristes de Sciences Po*, 2019, n°17, pp. 76-84.

Borghetti J.S., « L'accident généré par l'intelligence artificielle autonome », *Actes du colloque du Master 2 Droit privé général et du laboratoire de droit civil*, JCPG, 2017, n° spécial.

Boudinar-Zabaleta A., « Algorithmes et lignes directrices. Réflexions sur la codification automatisée des motifs des décisions administratives », *Droit Administratif*, avril 2019, n° 4, étude 7.

Boujeka A., « Vers un modèle de régulation des marchés financiers dans l'Union européenne », *Recueil Dalloz*, 2012, p. 1355.

Bourcier D., « De l'intelligence artificielle à la personne virtuelle : émergence d'une entité juridique ? », *Droit et société*, 2001, vol. 49, n°3, pp. 847-871.

Bourgeois M., Bounedjoum A., « Les apports de la loi pour une République numérique en matière d'accès et de réutilisation d'informations publiques », *La Semaine Juridique Administrations et Collectivités territoriales*, 5 décembre 2016, n° 48, p. 2307.

Boy L., « Réflexions sur le 'droit de la régulation' », *Recueil Dalloz*, 2001, p. 3031.

Brauneis R., Goodman E.P., « Algorithmic Transparency for the Smart City », *Yale Journal of Law & Technology*, vol. 20, n°103, 2018.

Buffelan J. P., « Le droit, l'informatique et la mathématique », *Journal de la société statistique de Paris*, 1974, tome 115, pp. 301-316.

Burk D.L., « Legal and Technical Standards in Digital Rights Management Technology », *Fordham Law Review*, 2005, vol. 74, 537.

Calandri L., « Pollution numérique et intelligence artificielle : variations autour du rapport de Monsieur le Député C. Villani, "Donner un sens à l'intelligence artificielle" », *Energie - Environnement – Infrastructures*, novembre 2018, n°11, 15.

Cameron J., Abouchar J., « The Precautionary Principle: A Fundamental Principle of Law and Policy for the Protection of the Global Environment », *Boston College International and Comparative Law Review*, 1991, vol. 14, 1–27.

Camy J., « Loi sur le devoir de vigilance et loi Sapin II : quelles obligations des entreprises ? », *La Semaine Juridique Entreprise et Affaires*, 18 mars 2021, n° 11, p. 1135.

Cans C., « Grande et petite histoire des principes généraux du droit de l'environnement dans la loi du 2 février 1995 », *Revue juridique de l'environnement*, 1995, p. 193.

Castets-Renard C., « Comment construire une intelligence artificielle responsable et inclusive ? », *Recueil Dalloz*, 2020, p. 225.

Castets-Renard C., « Le Livre blanc de la Commission européenne sur l'intelligence artificielle : vers la confiance ? », *Recueil Dalloz*, 2020, n°15, p. 837.

Castets-Renard C., « L'intelligence artificielle, les droits fondamentaux et la protection des données personnelles dans l'Union européenne et les États-Unis », *Revue de Droit International d'Assas*, 2019, 2, pp. 158-174.

Chander A., « How Law Made Silicon Valley », *Emory Law Journal*, 2014, vol. 63, 639.

Chander A., « The Racist Algorithm? », *Michigan Law Review*, 2017, 115.

Chevallier J., « Vers un droit post-moderne ? Les transformations de la régulation juridique », *Revue du droit public et de la science politique*, Librairie Générale de Droit et de Jurisprudence, 1998, pp. 659–714.

Choux E., Viano O., « Livre blanc du MEDEF : 40 propositions pour moderniser et simplifier le droit de l'environnement dans l'intérêt partagé de la protection de l'environnement et de la compétitivité des entreprises », *Bulletin du Droit de l'Environnement Industriel*, 1er juin 2017, n°69.

Citron D.K., « Technological Due Process », *Wahsington University Law Review*, 2008, vol. 85, issue 6, 1249.

Cluzel-Métayer L., « La loi pour une République numérique : l'écosystème de la donnée saisi par le droit », *AJDA*, 2017, p. 340.

Cluzel-Métayer L., « Les limites de l'open data », *AJDA*, 2016, p. 102.

Coglianese C., « The Limits of Performance-Based Regulation », *University of Michigan Journal of Law Reform*, 2017, vol. 50, n°3, pp. 525-564.

Cohen J.E., « What privacy is for », *Harvard Law Review*, 2013, 126, n°7, 1904–1933.

Contreras J.L., « Technical standards and "ex ante" disclosure: results and analysis of an empirical study », *Jurimetrics*, 2013, vol. 53, n°2, pp. 163-211.

Copain-Héritier C., « Le cadre européen de la protection des données entre forces et faiblesses intrinsèques », *Revue de l'Union européenne*, 2021, p. 163.

Coulon C., « Du robot en droit de la responsabilité civile : à propos des dommages causés par les choses intelligentes », *RCA*, 2016, Étude 6.

Courtois G., « Robots intelligents et responsabilité : quels régimes, quelles perspectives ? », *Dalloz IP/IT*, 2016, p. 287.

Courtois G., Mariez J. S., « Intelligence artificielle : quels objets de droit pour quel encadrement contractuel ? », *RLDA*, septembre 2019, n°151, p. 22, 6781.

Crichton C., « Intelligence artificielle : avis du CEPD sur le Livre blanc de la Commission », *Dalloz Actualité*, 17 juillet 2020, disponible en ligne : <https://www.dalloz-actualite.fr/flash/intelligence-artificielle-avis-du-cepd-sur-livre-blanc-de-commission>, consulté le 11 juillet 2021.

Crichton C., « Projet de règlement sur l'IA (I) : des concepts larges retenus par la Commission », *Dalloz Actualité*, 3 mai 2021.

Crichton C., « Publication par la Commission de son Livre blanc sur l'intelligence artificielle », *Dalloz Actualité*, 28 février 2020, disponible en ligne : <https://www.dalloz-actualite.fr/flash/publication-par-commission-de-son-livre-blanc-sur-l-intelligence-artificielle>, consulté le 11 juillet 2021

Crichton C., « Artificial Intelligence Act : avis conjoint des CEPD », *Dalloz Actualité*, 2 juillet 2021.

Crichton C., « Projet de règlement sur l'IA (II) : une approche fondée sur les risques », *Dalloz Actualité*, 4 mai 2021.

Dautieu T., Gabrié E., « Analyse de l'apport de la loi pour une République numérique à la protection des données à caractère personnel (1re partie) : L'ouverture de l'accès aux données publiques et sa conciliation avec la protection des données à caractère personnel », *Communication Commerce électronique*, décembre 2016, n° 12, étude 22.

Dautieu T., Gabrié E., « Analyse de l'apport de la loi pour une République numérique à la protection des données à caractère personnel (2e partie) : Les droits des personnes et les missions et pouvoirs de la CNIL », *Communication Commerce électronique*, janvier 2017, n° 1, étude 1.

De La Mure A., Mathon S., Fouqueau O., « Quelles limites juridiques à la libération des données ? Comment concilier open data et protection des données à caractère personnel ? », *La Semaine Juridique Administrations et Collectivités territoriales*, 22 janvier 2018, n° 3, 2031.

De Ravel d'Esclapon T., « La gouvernance des algorithmes dans le secteur financier : le point de vue de l'ACPR », *Revue de Droit bancaire et financier*, juillet 2020, n° 4, étude 12.

Derdevet M., « Repenser la mission d'ERDF à l'heure de la transition énergétique », *Énergie – Environnement – Infrastructures*, janvier 2016, n° 1, dossier 5.

Douville T., « L'émergence d'un droit commun de la cybersécurité », *Recueil Dalloz*, 2017, p. 2255.

Dreyfus M., « Principe de prévention », *Dictionnaire Collectivités territoriales et Développement Durable*, 2017, pp. 393-395.

Drobenko B., « La Convention d'Aarhus et le droit français », *Revue Juridique de l'Environnement*, numéro spécial, 1999, p. 37.

Dubreuil C.-A., « Règlementation des marchés financiers », *JurisClasseur Administratif*, LexisNexis, 2010, Fasc. 265.

Duclercq J.B., « Le droit public à l'ère des algorithmes », *Revue de Droit public*, 2017, n°5, p. 1401.

Ebers M., Hoch V.R.S., Rosenkranz F., *et al.*, « The European Commission's Proposal for an Artificial Intelligence Act—A Critical Assessment by Members of the Robotics and AI Law Society (RAILS) », *Multidisciplinary Scientific Journal*, 2021, 4, 589-603, 15.

Ebrahim T.Y., « National Cybersecurity Innovation », *West Virginia Law Review*, 2020, vol. 123, n° 2, pp. 483-546.

Eisenmann C., « Quelques problèmes de méthodologie des définitions et des classifications en science juridique », *Archives de philosophie du droit*, 1966, vol. 11, pp. 27-43.

Emmerechts S., « Droit de l'environnement et droit nucléaire : une symbiose croissante », *Bulletin de droit nucléaire*, 2009, vol. 2008, n°2.

Erdélyi O.J., Goldsmith J., « Regulating Artificial Intelligence: Proposal for a Global Solution », *Proceedings of the 2018 AAAI/ACM Conference on AI, Ethics, and Society*, 2018, 95-101.

Fiorino D.J., « Toward a New System of Environmental Regulation: The Case for an Industry Sector Approach », *Environmental Law*, 1996, vol. 26, n°2, pp. 457-486.

Fisher E., « Precaution, precaution everywhere: developing a "common understanding" of the precautionary principle in the EC », *Maastricht Journal of European and Comparative Law*, 2002, vol. 9, 7–28.

Fjeld J., Hilligoss H., Achten N., *et al.*, « Principled Artificial Intelligence : Mapping Consensus in Ethical and Rights-based Approaches to Principles for AI», *Berkman Klein Center for Internet & Society*, 2020, disponible en ligne : <https://dash.harvard.edu/bitstream/handle/1/42160420/HLS%20White%20Paper%20Final_v3.pdf?sequence=1&isAllowed=y>, consulté le 28 février 2022.

Fourmon A., « Ouverture des données énergétiques et big data », *Énergie - Environnement – Infrastructures*, Avril 2018, n° 4, comm. 22.

Frison Roche M.A., « Le Droit de la compliance », *Recueil Dalloz*, 29 septembre 2016, Etudes et commentaires, n°32, p. 1871.

Frison-Roche M.A., « Le droit de la régulation », *Recueil Dalloz*, 2001, n°7, pp. 610-616.

Gabuthy Y., « Analyse économique du droit : présentation générale », *Economie & Prévision*, 2013, n°202-203, pp. 1-8.

Gaillard E., « Principe de précaution – Droit interne », *JurisClasseur Environnement et Développement durable*, 2014, Fasc. 2410.

Gaillard E., « Principe de précaution : Systèmes juridiques internationaux et européens », *JurisClasseur Environnement et Développement durable*, LexisNexis, 1er mars 2015, Fasc. 4035.

Galan C., « The Certification as a Mechanism for Control of Artificial Intelligence in Europe », *SSRN Electronic Journal*, 11 septembre 2019, disponible en ligne : <https://ssrn.com/abstract=3451741>, consulté le 8 février 2022.

Galbois-Lehalle D., « Responsabilité civile pour l'intelligence artificielle selon Bruxelles : une initiative à saluer, des dispositions à améliorer », *Recueil Dalloz*, 2021, p. 87.

Galland J.-P., « Critique de la notion d'infrastructure critique », *Flux*, 2010, vol. 81, n°3, pp. 6-18, note 18.

Gasser U., « Recoding Privacy Law: Reflections on the Future Relationship among Law, Technology, and Privacy », *Harvard Law Review* : *Law, privacy & technology commentary series*, 2016, 130, n°2, 10.

Gemignani M., « Laying Down the Law to Robots », *San Diego Law Review*, 1984, vol. 21, 1045.

Georgiadis G., Poels G., « Towards a privacy impact assessment methodology to support the requirements of the general data protection regulation in a big data analytics context: A systematic literature review », *Computer Law & Security Review*, avril 2022, vol. 44, 105640.

Gremaud W., « Le droit administratif de la production électronucléaire », *RFDA*, 2021, n°4, p. 711.

Gruson D., « Enjeux juridiques de l'intelligence artificielle en santé : le stable et le mouvant », *Revues des Juristes de Sciences Po*, juin 2016, n° 21, p. 16.

Grynbaum L., « IA et assurance », *RLDA*, septembre 2019, n°151, p. 22, 6782.

Hélaine C., « Adaptation de la garantie légale de conformité pour les biens et les contenus et services numériques », *Dalloz Actualité*, 5 octobre 2021.

Henry X., « Le renouvellement de la jurisprudence. À propos du site de Cerclab », *JCP*, 2020, p. 938.

Hermitte M.-A., David V., « Évaluation des risques et principe de précaution », *LPA*, 30 nov. 2000, p. 13.

Hirsch D.D., « Protecting the Inner Environment: What Privacy Regulation Can Learn from Environmental Law », *Georgia Law Review*, 2006, vol. 41, n°1.

Huglo C., « Comment adapter le droit de l'environnement aux immenses défis lancés à l'humanité par le changement climatique ? », *Bulletin du Droit de l'Environnement Industriel*, 1 mars 2019, n°80.

Huhta K., « Smartening up while keeping safe? Advances in smart metering and data protection under EU law », *Journal of Energy & Natural Resources Law*, 2020, vol. 38, n°1, 5–22.

Jamay F., « Principe de participation », *JurisClasseur Environnement et Développement durable*, version mise à jour du 3 août 2021, Fasc. 2440.

Johnson S.M., « Economics v. Equity II: The European Expérience », *Washington & Lee Law Review*, 2001, vol. 58, 417.

Jourdain P., « Faute délictuelle et manquement contractuel : des relations complexes. Illustration à travers les fautes délictuelles de l'entrepreneur et du mandataire », *Revue Trimestrielle de Droit civil*, 1995, p. 895.

Kaminski M.E., « Binary Governance: Lessons from the GDPR's Approach to Algorithmic Accountability », *Southern California Law Review*, 2019, vol. 92, n°6, 1529–1616.

Kaminski M.E., « The Right to Explanation, Explained », *Berkeley Technology Law Journal*, 2018, vol. 34, n°1.

Kaminski M.E., Malgieri G., « Algorithmic Impact Assessments under the GDPR: Producing Multi-Layered Explanations », *SSRN Electronic Journal*, 2019.

Kaminski M.E., Urban J.M., « The Right to Contest AI », *Columbia Law Review*, 16 novembre 2021, vol. 121, n°7.

Kaufmann C., Weber R.H., « The Role of Transparency in Financial Regulation », *Journal of International Economic Law*, septembre 2010, vol. 13, n°3, pp. 779-797.

Keller J., « Le Data Act : de nouvelles règles de partage des données », *Dalloz Actualité*, 8 mars 2022.

Kramer I.R., « The Birth of Privacy Law: A Century Since Warren and Brandeis », *Cath. U. Law Review*, 1990, 39, 703.

Krämer L., « Transnational Access to Environmental Information », *Transnational Environmental Law*, 2012, vol. 1, n°, 95–104.

Kromarek P., « Le principe de précaution vu par l'industrie », *Droit de l'environnement*, n° 7-8, 2001, p. 189.

Labbée X., « L'homme augmenté », *Recueil Dalloz*, 2012, p. 2323.

Lacroix De Sousa S., « L'octroi de crédit à l'épreuve des droits fondamentaux », *Revue de Droit bancaire et financier*, novembre 2018, n° 6, dossier 41.

Lamoureux M., « La causalité juridique à l'épreuve des algorithmes », *JCPG*, 2016, p.731.

Laprès D.A., « Le droit, la morale et l'éthique dans la gestion des entreprises », *Revue Lamy droit des affaires*, 1er juin 2012, nº 72.

Larrieu J., « Robot et propriété intellectuelle », *Dalloz IP/IT*, 2016, p. 291.

Lasmoles O., « Difficulties faced by the legal system in coming to terms with blockchains », *Revue internationale de droit économique*, 2018, vol. 4, 453-469.

Levin H.J., « The Limits of Self-Regulation », *Columbia Law Review*, avril 1967, vol. 67, n° 4, pp. 603-644.

Loiseau G., « Des robots et des hommes », *Recueil Dalloz*, 2015, p. 2369.

Loiseau G., Bourgeois M., « Du robot en droit à un droit des robots », *JCP*, 2014, p. 1231.

Maggiolino M., « EU Trade Secrets Law and Algorithmic Transparency », *Bocconi Legal Studies Research Paper*, 31 mars 2019, n° 3363178, disponible en ligne : <https://ssrn.com/abstract=3363178>, consulté le 7 août 2021.

Magrani E., « New Perspectives on Ethics and the Laws of Artificial Intelligence », *Journal of Internet Regulation*, 2019, vol. 8, n°3, 19.

Maisnier-Boché L., « Anonymisation : que faire pour sortir de l'impasse ? », *Expertises des systèmes d'information*, septembre 2016, n° 416, 296–300.

Maisnier-Boche L., Botchorichvili N., « Projet de loi « pour une République numérique » : quels impacts sur la protection des données personnelles ? », *RLDI*, janvier 2016, n°3911.

Marignol L., « Principe de responsabilité et action en responsabilité dans le Règlement général sur la protection des données », *RLDI*, 1er janvier 2020, n°166.

Marsden C., « Internet co-regulation and constitutionalism : Towards European judicial review », *International Review of Law Computers and Technology*, 2012, vol. 26, n°2, 211-228.

Martial-Braz N., « L'apport de l'intelligence artificielle à la banque », *Revue de Droit bancaire et financier*, novembre 2019, n°6, comm. 78.

Martial-Braz N., « L'intelligence artificielle bientôt régulée. L'incidence du AI Act dans le secteur financier », *Revue de Droit bancaire et financier*, mai 2021, n°3, comm. 81.

Martial-Braz N., « Le profilage », *Communication Commerce Electronique*, avril 2018, p. 70 et s., dossier spécial « Entrée en vigueur du Règlement général sur la protection des données ».

Martial-Braz N., « Les nouveaux droits des individus consacrés par la loi pour une République numérique. Quelles innovations ? Quelle articulation avec le Règlement européen ? », *Dalloz IP/IT*, 2016, p. 525.

Martin G.J., « Les angles morts de la doctrine juridique environnementaliste », *Revue juridique de l'environnement*, Lavoisier, 2020/1, 45, no 67-80.

Mathey N., Bourdeaux G., « Vers une régulation des FinTechs ? », *Revue de Droit bancaire & financier*, mars 2017, vol. 2, dossier 15.

Mathis B., « Proposition de règlement européen sur l'intelligence artificielle : le regard d'un praticien », *RLDI*, 2022, n°192, 4179, p. 44.

Maupin E., « Adoption de la loi sur la cyber-sécurité, les armes à feu et Galiléo », *AJDA*, 2018, p. 368.

Maury E., « Actualité Informatique et Libertés – Transfert de données à caractère personnel vers des pays tiers », *AJDA*, 2022, p. 1433

Maynard A.H., « Environnement : protection contre les pollutions et nuisances », *Dalloz*, 1999, n°5450, 35 et s.

Mazeau L., « Intelligence artificielle et responsabilité civile : le cas des logiciels d'aide à la décision en matière médicale », *Revue pratique de la prospective et de l'innovation*, Avril 2018, n°1, p. 6.

Mekki M., « Les fonctions de la responsabilité civile à l'épreuve du numérique : l'exemple des logiciels prédictifs », *Dalloz IP/IT*, 2020, p. 672.

Mendoza-Caminade A., « Le droit confronté à l'intelligence artificielle des robots : vers l'émergence de nouveaux concepts juridiques ? », *Recueil Dalloz*, 2016, p. 445.

Meneceur Y., « DataJust face aux défis de l'intelligence artificielle », *JCP*, 2020, p. 1087.

Migayron S., « Pratique contentieuse : Intelligence artificielle : qui sera responsable ? », *Communication commerce électronique*, avril 2018, n°4.

Moliner-Dubost M., « Nucléaire », *JurisClasseur Administratif*, version mise à jour le 14 avril 2021, Fasc. 378.

Molinier J., « Primauté du droit de l'Union européenne », *Répertoire de droit européen*, Dalloz, septembre 2011, pp. 2–20.

Morel-Maroger J., « La protection des données personnelles des clients des banques : Bilan et perspectives », *Revue de Droit bancaire et financier*, mars 2011, p. 7 et s.

Mylrea M., « Smart energy-internet-of-things opportunities require smart treatment of legal, privacy and cybersecurity challenges », *The Journal of World Energy Law & Business*, avril 2017, vol. 10, n°2, pp. 147–158.

Pagallo U., Casanovas P., Madelin R., « The Middle-out Approach: Assessing Models of Legal Governance in Data Protection, Artificial Intelligence, and the Web of Data », *The Theory and Practice of Legislation*, 2 janvier 2019, vol. 7, n°1, 25.

Pasquale F., « Restoring Transparency to Automated Authority », *Journal on Telecommunications and High Technology Law*, 2011, vol. 9, n° 235, pp. 235-256.

Pavlou P.A., « Consumer Acceptance of Electronic Commerce: Integrating Trust and Risk with the Technology Acceptance Model », *International Journal of Electronic Commerce*, 2003, vol. 7, n°3, pp. 101-134.

Petit Y., « Environnement », *Répertoire de droit international*, Dalloz, Janvier 2010, pp. 89-92.

Pike H., Khan F., Amyotte P., « Precautionary Principle (PP) versus As Low As Reasonably Practicable (ALARP): Which one to use and when », *Process Safety and Environmental Protection*, 2020, vol. 137, 158-168.

Poncelet C., « Access to Justice in Environmental Matters : Does the European Union Comply with its Obligations ? », *Journal of Environmental Law*, 2012, vol. 24, n°2, 287–309.

Pontier J.M., « La certification, outil de modernité normative », *Recueil Dalloz*, 1996, n°41, pp. 355-360.

Poulenard H., « L'encadrement juridique des algorithmes mis en œuvre sur les marchés financiers », *Revue Lamy droit des affaires*, 1er décembre 2018, nº 143.

Quoy N., Boullet A., « Pratique contractuelle. L'encadrement contractuel de l'intelligence artificielle », *Communication Commerce électronique*, février 2020, n°2, 3.

Randrianirina I., « Plaidoyer pour un nouveau droit de propriété intellectuelle sur les productions générées par intelligence artificielle », *Recueil Dalloz*, 2021, p. 91.

Rapp L., Terneyre P., « Démarche de/par la conformité… diversité des expressions », *Lamy droit public des affaires*, pp. 615-616.

Reboul-Maupin N., « La prévention des risques technologiques : aspects juridiques », *Les Petites Affiches*, 16 décembre 2004, n° 251, pp. 6-13.

Rechtschaffen C., « Deterrence vs. Cooperation and the Evolving Theory of Environmental Enforcement », *Southern California Law Review*, 1997, n°71, pp. 1181-1272.

Ringe W.G., Ruof C., « Regulating fintech in the EU: the Case for a Guided Sandbox », *European Journal of Risk Regulation*, septembre 2020, vol. 11, n°3, pp. 604-629.

Rochfeld J., « L'encadrement des décisions prises par algorithme », *Dalloz IP/IT*, 2018, p. 474.

Rousseau S., « La réglementation des cryptomonnaies (ICOs) au Canada : la protection des investisseurs et le bon fonctionnement du marché dans le bac à sable réglementaire », *Revue internationale des services financiers*, 2018, n°1, p. 15.

Rouvière F., « L'intelligence artificielle au risque du mythe », *Revue Trimestrielle de Droit Civil*, Dalloz, 2020, p. 990.

Ruiz Fabri H., « La prise en compte du principe de précaution par l'OMC », *Revue juridique de l'environnement*, 2000, numéro spécial, p. 55.

Salzman J., « Creating Markets for Ecosystem Services: Notes from the Field », *New York University Law Review*, 2005, vol. 80, 870.

Saud Neto S.J., Tollet N., « Cartographie des risques : Retour d'expérience brésilien pour réussir l'exercice », *Cahiers de droit de l'entreprise*, mars 2017, n° 2, dossier 7.

Savin P., Martinet Y., « Risques technologiques et réparation des dommages : points saillants de la loi du 30 juillet 2003 », *Les Petites Affiches*, 10 octobre 2003, n°203, p. 4.

Scarpelli U., « La natura della metodologica giuridica », *Rivista internazionale de filosofia del Diritto*, 1956, XXXIII, pp. 247–255.

Schauss M., « Systèmes experts et droit », *Revue interdisciplinaire d'études juridiques*, 1987, vol. 18, n°1, pp. 101-114.

Schwartz P.M., Janger E.J., « Notification of Data Security Breaches », *Michigan Law Review*, 2006, n°105, 913-984.

Sénéchal J., « Responsabilisation ab initio, régulation ex ante et responsabilités a posteriori : le coeur des débats européens sur les systèmes d'intelligence artificielle, hors et dans le secteur du commerce électronique », *Dalloz IP/IT*, 2020, p. 667.

Shea T., « Regulation of Takeovers in the United Kingdom », *Brooklyn Journal of International Law*, 1990, vol. 16, n°1.

Smatt-Pinelli S., « Focus : des regulatory sandbox pour l'innovation juridique », *Revue pratique de la prospective et de l'innovation*, LexisNexis, juillet 2021, n°1, p. 3.

Smuha N.A., « Beyond the Individual: Governing AI's Societal Harm », *Internet Policy Review*, 30 septembre 2021, vol. 10, n°3.

Smuha N.A., Ahmed-Rengers E., Harkens A., *et al.*, « How the EU Can Achieve Legally Trustworthy AI: A Response to the European Commission's Proposal for an Artificial Intelligence Act », *SSRN Electronic Journal*, 5 août 2021, disponible en ligne : <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3899991>, consulté le 21 janvier 2022.

Smyth S.M., « The Greening of Canadian Cyber Laws: What Environmental Law can Teach and Cyber Law can learn », *International Journal of Cyber Criminology*, 2014, vol. 8, n°2, 111-155.

Spagnuelo D., Ferreira A., Lenzini G., « Accomplishing Transparency within the General Data Protection Regulation », *Proceedings of the 5th International Conference on Information Systems Security and Privacy* (ICISSP 2019), pp. 114-125.

Stein A.L., « Artificial Intelligence and Climate Change », *Yale Journal on Regulation*, 2020, 37, n°890.

Stewart R.B., « A New Generation of Environmental Regulation ? », *Capital University Law Review*, 2001, vol. 29, 21.

Stone C., « Is there a precautionary principle ? », *Environmental Law Reporter*, 2001, vol. 7, p. 107.

Surden H., « Artificial Intelligence and Law: An Overview », *Georgia State University Law Review*, 28 juin 2019, vol. 35, 1305.

Sutterlin O., « Synthèse – Principe de précaution », *LexisNexis*, Encyclopédies, 2019.

Tarello G., « Sur la spécificité du raisonnement juridique », *Archives de philosophie du droit et de philosophie sociale*, 1972, n°7, p. 105.

Trébulle F.G., « Droit de l'environnement », *Recueil Dalloz*, 2010, n°2468.

Tunc A., « Garde du comportement et garde de la structure dans la responsabilité du fait des choses inanimées », *JCP*, 1957, I, p. 1384.

Untermaier-Kerléo E., « Les nouveaux visages de la décision administrative : d'une administration assistée à une administration automatisée », *La Semaine Juridique Administrations et Collectivités territoriales*, 17 décembre 2018, n° 50, p. 2339.

Veale M., Zuiderveen Borgesius F., « Demystifying the Draft EU Artificial Intelligence Act », *Computational Law Review International*, juillet 2021, vol. 22, n°4.

Villafranco J., « Self-Regulation in the Big Data and AI Space », *The Judges' Journal*, 2020, vol. 59, n°1, 32-35.

Viney G., « Le principe de précaution et la responsabilité civile des personnes privée », *Recueil Dalloz*, 2007, p. 1542.

Vivant M., « Publication des travaux du CNNum sur l'environnement et le numérique », *Revue Lamy Droit de l'Immatériel*, 1er août 2020, n°173.

Vivant M., « Le droit des systèmes-experts », *Congrès sur l'information et la documentation*, 1987, 7, pp. 159-163.

Wachter S., Mittelstadt B., Floridi L., « Why a right to explanation of automated decision-making does not exist in the General Data Protection Regulation », *International Data Privacy Law*, 2017, 7(2), 76–99.

Wachter S., Mittelstadt B., Russell C., « Counterfactual explanations without opening the black box: Automated decisions and the GDPR », *Harvard Journal of Law & Technology*, 2017, vol. 31, n°2, 841–888.

Warren S.D., Brandeis L.D., « The Right to Privacy », *Harvard Law Review*, 1890, vol. 4, N°5, 193-220.

Werbach K., « Higher Standards Regulation in the Network Age », *Harvard Journal of Law & Technology*, 2009, vol. 23, n°1, 179.

Yeung K., Howes A., Pogrebna G., « AI Governance by Human Rights-Centred Design, Deliberation and Oversight: An End to Ethics Washing », in *The Oxford Handbook of AI Ethics*, dir. M. Dubber, F. Pasquale, Oxford University Press, 2019.

Zuiderveen Borgesius F., « Strengthening legal protection against discrimination by algorithms and artificial intelligence », *The International Journal of Human Rights*, 2020, 24:10, 1572-1593.

B/ Sources extra-juridiques

Abioye O.S., Oyedele L.O., Akanbi L., *et al.*, « Artificial intelligence in the construction industry: A review of present status, opportunities and future challenges », *Journal of Building Engineering*, 2021, vol. 44, 103299.

Abowd J.M., Vilhuber L., « How Protective Are Synthetic Data ? », in *Privacy in Statistical Databases*, dir. J. Domingo-Ferrer, Y. Saygın, Springer, Berlin, 2008, 5262.

Ahlgren B., Hidell M., Ngai E., « Internet of Things for Smart Cities: Interoperability and Open Data », *IEEE Internet Computing*, 2016, vol. 20, n°6, pp. 52-56.

Amable B., Demmou L., Ledezma I., « L'impact de la règlementation sur l'innovation : une analyse des performances selon la proximité à la frontière technologique », *Economie et prévision*, 2011, 1-2, 197-198.

Andrae A.S.G., Edler T., « On Global Electricity Usage of Communication Technology: Trends to 2030 », *Challenges*, 2015, vol. 6, n°1, pp. 117-138.

Astion M.L., Wener M.H., Thomas R.G., *et al.*, « Overtraining in neural networks that interpret clinical data », *Clinical Chemistry*, Volume 39, Issue 9, 1 September 1993, pp. 1998-2004.

Aven T., « On the ethical justification for the use of risk acceptance criteria », *Risk Analysis*, 2007, vol. 27, n°2, pp. 303-312.

Awad E., Dsouza S., Kim R., *et al.*, « The Moral Machine experiment », *Nature*, 2018, 563, 59-64.

Barbu E.M., Feleaga N., Feleaga L., « Quelles normes IAS/IFRS utiliser pour le reporting environnemental ? », *Revue Française de Comptabilité*, février 2011, n°440.

Barnes S.B., « A privacy paradox: Social networking in the United States », *First Monday*, 2006, vol. 11, n°9.

Benbouzid B., « Quand prédire, c'est gérer. La police prédictive aux États-Unis », *Réseaux*, 2018, vol.5, n° 211, pp. 221-256.

Bernstone C., Heyden A., « Image analysis for monitoring of crack growth in hydropower concrete structures », *Measurement*, juillet 2009, vol. 42, n°6, pp. 878–893.

Bertail P., Bounie D., Clémençon S., Waelbroeck P., « Algorithmes : biais, discrimination et équité », *Telecom Paris Tech*, février 2019, disponible en ligne : <https://www.telecom-paris.fr/wp-content-EvDsK19/uploads/2019/02/Algorithmes-Biais-discrimination-equite.pdf>, consulté le 16 mars 2020.

Bilbao I., Bilbao J., « Overfitting problem and the over-training in the era of data: Particularly for Artificial Neural Networks », *2017 Eighth International Conference on Intelligent Computing and Information Systems (ICICIS)*, 2017, pp. 173-177.

Bolón-Canedo V., Sánchez-Maroño N., Alonso-Betanzos A., « A review of feature selection methods on synthetic data », *Knowledge and Information System*, 2013, vol. 34, 483–519.

Borning A., « What Pushes Back Considering Materiality in IT ? », *Limits*, 2018, disponible en ligne : <https://computingwithinlimits.org/2018/papers/limitsl8-borning.pdf>, consulté le 7 avril 2021.

Boujemaa N., « La définition de l'intelligence artificielle, enjeu juridico-commercial », *Le Monde*, 15 décembre 2021.

Bourg D., « L'éco-scepticisme et le refus des limites », *Études*, 2010, vol. 7, n°413, p. 29.

Boyer-Allirol B., « Faut-il mieux réglementer le reporting extrafinancier ? », *Revue française de gestion*, 2013, vol. 8, n°237, pp. 73-95.

Brown S., Davidovic J., Hasan A., « The algorithm audit: Scoring the algorithms that score us », *Big Data & Society*, 2021, 1-8.

Bulut M., Özcan E., « A new approach to determine maintenance periods of the most critical hydroelectric power plant equipment », *Reliability Engineering & System Safety*, 2021, vol. 205, 107238.

Buzelay A., « A propos du secteur des hautes technologies en Europe », *Revue de l'Union européenne*, 2022, p. 167.

Camero A., Alba E., « Smart City and information technology: A review », *Cities*, 2019, vol. 93, pp. 84-94.

Cardon D., Cointet J. P., Mazières A., « La revanche des neurones : L'invention des machines inductives et la controverse de l'intelligence artificielle », *Réseaux*, 2018, n° 211, 5, 173.

Cath C., Wachter S., Mittelstadt B., *et al.*, « Artificial intelligence and the 'Good Society' : the US, EU, and UK approach », *Science and Engineering Ethics*, 2018, vol. 24, 505-528.

Chambru M., « La publicisation du risque nucléaire par les usages protestataires du droit », *Sciences de la société*, 2017, 100, pp. 79-91.

Chambru M., « La protestation antinucléaire par-delà les frontières : mutations et temporalités de l'enjeu européen », *Communication & Organisation*, Presses Universitaires de Bordeaux, 2020, vol. 57, pp. 121-133.

Chapi K., Singh V.P., Shirzadi A., *et al.*, « A novel hybrid artificial intelligence approach for flood susceptibility assessment », *Environmental Modelling & Software*, 2017, vol. 95, pp. 229–245.

Chen C., Duan S., Cai T., Liu B., « Online 24-h solar power forecasting based on weather type classification using artificial neural network », *Solar Energy*, novembre 2011, vol. 85, n°11, pp. 2856–2870.

Chen Z., Liu B., « Lifelong Machine Learning », *Synthesis Lectures on Artificial Intelligence and Machine Learning*, 2nd ed., août 2018, vol. 12, n°3, pp. 1-207.

Choné F., « L'énergéticien du XXIe siècle : le numérique au service du consommateur et de la transition énergétique, *Annales des Mines - Responsabilité et environnement*, 2017, vol. 87, n°3, 43.

Clare E.M., « Moral Crumple Zones: Cautionary Tales in Human-Robot Interaction », *Engaging Science, Technology, and Society*, 23 mars 2019, n°5.

Coase R.H., « The nature of the firm », *Economica*, novembre 1937, vol. 4, n° 16, 386-405.

Corchado J.M., Aiken J., « Hybrid artificial intelligence methods in oceanographic forecast models », *IEEE Transactions on Systems, Man, and Cybernetics*, novembre 2022, vol. 32, n°4, pp. 307–313.

Cortade T., Poudou J.-C., « Les plateformes numériques d'échange d'électricité », *Annales des Mines – Enjeux numériques*, septembre 2021, n°15.

D'Amico B., Myers R.J., Sykes J., *et al.*, « Machine Learning for Sustainable Structures: A Call for Data », *Structures*, juin 2019, vol. 19, n°1, 4.

Dhar P., « The Carbon Impact of Artificial Intelligence », *Nature Machine Intelligence*, 2020, vol. 2, n°8, 423.

Dhar V., « The future of Artificial Intelligence », *Big Data*, 2016, vol. 4, n°1.

Di Santo K.G., Di Santo S.G., Monaro R.M., *et al.*, « Active demand side management for households in smart grids using optimization and artificial intelligence », *Measurement*, 2018, vol. 115, pp. 152–161.

Dobbe R., « System Safety and Artificial Intelligence », *2022 ACM Conference on Fairness, Accountability, and Transparency (FAccT '22) Proceedings*, Association for Computing Machinery, 2022, New York, USA, p. 1584.

Došilović F.K., Brčić M., Hlupić N., « Explainable artificial intelligence: A survey », *41st International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO)*, 2018, pp. 0210-0215.

Falco G., Shneiderman B., Badger J., *et al.*, « Governing AI safety through independent audits », *Nature Machine Intelligence*, 2021, vol. 3, n°7, pp. 566–571.

Faria E., Pereira C., « Nuclear fuel loading pattern optimisation using a neural network », *Annals of Nuclear Energy*, 2003, vol. 30, pp. 603–613.

Ferreboeuf H., « Pour une sobriété numérique », *Futuribles*, 2019, vol. 429, n°2, 15.

Fjelland R., « Why general artificial intelligence will not be realized », *Humanities Social Sciences Communications*, 2020, vol. 7, n°10.

Fourquet-Courbet M.-P., Courbet D., « Anxiété, dépression et addiction liées à la communication numérique », *Revue française des sciences de l'information et de la communication*, 2017, 11.

Gade K., Geyik S.C., Kenthapadi K., *et al.*, « Explainable AI in Industry », *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (KDD '19)*, 2019, Association for Computing Machinery, New York, 3203.

Galperin A., Nissan E., « Application of a heuristic search method for generation of fuel reload configurations », *Nuclear Science Engineering*, 1988, vol. 99, pp. 343-352.

Giannakaris P., Trakadas P., Zahariadis T., *et al.*, « Using Smart Contracts in Smart Energy Grid Applications », *Proceedings of the International Scientific Conference*, Novi Sad, Serbia, Singidunum University, 2019, 597–602.

Giles M., « Is Al the Next Big Climate-Change Threat? We Haven't a Clue », *MIT Technology Review*, 29 juillet 2019, disponible en ligne : <https://www.technologyreview.com/2019/07/29/663/ai-computing-cloud-computing-microchips>, consulté le 5 avril 2021.

Gomez-Fernandez M., Higley K., Tokuhiro A., *et al.*, « Status of Research and Development of Learning-Based Approaches in Nuclear Science and Engineering: A Review », *Nuclear Engineering and Design*, avril 2020, vol. 359, 110479.

Grenard A., « Normalisation, certification : quelques éléments de définition », *Revue d'économie industrielle*, 1996, vol. 75, pp. 45-60.

Guiraud E., « Le rôle de l'éthique dans la mise en place d'une certification pour l'utilisation d'algorithmes dans le système juridique », *Ethique publique*, 2019, vol. 21, n°1.

Gunning D., Stefik M., Choi J., *et al.*, « XAI : Explainable artificial intelligence », *Science Robotics*, 18 décembre 2019, vol. 4, n°37.

Gupta A., Vedaldi A., Zisserman A., « Synthetic Data for Text Localisation in Natural Images », *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, pp. 2315-2324.

Hall R.E., Bowerman B., Braverman J., *et al.*, « The vision of a smart city », *2nd International Life Extension Technology Workshop*, Paris, 28 septembre 2000.

Hardin G., « The Tragedy of the Commons », *Science*, 1968, vol. 162, 1243.

Harvey S., « Des réseaux intelligents, stables et fiables : le rôle des réseaux intelligents et de l'électronucléaire dans les systèmes énergétiques bas carbone », *IAEA Bulletin*, septembre 2020, vol. 61, n°3.

Hashemian H. M., « State-of-the-Art Predictive Maintenance Techniques », *IEEE Transactions on Instrumentation and Measurement*, Janvier 2011, vol. 60, n°1, pp. 226–236.

Helm D., « Regulatory Reform, Capture, and the Regulatory Burden », *Oxford Review of Economics*, 2006, vol. 22, n°169.

Henry J.-Y., « La sûreté des logiciels dans les installations nucléaires », *Revue Contrôle de l'Autorité de sûreté nucléaire*, octobre 1999, n°131, Octobre 1999, Dossier : Les systèmes informatiques dans l'industrie nucléaire.

Heurix J., Zimmermann P., Neubauer T., Fenz S., « A taxonomy for privacy enhancing technologies », *Computers & Security*, 2015, vol. 53, pp. 1–17.

Hinds J., Williams E. J., Joinson A. N., « 'It wouldn't happen to me' : Privacy concerns and perspectives following the Cambridge Analytica scandal », *International Journal of Human-Computer Studies*, 2020, 143, 102498.

Il I., Yongbeom K., Hyo-Joo H., « The effects of perceived risk and technology type on users' acceptance of technologies », *Information & Management*, 2008, vol. 45, n°1, pp. 1-9.

Ivanov S., « Artificial Intelligence will be able to deal with blackouts in power grids », *Hi-Tech*, 2019, n°1, pp. 28-31.

Jain A., Patel H., Nagalapatti L., *et al.*, « Overview and Importance of Data Quality for Machine Learning Tasks », *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, 2020, 3561.

Ji-Hwan L., Chi-Hoon S., « Effects of trust and perceived risk on user acceptance of a new technology service », *Social Behavior and Personality : an international Journal*, 2013, vol. 41, n°4, pp. 586-597.

Jobin A., Ienca M., Vayena E., « The Global Landscape of AI Ethics Guidelines », *Nature Machine Intelligence*, septembre 2019, vol. 1, n°9, 389-399.

Kalfon L., « Des machines à gouverner ? », *Le Monde*, 11 octobre 1980, archives.

Kalogirou S.A., « Artificial neural networks in renewable energy systems applications: a review », *Renewable and Sustainable Energy Reviews*, décembre 2001, vol. 5, n°4, pp. 373–401.

Kaptein M., Eckles D., « Selecting Effective Means to Any End: Futures and Ethics of Persuasion Profiling », *Persuasive Technology*, 2010, 82–93.

Kim H.G., Chang S.H., Lee B.H., « Optimal fuel loading pattern design using an artificial neural network and a fuzzy rule-based system », *Nuclear Science Engineering*, 1993, vol. 115, pp. 152-163.

Kruse J., Schäfer B., Witthaut D., « Revealing drivers and risks for power grid frequency stability with explainable AI », *Patterns*, vol. 2, n°11, 2021.

Kurukuru V.S.B., Haque A., Khan M.A., *et al.*, « A Review on Artificial Intelligence Applications for Grid-Connected Solar Photovoltaic Systems », *Energies*, 2021, vol. 14, n°15, 4690.

Kwet M., « Digital colonialism: US empire and the new imperialism in the Global South », *Race & Class*, 2019, vol. 60, n°4, 3-26.

Lagarrigue J., Lebe G., « Ethique ou morale », *Recherche & Formation*, 1997, 24, pp. 121-130.

Laperrière R., « L'informatique et les droits des personnes », *Cahiers de recherche sociologique*, 1993, n°21, pp. 53–77.

Lavrijssen S., Carrillo Parra A., « Radical prosumer innovations in the electricity sector and the impact on prosumer regulation », *Sustainability*, 2017, vol. 9, 1207.

Lavrijssen S., Espinosa Apráez B., ten Caten T., « The Legal Complexities of Processing and Protecting Personal Data in the Electricity Sector », *Energies*, 2022, vol. 15, n°3, 1088.

Lehman-Wilzig S.N., « Frankenstein unbound: Towards a legal definition of artificial intelligence », *Futures*, 1981, vol. 13, n°6, pp. 442-457.

Leveson N., « The use of safety case in certification and regulation », *MIT ESD Working paper series*, novembre 2011, n°2011-13.

Louis V., Baron C., « Vers une certification continue des logiciels critiques en aéronautique », *Techniques de l'ingénieur*, coll. Technologies logicielles Architectures des systèmes, novembre 2019, n°h8060.

Lundberg S.M., Erion G., Chen H., *et al.*, « From local explanations to global understanding with explainable AI for trees », *Nature Machine Intelligence*, 2020, vol. 2, 56–67.

MacCarthy M., « An Examination of the Algorithmic Accountability Act of 2019 », *The Transatlantic Working Group Papers Series*, 24 octobre 2019.

Mahmoud M.A., Nasir N.R., Gurunathan M., *et al.*, « The Current State of the Art in Research on Predictive Maintenance in Smart Grid Distribution Network: Fault's Types, Causes, and Prediction Methods – A Systematic Review », *Energies*, 2021, vol. 14, 5078.

Maisnier-Boché L., « Anonymisation : que faire pour sortir de l'impasse ? », *Expertises des systèmes d'information*, septembre 2016, n° 416, 296–300.

Mandard S., « Le gouvernement présente un plan sans contrainte pour limiter l'impact écologique du numérique », *Le Monde*, 24 février 2021.

Manokha I., « Le scandale Cambridge Analytica contextualisé : le capital de plateforme, la surveillance et les données comme nouvelle « marchandise fictive » », *Cultures & Conflits*, 2018, 109, 39-59.

Marot A., Rozier A., Dussartre M., *et al.*, « Towards an AI assistant for human grid operators », *Hybrid Human Artificial Intelligence (HHAI) Conference*, juin 2022, Amsterdam.

McNally P., Inayatullah S., « The rights of robots: Technology, culture and law in the 21st century », *Futures*, 1988, Vol. 20, n°2, pp. 119-136.

Mechin A., Pioch S., « Séquence ERC : comment améliorer l'utilisation des méthodes de dimensionnement de la compensation écologique ? », *Revue électronique en sciences de l'environnement*, décembre 2014, vol. 14, n°3.

Mogilenko A., « Application of artificial intelligence algorithms in the global energy industry », *Energy and Industry of Russia*, 2018, n°7.

Mohaghegh S.D., « Recent Developments in Application of Artificial Intelligence in Petroleum Engineering », *Petroleum Tech*, 2005, 57, 86.

Ntoutsi E., Fafalios P., Gadiraju U., *et al.*, « Bias in data-driven artificial intelligence systems : an introductory survey », *WIREs Data Mining Knowlede Discovery*, décembre 2020, 10:1356.

Orts E.W., « A Reflexive Model of Environmental Regulation », *Business Ethics Quarterly*, octobre 1995, vol. 5, n°4, pp. 779-794.

Parisi G.I., Kemker R., Part J.L., *et al.*, « Continual lifelong learning with neural networks : A review », *Neural Networks*, 2019, vol. 113, pp. 54-71.

Pégny M., Ibnouhsein I., « Quelle transparence pour les algorithmes d'apprentissage machine ? », *Revue d'intelligence artificielle,* 28 août 2018, 32, n°4, 447, 78.

Pereira R., Couto M., Ribeiro F., *et al.*, « Energy efficiency across programming languages: how do energy, time, and memory relate ? », *Proceedings of the 10th ACM SIGPLAN International Conference on Software Language Engineering (SLE 2017)*, Association for Computing Machinery, New York, NY, USA, 256–267.

Photopoulos J., « Fighting algorithm bias », *Phys. World*, 2021, vol. 34, n°5, 42.

Pinto G., Castor F., « Energy efficiency: A new concern for application software developers », *Communications of the ACM*, Novembre 2017.

Ramchurn S. D., Vytelingum P., Rogers A., Jennings N. R., « Putting the 'smarts' into the smart grid: a grand challenge for artificial intelligence », *Communications of the ACM*, 2012, vol. 55, n°4, pp. 86–97.

Raza M. Q., Khosravi A., « A review on artificial intelligence based load demand forecasting techniques for smart grid and buildings », *Renewable and Sustainable Energy Reviews*, 2015, vol. 50, pp. 1352–1372.

Reforgiato Recupero D., « Toward a Green Internet », *Science*, 29 mars 2013, 339, n°6127, 1533.

Rigas E.S., Ramchurn S.D., Bassiliades N., « Managing Electric Vehicles in the Smart Grid Using Artificial Intelligence: A Survey », *IEEE Transactions on Intelligent Transportation Systems*, août 2015, vol. 16, n°4, pp. 1619–1635.

Roberts H., Cowls J., Morley J., *et al.*, « The Chinese approach to artificial intelligence : an analysis of policy, ethics and regulation », *AI & Society*, 2021, vol. 36, 59-77.

Rolnick D., Donti P.L., Kaack L.H., *et al.*, « Tackling Climate Change with Machine Learning », *ArXiv*, 5 novembre 2019, 1906.05433, disponible en ligne : <http://arxiv.org/abs/1906.05433>, consulté le 8 avril 2021.

Rudin C., Waltz D., Anderson R.N., « Machine Learning for the New York City Power Grid », *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2012, vol. 34, n°2, pp. 328–345.

Rudolph A., Voget S., Mottok J., « A consistent safety case argumentation for artificial intelligence in safety related automotive systems », *ERTS 2018*, janvier 2018, Toulouse, France, hal-02156048.

Rugraff E., « La politique industrielle de l'UE face à son décrochage technologique », *Bulletin de l'Observatoire des Politiques Économiques en Europe*, 2019, vol. 41, 33–45.

Sankaran S., Zhang C., Gutierrez Lopez M., *et al.*, « Respecting Human Autonomy through Human-Centered AI », *Proceedings of the 11th Nordic Conference on Human-Computer Interaction: Shaping Experiences, Shaping Society (NordiCHI '20)*, 2020, Association for Computing Machinery, New York, 134.

Schlesinger A., O'Hara K.P., Taylor A.S., « Let's Talk About Race: Identity, Chatbots, and AI », *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, Association for Computing Machinery, New York, 2018, Paper 315, 1–14.

Schwarzschild A., Goldblum M., Gupta A., *et al.*, « Just How Toxic is Data Poisoning? A Unified Benchmark for Backdoor and Data Poisoning Attacks », *Proceedings of the 38th International Conference on Machine Learning*, PMLR, 2021, vol. 139, 9389-9398.

Siegrist M., « The influence of trust and perceptions of risks and benefits on the acceptance of gene technology », *Risk Analysis*, 2000, vol. 20, n°2, pp. 195-203.

Sjöberg J., Ljung L., « Overtraining, regularization and searching for a minimum, with application to neural networks », *International Journal of Control*, 1995, vol. 62, n°6, pp. 1391-1407.

Slovic P., Peters E., « Risk Perception and Affect », *Current Directions in Psychological Science*, 2006, vol. 15, 6, 322-325.

Snow J., « How Artificial Intelligence Can Tackle Climate Change », *National Geographic*, 18 juillet 2019, disponible en ligne : <https://www.nationalgeographic.com/environment/2019/07/artificial-intelligence-climate-change>, consulté le 5 avril 2021.

Solans V., Rochman D., Brazell C., *et al.*, « Optimisation of used nuclear fuel canister loading using a neural network and genetic algorithm », *Neural Computer & Applications*, 2021, vol. 33, 16627–16639.

Sozontov A., Ivanona M., Gibadullin A., « Implementation of artificial intelligence in the electric power industry », *E3S Web of Conferences*, 2019, vol. 114, n°01009.

Steiner F., « L'industrie de l'électricité : réglementation, structure du marché et performances », *Revue économique de l'OCDE*, 2001, vol. nº32, n°1, pp. 159–201.

Strubel E., « Energy and Policy Considerations for Deep Learning in NLP », *57th Annual meeting of the Association for Computational Linguistics*, 5 juin 2019.

Suman S., « Artificial Intelligence in Nuclear Industry: Chimera or Solution? », *Journal of Cleaner Production*, 1er janvier 2021, vol. 278, 124022.

Swire P.P., Litan R.E., « None of your business: world data flows, electronic commerce, and the European privacy directive », *Brookings inst. Press*, 1998.

Toews R., « Deep Learning's Carbon Emissions Problem », *Forbes*, 17 juin 2020, disponible en ligne : <https://www.forbes.com/sites/robtoews/2020/06/17/deep-learnings-climate-change-problem/>, consulté le 7 mai 2021.

Truby J., « Decarbonizing Bitcoin: Law and Policy Choices for Reducing the Energy Consumption of Blockchain Technologies and Digital Currencies », *Energy Research & Social Science*, octobre 2018, vol. 44, 399.

Van der Mei A., Doomernik J.-P., « Artificial intelligence potential in power distribution system planning », *24th International Conference & Exhibition on Electricity Distribution (CIRED)*, 12–15 juin 2017, session 5.

Venema L., « Code of conduct for using AI in healthcare », *Nature Machine Intelligence*, 2019, vol. 1, 265-266.

Vinuesa R., Azizpour H., Leite I., *et al.*, « The Role of Artificial Intelligence in Achieving the Sustainable Development Goals », *Nature Communications*, décembre 2020, 11, n°1, 233.

Wang L., « Issues on software testing for safety-critical real-time automation systems », *The 23rd Digital Avionics Systems Conference*, IEEE, 2004, Cat. n°04CH37576, pp. 530-101.

Wang Y., Kosinski M., « Deep Neural Networks are more accurate that humans at detecting sexual orientation from facial images », *Journal of personality and social psychology*, février 2018, vol. 114, n°2, 246-257.

Wassyng A., Maibaum T., Lawford M., Bherer H., « Software Certification : Is There a Case against Safety Cases ? », in *Foundations of Computer Software. Modeling, Development, and Verification of Adaptive Systems*, dir. R. Calinescu, E. Jackson, Monterey Workshop 2010, Lecture Notes in Computer Science, Springer, 2011, vol. 6662.

Yampolskiy R., « Unexplainability and Incomprehensibility of AI », *Journal of Artificial Intelligence and Consciousness*, juillet 2020, vol. 7, n°2, 1-15.

Yeung A.W.K., « The "As Low As Reasonably Achievable" (ALARA) principle: a brief historical overview and a bibliometric analysis of the most cited publications », *Radioprotection*, juin 2019, vol. 54, n°2, p. 103-109.

Zhang C., Wu J., Long C., *et al.*, « Review of Existing Peer-to-Peer Energy Trading Projects », *Elsevier Energy Procedia*, 2017, vol. 105, 2563–2568.

§IV : Rapports, livres blancs, référentiels et communications

ACCESS NOW, *Two years under the EU GDPR: an implementation progress report*, Rapport, 2020, disponible en ligne : <https://www.accessnow.org/cms/assets/uploads/2020/05/Two-Years-Under-GDPR.pdf>, consulté le 24 juin 2022.

ACPR, *Gouvernance des algorithmes d'intelligence artificielle dans le secteur financier*, Document de réflexion, disponible en ligne : <https://acpr.banque-france.fr/sites/default/files/medias/documents/20200612_gouvernance_evaluation_ia.pdf>, consulté le 24 mai 2021.

ADA LOVELACE INSTITUTE, *Examining the black box : Tools for assessing algorithmic systems*, Rapport, 2020, disponible en ligne : <https://www.adalovelaceinstitute.org/wp-content/uploads/2020/04/Ada-Lovelace-Institute-DataKind-UK-Examining-the-Black-Box-Report-2020.pdf>, consulté le 3 août 2021.

ADEME, *L'analyse du cycle de vie*, dossier thématique, 18 juin 2018, disponible en ligne : <https://www.ademe.fr/expertises/consommer-autrement/passer-a-laction/dossier/lanalyse-cycle-vie/quest-lacv>, consulté le 7 mai 2021.

ADEME, *La face cachée du numérique – Réduire les impacts du numérique sur l'environnement*, rapport, janvier 2021.

ADEME, *TIC et impacts environnementaux*, dossier thématique, disponible en ligne : < https://communication-responsable.ademe.fr/digital-eco-responsable/tic-et-impacts-environnementaux/tic-axes-daction-pour-reduire-les-impacts>, consulté le 6 mai 2021.

AFNOR, *ISO-TR 14069, « Guide d'application de la norme 14064-1 WD3 »*, Mars 2011.

AGENCE DES SYSTEMES D'INFORMATION PARTAGES DE SANTÉ, *Cadre d'interopérabilité des SIS*, référentiel technique, 13 novembre 2012, V1.3.1, 24 p.

Aghion P., Bergeaud A., Van Reenen J., « The impact of regulation on innovation », Document de travail pour la Banque de France, janvier 2021, n°804, disponible en ligne : <https://publications.banque-france.fr/leffet-des-regulations-sur-linnovation>, consulté le 24 janvier 2022.

AI NOW INSTITUTE, *2019 Annual Report*, rapport, décembre 2019, disponible en ligne : < https://ainowinstitute.org/AI_Now_2019_Report.pdf>, consulté le 12 février 2022.

AIEA, *La sûreté des installations nucléaires*, 1993, collection Sécurité n°110.

AIEA, *Logiciels destinés aux systèmes programmés important s pour la sûreté des centrales nucléaires*, 2004, collection Normes de sûreté n° NS-G-1.1.

AIEA, *Systèmes d'instrumentation et de contrôle-commande importants pour la sûreté des centrales nucléaires*, 2005, collection Normes de sûreté nº NS-G-1.3.

AMERICAN SOCIETY OF MECHANICAL ENGINEERS (ASME), *Standard for Verification and Validation in Computational Fluid Dynamics and Heat Transfer*, VV 20–2009(R2021), 100 p., disponible en ligne : <https://www.asme.org/codes-standards/find-codes-standards/v-v-20-standard-verification-validation-computational-fluid-dynamics-heat-transfer>, consulté le 18 septembre 2022.

ANSSI, « L'homologation de sécurité », *Guide pratique*, Version 1.0, août 2014, n°20140821-1128.

ANSSI, « Recommandations pour la mise en place de cloisonnement système », *Guide pratique*, 14 décembre 2017, ANSSI-PG-040.

ANSSI, *Maîtriser la SSI pour les systèmes industriels : la cybersécurité des systèmes industriels*, guide pratique, 2012, disponible en ligne : <https://www.ssi.gouv.fr/uploads/IMG/pdf/2012-06-19_CP-Guide_SCADA.pdf>, consulté le 7 juillet 2022.

ANSSI, *Rapport d'activité 2018*, p. 50, disponible en ligne : <https://www.ssi.gouv.fr/uploads/2019/04/anssi_rapport_annuel_2018.pdf>, consulté le 18 septembre 2022.

ANSSI, *Recommandations pour la protection des systèmes d'information essentiels*, guide ANSSI, 18 décembre 2020, ANSSI-PA-085, Annexe A.

ANSSI, *Référentiel « PRIS v2 » - Prestataires de réponse aux incidents de sécurité*, Référentiel d'exigences techniques, 2 août 2017, 53 p.

ANSSI, *Référentiel d'exigences « Prestataires de services d'informatique en nuage » (SecNumCloud)*, version 3.2 du 21 septembre 2021.

ANSSI, *Référentiel général de sécurité v2.0*, Référentiel d'exigences techniques, 13 juin 2014, 25 p.

ARCEP, *L'empreinte environnementale des réseaux*, dossier thématique sur le site officiel de l'Arcep, 19 mars 2021, disponible en ligne : <https://www.arcep.fr/la-regulation/grands-dossiers-thematiques-transverses/lempreinte-environnementale-des-reseaux.html>, consulté le 1er mai 2021.

ASN, *Conception des réacteurs à eau sous pression*, Guide ASN n°22, 19 juillet 2017.

ASN, *Qualification des outils de calcul scientifique utilisés dans la démonstration de sûreté nucléaire – 1re barrière*, Guide pratique, 2017, n° 28.

ASSEMBLÉE NATIONALE, *Archives des débats parlementaires sur le projet de loi Informatique et Libertés du 4 octobre 1977*, publiées au JORF n°79 A.N. du 5 octobre 1977.

Bellot L., *De la smart city au territoire d'intelligence(s) : l'avenir de la smart city*, rapport au Premier ministre, avril 2017, disponible en ligne : <https://www.gouvernement.fr/sites/default/files/document/document/2017/04/rapport_smart_city_luc_belot_avril_2017_definitif.pdf>, consulté le 28 avril 2022.

Bouchoux C., *Refonder le droit à l'information publique à l'heure du numérique : un enjeu citoyen, une opportunité stratégique*, Rapport d'information n° 589 fait au nom de la mission commune d'information du Sénat sur l'accès aux documents administratifs, 5 juin 2014.

Brundage M., Avin S., Clark J., *et al.*, *The malicious use of artificial intelligence : forecasting, prevention, mitigation*, rapport, février 2018, 101 p., disponible en ligne : <https://arxiv.org/ftp/arxiv/papers/1802/1802.07228.pdf>, consulté le 14 novembre 2019.

C. Villani, *Donner un sens à l'intelligence artificielle*, Rapport dans le cadre d'une mission parlementaire du 8 septembre 2017 au 8 mars 2018 confiée par le Premier Ministre Edouard Philippe, *La Documentation Française*, 8 mars 2018.

CAPGEMINI INVENT, *Open data maturity report 2021*, rapport commandité par la Commission européenne, 7ème ed., 17 décembre 2021.

CEN et CENELEC, *Position Paper – Proposal for a Regulation laying down harmonized rules on artificial intelligence (Artificial Intelligence Act-COM 2021/206)*, papier de position, octobre 2021, disponible en ligne : <https://www.cencenelec.eu/media/CEN-CENELEC/AreasOfWork/Position%20Paper/2021/positionpaper_aia_2021.pdf>, consulté le 12 février 2022.

CENTER FOR DATA INNOVATION, *How much will the Artificial Intelligence Act cost Europe ?*, Rapport, juillet 2021, disponible en ligne : <https://www2.datainnovation.org/2021-aia-costs.pdf>, consulté le 9 juin 2022.

CEPD, *Recommandations n°01/2020 sur les mesures qui complètent les instruments de transfert destinés à garantir le respect du niveau de protection des données à caractère personnel de l'UE*, 10 novembre 2020.

CISPE.CLOUD, *Code de Conduite CISPE relatif à la Protection des Données*, 9 février 2021, disponible en ligne : <https://www.cnil.fr/sites/default/files/atoms/files/code_de_conduite_des_fournisseurs_dinfrastructures_cloud_relatif_a_la_protection_des_donnees_-_cispe_-_version_francaise.pdf>, consulté le 25 janvier 2022.

CJUE, *Press release n°105/22*, Case C-817/19, Ligue des droits humains, 21 juin 2022.

CNIL, *Comment permettre à l'homme de garder la main ? Les enjeux éthiques des algorithmes et de l'intelligence artificielle*, Synthèse du débat public animé par la CNIL dans le cadre de la mission de réflexion éthique confiée par la loi pour une République Numérique, 15 décembre 2017.

CNIL, *Pack de conformité sur les compteurs communicants*, mai 2014, disponible en ligne : <https://www.cnil.fr/sites/default/files/typo/document/Pack_de_Conformite_COMPTEURS_COMMUNICANTS.pdf>, consulté le 11 mai 2022.

CNRS, *Comment l'intelligence artificielle va changer nos vies*, dossier thématique, disponible en ligne : <https://lejournal.cnrs.fr/dossiers/comment-lintelligence-artificielle-va-changer-nos-vies>, consulté le 3 mars 2022.

COLOMBUS CONSULTING, *Data & IA : Quels enjeux et bénéfices concrets pour le secteur de l'énergie ?*, rapport, 2019, disponible en ligne : <https://colombus-consulting.com/nos-publications/data-ia-quels-enjeux-et-benefices-concrets-pour-le-secteur-energie/>, consulté le 5 février 2020.

COMMISSION DE L'ENVIRONNEMENT, DE LA SANTÉ PUBLIQUE ET DE LA SÉCURITÉ ALIMENTAIRE (ENVI), *Avis sur la proposition de règlement du Parlement européen et du Conseil établissant des règles harmonisées concernant l'intelligence artificielle (législation sur l'intelligence artificielle) et modifiant certains actes législatifs de l'Union (COM(2021)0206 – C9-0146/2021 – 2021/0106(COD))*.

COMMISSION DE RÉGULATION DE L'ÉNERGIE (CRE), *Rapport du comité d'études relatif aux données dont disposent les gestionnaires de réseaux et d'infrastructures d'énergie*, rapport, 18 mai 2017.

COMMISSION DE RÉGULATION DE L'ÉNERGIE (CRE), *Synthèse de l'étude sur les perspectives stratégiques de l'énergie*, rapport, 2018, disponible en ligne : <http://fichiers.cre.fr/Etude-perspectives-strategiques/1SyntheseGenerale/Perspectives_Strategiques_du_secteur_de_l_energie_Synthese_generale_FR.pdf>, consulté le 5 février 2020.

COMMISSION ÉCONOMIQUE DES NATIONS UNIES POUR L'EUROPE, *La Convention d'Aarhus : guide d'application*, Rapport, 2ème édition, 2014, 280 p., disponible en ligne : <https://unece.org/DAM/env/pp/Publications/Aarhus_Implementation_Guide_FRE_interactive.pdf>, consulté le 11 décembre 2021.

COMMISSION EUROPÉENNE, « Commission collects views on making liability rules fit for the digital age, Artificial Intelligence and circular economy », *Site officiel de la Commission*

européenne, 19 octobre 2021, disponible en ligne : <https://digital-strategy.ec.europa.eu/en/news/commission-collects-views-making-liability-rules-fit-digital-age-artificial-intelligence-and>, consulté le 13 février 2022.

COMMISSION EUROPÉENNE, « Commission proposes measures to boost data sharing and support European data spaces », *Communiqué de presse accompagnant la proposition du Data Governance Act*, 25 novembre 2020.

COMMISSION EUROPÉENNE, « De nouvelles règles et actions en faveur de l'excellence et de la confiance dans l'intelligence artificielle », *Site de la représentation française auprès de la Commission européenne*, 21 avril 2021.

COMMISSION EUROPÉENNE, « Questions & Answers: EU-U.S. Data Privacy Framework, draft adequacy decision », *Site officiel de la Commission européenne*, 13 décembre 2022, disponible en ligne : <https://ec.europa.eu/commission/presscorner/detail/en/qanda_22_7632>, consulté le 16 janvier 2023.

COMMISSION EUROPÉENNE, *Action plan on the digitalisation of the energy sector*, 22 juillet 2021, Ares(2021)4720847.

COMMISSION EUROPÉENNE, *Communication de la Commission du 28 avril 2017 sur l'accès à la justice en matière d'environnement*, C(2017) 2616 final.

COMMISSION EUROPÉENNE, *Communication from the Commission to the Council and the European Parliament on Environmental Agreements*, 1996, COM (96)561.

COMMISSION EUROPÉENNE, *Communication sur le recours au principe de précaution*, COM(2000) 1 final, 2 février 2000.

COMMISSION EUROPÉENNE, *Digitalising the Energy System*, Consultation publique pour un plan d'action sur la numérisation du secteur de l'énergie, octobre 2021.

COMMISSION EUROPÉENNE, *Le pacte vert européen*, Communication de la Commission au Parlement européen, au Conseil européen, au Conseil, au Comité économique et social européen et au Comité des régions, 11 décembre 2019, COM(2019) 640 final.

COMMISSION EUROPÉENNE, *L'intelligence artificielle pour l'Europe*, Communication de la Commission au Parlement européen, au Conseil européen, au Conseil, au Comité économique et sociale européen et au Comité des régions, 25 avril 2018, COM (2018) 237 final.

COMMISSION EUROPÉENNE, *Livre blanc du 19 février 2020 sur l'Intelligence Artificielle – Une approche européenne axée sur l'excellence et la confiance*, 19 février 2020, COM(2020) 65.

COMMISSION EUROPÉENNE, *Public consultation on the AI White Paper : Final report*, novembre 2020, disponible en ligne : < https://digital-strategy.ec.europa.eu/en/white-paper-artificial-intelligence-public-consultation-towards-european-approach-excellence-and>, consulté le 5 juin 2022.

COMMISSION EUROPÉENNE, *Rapport sur les conséquences de l'intelligence artificielle, de l'internet des objets et de la robotique sur la sécurité et la responsabilité*, Rapport de la Commission au Parlement européen, au Conseil et au Comité économique et social européen, 19 février 2020, COM(2020) 64 final.

COMMISSION EUROPÉENNE, *Un plan coordonné dans le domaine de l'intelligence artificielle*, Communication de la Commission au Parlement européen, au Conseil européen, au Conseil, au Comité économique et social européen et au Comité des régions, 7 décembre 2018, COM (2018) 795 final.

COMMISSION EUROPÉENNE, *Une stratégie européenne pour les données*, communication de la Commission au Parlement européen, au Conseil, au Comité économique et social européen et au Comité des régions, 19 février 2020, COM(2020) 66 final.

COMMISSION EUROPÉENNE, *Data protection as a pillar of citizens' empowerment and the EU's approach to the digital transition – two years of application of the GDPR*, document de travail accompagnant la communication de la Commission au Parlement et au Conseil, 24 juin 2020, COM(2020) 264 final.

COMMITTE ON LEGAL AFFAIRS (JURI), *Draft opinion on the proposal for a regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union Legislative Acts (COM(2021)0206 – C9-0146/2021 – 2021/0106(COD)*, 2 mars 2022, 2021/0106(COD).

COMMITTEE ON INDUSTRY, RESEARCH AND ENERGY (ITRE), *Draft opinion on the proposal for a regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts (COM(2021)0206 – C9-0146/2021 – 2021/0106(COD))*, 3 mars 2022, 2021/0106(COD).

COMMITTEE ON THE INTERNAL MARKET AND CONSUMER PROTECTION (IMCO) AND COMMITTEE ON CIVIL LIBERTIES, JUSTICE AND HOME AFFAIRS (LIBE), *Draft report on the proposal for a regulation of the European Parliament and of the Council on harmonised rules on Artificial Intelligence (Artificial Intelligence Act) and amending certain Union Legislative Acts (COM2021/0206 – C9-0146/2021 – 2021/0106(COD)*, 20 avril 2022, 2021/0106(COD).

CONSEIL D'ÉTAT, *Intelligence artificielle et action publique : construire la confiance, servir la performance*, Etude à la demande du premier Ministre, 30 août 2022, 360 p.

CONSEIL D'ÉTAT, *Le numérique et les droits fondamentaux,* Etude annuelle 2014, La Documentation française, 64, 447 p.

CONSEIL D'ÉTAT, *Révision de la loi de bioéthique : quelles options pour demain ?*, Etude à la demande du Premier Ministre par la section du rapport et des études, 28 juin 2018.

CONSEIL DE L'EUROPE, *Algorithmes et droits humains*, Etude menée par le comité d'experts sur les intermédiaires d'internet MSI-NET, 2017, DGI (2017)12, disponible en ligne : <https ://rm.coe.int/algorithms-and-human-rights-fr/1680795681>, consulté le 8 mars 2020.

CONSEIL DE L'EUROPE, *Lignes directrices sur l'intelligence artificielle et la protection des données*, Rapport, novembre 2019, disponible en ligne : <https://rm.coe.int/2018-lignes-directrices-sur-l-intelligence-artificielle-et-la-protecti/168098e1b8>, consulté le 3 août 2021.

CONSEIL DE L'EUROPE, *Possible introduction of a mechanism of certifying artificial intelligence tools and services in the sphere of justice and the judiciary*, Etude de faisabilité de la Commission européenne pour l'efficacité de la justice, 8 décembre 2020, CEPEJ(2020)15Rev, disponible en ligne : <https://rm.coe.int/feasability-study-en-cepej-2020-15/1680a0adf4>, consulté le 8 février 2022.

CONSEIL DE L'UE, *Conclusions de la présidence – La charte des droits fondamentaux dans le contexte de l'intelligence artificielle et du changement numérique*, 11481/20, 2020.

CONSEIL DE L'UE, *Orientation générale sur la Proposition de règlement du Parlement européen et du Conseil établissant des règles harmonisées concernant l'intelligence artificielle (législation sur l'intelligence artificielle) et modifiant certains actes législatifs de l'Union*, 25 novembre 2022, 2021/0106(COD).

CONSEIL DE L'UE, *Presidency compromise text (Articles 16-29) on the proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (AI Act) and amending certain Union legislative acts*, n° 2021/0106(COD), 3 février 2022, n°5756/22, Présidence française.

CONSEIL DE L'UE, *Presidency compromise text (Articles 40-52) on the proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (AI Act) and amending certain Union legislative acts*, n° 2021/0106(COD), 15 février 2022, n°6239/22, Présidence française.

CONSEIL DE L'UE, *Presidency compromise text on the proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial*

intelligence (AI Act) and amending certain Union legislative acts, 29 Novembre 2021, n° 2021/0106(COD), n°14278/21, Présidence slovène.

CONSEIL DE L'UE, « Le Conseil approuve l'acte sur la gouvernance des données », *Communiqué de presse*, 16 mai 2022.

CONSEIL NATIONAL DU NUMÉRIQUE, *Ambition numérique – pour une politique française et européenne de la transition numérique*, rapport remis au Premier Ministre, juin 2015.

CONSEIL NATIONAL DU NUMÉRIQUE, *Feuille de route sur l'environnement et le numérique - 50 mesures pour un agenda national et européen sur un numérique responsable c'est-à-dire sobre et au service de la transition écologique et solidaire et des objectifs de développement durable*, Rapport remis à la ministre de la Transition écologique et solidaire et au secrétaire d'État chargé du Numérique, juillet 2020, disponible en ligne : <https://cnnumerique.fr/environnement_numerique>, consulté le 01/04/2021.

COUR D'APPEL DE PARIS, *La réforme du droit français de la responsabilité civile et les relations économiques*, Rapport, avril 2019, disponible en ligne : <http://www.justice.gouv.fr/art_pix/Rapport_CA_PARIS_reforme_responsabilite_civile.pdf> , consulté le 20 mai 2021.

DGE, *Intelligence artificielle : État de l'art et perspectives pour la France*, rapport commandité par le Pôle interministériel de prospective et d'anticipation des mutations économiques, février 2019, 333 p.

ENISA, *Securing Machine Learning Algorithms*, rapport, 14 décembre 2021, 70 p., disponible en ligne : <https://www.enisa.europa.eu/publications/securing-machine-learning-algorithms>, consulté le 12 février 2022.

ETSI, *Securing Artificial Intelligence (SAI) : Problem Statement*, rapport, décembre 2020, GR SAI 004, V1.1.1.

Eurelectric, *AI Insights : The Power Sector in a Post-Digital Age*, Rapport, 26 novembre 2020, disponible en ligne : <https://www.eurelectric.org/media/5016/ai-insights-final-report-26112020.pdf>, consulté le 6 mai 2021.

EUROPEAN PARLIAMENTARY RESEARCH SERVICE, *The impact of the GDPR on AI*, Rapport de recherche, Juin 2020, n° 641530.

Euzet C., *Rapport fait au nom de la commission des lois constitutionnelles, de la législation et de l'administration générale de la république, sur le projet de loi, adopté par le sénat après engagement de la procédure accélérée, portant diverses dispositions d'adaptation au droit de l'Union européenne dans le domaine de la sécurité (n° 530)*, texte n°554.

EXPERT GROUP ON LIABILITY AND NEW TECHNOLOGIES, *Liability for AI and other emerging digital technologies*, 21 novembre 2019, disponible en ligne : <https://op.europa.eu/en/publication-detail/-/publication/1c5e30be-1197-11ea-8c1f-01aa75ed71a1/language-en/format-PDF>, consulté le 18 janvier 2020.

GAIA-X, *Policy Rules Document*, 21 avril 2022, disponible en ligne : < https://gaia-x.eu/sites/default/files/2022-04/Gaia-X_Policy%20Rules_Document_v22.04_Final.pdf>, consulté le 3 mai 2022.

GOOGLE, *Consultation on the EU AI Act Proposal : Google's submission*, 15 juillet 2021, disponible en ligne : <https://ec.europa.eu/info/law/better-regulation/have-your-say/initiatives/12527-Artificial-intelligence-ethical-and-legal-requirements/F2662492_en>, consulté le 19 juin 2022.

GOOGLE, *Consultation on the white paper on AI - a European approach, Google's submission*, 28 mai 2020, disponible en ligne : <https://www.blog.google/documents/77/Googles_submission_to_EC_AI_consultation_1.pdf/>, consulté le 5 juin 2022.

Green Concept, *Livre blanc sur l'écoconception numérique*, 21 février 2020, disponible en ligne : <http://www.greenconcept-innovation.fr/wp-content/uploads/2020/02/greenconcept_21022020.pdf>, consulté le 7 mai 2021.

GREENPEACE, *Clicking Clean*, Rapport, 2017, disponible en ligne : <http://www.clickclean.org/france/fr/>, consulté le 8 avril 2021.

GREENPEACE, *Greenpeace Report : Oil in the Cloud*, Rapport, 19 mai 2020, disponible en ligne : <https://www.greenpeace.org/usa/reports/oil-in-the-cloud/>, consulté le 8 avril 2021.

GROUPE CONSULTATIF INTERNATIONAL POUR LA SÛRETÉ NUCLÉAIRE, *Basic Safety Principles for Nuclear Power Plants*, 1999, 75-INSAG-3 Rev. 1, INSAG-12.

GROUPE CONSULTATIF INTERNATIONAL POUR LA SÛRETÉ NUCLÉAIRE, *La défense en profondeur en sûreté nucléaire*, 1997, INSAG-10.

GROUPE D'EXPERTS DE HAUT NIVEAU EN IA, *A Definition of AI : Main capabilities and disciplines*, Rapport du groupe d'experts indépendants établis par la Commission européenne, 8 avril 2019, disponible en ligne : <https://digital-strategy.ec.europa.eu/en/library/definition-artificial-intelligence-main-capabilities-and-scientific-disciplines>, consulté le 18 janvier 2020.

GROUPE D'EXPERTS DE HAUT NIVEAU EN IA, *Lignes directrices du 8 avril 2019 pour une IA digne de confiance*, 8 avril 2019.

GROUPE D'EXPERTS DE HAUT NIVEAU EN IA, *Policy and Investment Recommendations for Trustworthy AI*, 26 juin 2019.

GROUPE D'EXPERTS DE HAUT NIVEAU EN IA, *The assessment list for trustworthy artificial intelligence (ALTAI) for self-assessment*, 14 septembre 2020.

GROUPE DE TRAVAIL « ARTICLE 29 » SUR LA PROTECTION DES DONNÉES, *Lignes directrices relatives à la prise de décision individuelle automatisée et au profilage aux fins du règlement (UE) 2016/679*, 3 octobre 2017, n°WP251.

GROUPE DE TRAVAIL « ARTICLE 29 » SUR LA PROTECTION DES DONNÉES, *Lignes directrices concernant l'analyse d'impact relative à la protection des données (AIPD) et la manière de déterminer si le traitement est « susceptible d'engendrer un risque élevé » aux fins du règlement (UE) 2016/679*, adoptées le 4 avril 2017, WP 248 rév. 01.

HEALTH AND SAFETY EXECUTIVE, « The Tolerability of Risk from Nuclear Power Stations », Document officiel, 1988, disponible en ligne : <https://www.onr.org.uk/documents/tolerability.pdf>, consulté le 16 mai 2020.

HEALTH AND SAFETY EXECUTIVE, *Reducing risks, protecting people : HSE's decision making process*, HSE books, 2001, 88 p., disponible en ligne : <https://www.hse.gov.uk/managing/theory/r2p2.pdf>, consulté le 9 septembre 2021.

ICO et ALAN TURING INSTITUTE, *Explaining decisions made with AI : Part 1*, draft guidance, 2019, version 1.1., disponible en ligne : <https://ico.org.uk/media/about-the-ico/consultations/2616434/explaining-ai-decisions-part-1.pdf>, consulté le 23 septembre 2022.

Iddri, GreenIT, FING et WWF France, *Livre blanc Numérique et environnement*, 2018, disponible en ligne : <https://www.wwf.fr/sites/default/files/doc-2018-03/180319_livre_blanc_numerique_environnement.pdf>, consulté le 11 mars 2020.

INSTITUT MONTAIGNE, *Algorithmes : contrôle des biais S.V.P.*, rapport, mars 2020.

INSTITUT MONTAIGNE, *Transition énergétique : faisons jouer nos réseaux*, rapport, 2019, disponible en ligne : <https://www.institutmontaigne.org/ressources/pdfs/publications/transition-energetique-faisons-jouer-nos-reseaux-rapport.pdf>, consulté le 5 février 2020.

IRSN, *Approche comparative entre sûreté et sécurité nucléaires*, Rapport IRSN 2009/117, 21 avril 2009, 26 p.

ISO, *ISO 14064-1:2018, « Gaz à effet de serre — Partie 1: Spécifications et lignes directrices, au niveau des organismes, pour la quantification et la déclaration des émissions et des suppressions des gaz à effet de serre »*.

ISO, *Norme ISO 14044:2006 « Management environnemental »*.

ISO, *Norme ISO 5001:2018 « Systèmes de management de l'énergie ».*

ISO, *Norme ISO/CEI 51:2014* « *Aspects liés à la sécurité – Principes directeurs pour les inclure dans les normes ».*

ISO, *Norme ISO/IEC 2382:2015, Technologies de l'information.*

ISO, *Norme ISO/IEC 27000:2018 « Technologies de l'information – Techniques de sécurité – Systèmes de management de la sécurité de l'information – Vue d'ensemble et vocabulaire »*, février 2018, 5ème ed., 29 p.

ISO, *Norme ISO/IEC 27001 « Management de la sécurité de l'information »*, octobre 2013, 2ème ed., 23 p.

ISO, *Norme ISO/TR 14062:2002 « Management environnemental — Intégration des aspects environnementaux dans la conception et le développement de produit ».*

Lepetit J.-F., *Rapport sur le risque systémique*, rapport du Ministère de l'Économie, de l'Industrie et de l'Emploi, *Documentation française*, 2010, 108 p.

Mamalet F., Jenn E., Flandin G., *et al.*, *Machine learning in certified systems*, livre blanc, ANITI – IRT Saint-Exupéry, 2021.

Martuzzi M., Tickner J.A., *The Precautionary Principle: Protecting Public Health, the Environment and the Future of Our Children*, World Health Organization Regional Office for Europe, 2004, disponible en ligne : < https://www.euro.who.int/__data/assets/pdf_file/0003/91173/E83079.pdf>, consulté le 30 avril 2021.

MICROSOFT, *Microsoft Responsible AI Standard v2*, document officiel, juin 2022, p. 21, « Reliability & Safety Goals », disponible en ligne : <https://query.prod.cms.rt.microsoft.com/cms/api/am/binary/RE4ZPmV>, consulté le 7 juillet 2022.

NATIONAL RESEARCH COUNCIL, *Assessing the Reliability of Complex Models: Mathematical and Statistical Foundations of Verification, Validation, and Uncertainty Quantification*, The National Academies Press, 2012, 144 p.

NATIONAL TRANSPORTATION SAFETY BOARD (NTSB), « 'Inadequate Safety Culture' Contributed to Uber Automated Test Vehicle Crash - NTSB Calls for Federal Review Process for Automated Vehicle Testing on Public Roads », *Communiqué de presse du NTSB Office of Safety Recommendations and Communications*, 19 novembre 2019.

Rolls Royce, *The Aletheia Framework*, 14 décembre 2020, disponible en ligne : <https://www.rolls-royce.com/~/media/Files/R/Rolls-Royce/documents/stand-alone-pages/aletheia-framework-booklet-2021.pdf>, consulté le 28 février 2022.

SECRETARIAT GENERAL DE LA DEFENSE NATIONALE, *Instruction générale interministérielle relative à la sécurité des activités d'importance vitale*, n°6600/SGDSN/PSE/PSN, 7 janvier 2014.

SECRETARIAT GENERAL DE LA DEFENSE NATIONALE, *La défense en profondeur appliquée aux systèmes d'information*, Mémento, 19 juillet 2004, version 1.1, disponible en ligne : <https://www.ssi.gouv.fr/uploads/IMG/pdf/mementodep-v1-1.pdf>, consulté le 16 septembre 2022.

SENAT, *Etude d'impact sur le projet de loi portant diverses dispositions d'adaptation au droit de l'union européenne dans le domaine de la sécurité*, 17 novembre 2017, NOR : INTX1728622L/Bleue-1.

SENAT, *Projet de loi de finances 2021 : Economie*, Avis n°139 au nom de la commission des affaires économiques sur le projet de loi de finances, adopté par l'Assemblée nationale, pour 2021.

SENAT, *Rapport d'information sur l'empreinte environnementale du numérique*, rapport sénatorial, 24 juin 2020.

SERVICE DE RECHERCHE DU PARLEMENT EUROPEEN, *Le principe de précaution – définitions, applications et gouvernance*, décembre 2015, PE 573.876, disponible en ligne : <https://www.europarl.europa.eu/RegData/etudes/IDAN/2015/573876/EPRS_IDA%282015%29573876_FR.pdf>, consulté le 30 avril 2021.

SOCIETE FRANÇAISE DES MECANICIENS, *Guide de validation des progiciels de calcul des structures*, guide AFNOR technique, 1990, 372 p.

Thales, *Charte éthique du numérique*, document officiel, septembre 2021, disponible en ligne : <https://www.thalesgroup.com/sites/default/files/2021-10/Charte%20%C3%A9thique%20du%20num%C3%A9rique.pdf>, consulté le 11 octobre 2021.

THE NORVEGIAN DATA PROTECTION AUTHORITY, *Artificial intelligence and privacy*, Rapport, Janvier 2018, disponible en ligne : <https://www.datatilsynet.no/globalassets/global/english/ai-and-privacy.pdf>.

The Shift Project, *Climat : l'insoutenable usage de la vidéo en ligne*, rapport, juillet 2019, disponible en ligne : <https://theshiftproject.org/wp-content/uploads/2019/07/2019-01.pdf>, consulté le 17 mars 2020.

The Shift Project, *Déployer la sobriété numérique*, rapport, disponible en ligne : <https://theshiftproject.org/wp-content/uploads/2020/01/2020-01.pdf>, consulté le 11 mars 2020.

The Shift Project, *Pour une sobriété numérique*, octobre 2018, disponible en ligne : <https://theshiftproject.org/wp-content/uploads/2018/11/Rapport-final-v8-WEB.pdf>, consulté le 10 avril 2021.

U.S. CONGRESS, OFFICE OF TECHNOLOGY ASSESSMENT, *Environmental Policy Tools: A User's Guide*, U.S. Government Printing Office, OTA-ENV-634, Washington DC, September 1995, 224p.

UK OFFICE FOR ARTIFICIAL INTELLIGENCE, « New UK initiative to shape global standards for Artificial Intelligence », *Communiqué de presse*, 12 janvier 2022, disponible en ligne : <https://www.gov.uk/government/news/new-uk-initiative-to-shape-global-standards-for-artificial-intelligence>, consulté le 18 juillet 2022.

Van der Leyen U., *A Union that strives for more. My agenda for Europe*, Political guidelines for the next European Commission 2019-2024, 2019.

Villani C., *Donner un sens à l'intelligence artificielle*, Rapport dans le cadre d'une mission parlementaire du 8 septembre 2017 au 8 mars 2018 confiée par le Premier Ministre Edouard Philippe, La Documentation Française, 8 mars 2018.

Viney G., Kourilsky P., *Le principe de précaution*, Rapport au Premier Ministre, 15 octobre 1999.

Zuiderveen Borgesius F., *Discrimination, Artificial Intelligence and Algorithmic Decision-Making*, Rapport à la demande du Conseil de l'Europe, 2018, disponible en ligne : <https://rm.coe.int/discrimination-artificial-intelligence-and-algorithmic-decision-making/1680925d73>, consulté le 26 novembre 2021.

§V : Articles et publications web

ADEME, « Base Carbone », *Bilan GES, Site de l'ADEME (blog)*, disponible en ligne : <https://www.bilans-ges.ademe.fr/fr/accueil/contenu/index/page/presentation/siGras/0>, consulté le 8 octobre 2021.

ADEME, « Bilan GES Organisation », *Bilan GES – Site de l'ADEME (blog)*, disponible en ligne : <https://www.bilans-ges.ademe.fr/fr/accueil/contenu/index/page/bilan%2Bges%2Borganisation/siGras/1>, consulté le 8 octobre 2021.

ADEME, « Etapes d'un bilan GES », *Bilan GES – Site de l'ADEME (blog)*, disponible en ligne : <https://www.bilans-

ges.ademe.fr/fr/accueil/contenu/index/page/Etapes%2Bbilan%2BGES/siGras/0>, consulté le 8 octobre 2021.

AFNOR, « EDF obtient la certification ISO 50001 pour son parc de datacenters », *communiqué de presse*, 19 février 2016, disponible en ligne : <https://www.afnor.org/presse_fev2016/edf-obtient-la-certification-iso-50001-pour-son-parc-de-datacenters/>, consulté le 7 mai 2021.

AFNOR, *Normes en cours de conception pour le terme « intelligence artificielle »*, disponible en ligne : <https://norminfo.afnor.org/search?term=intelligence+artificielle>, consulté le 12 février 2022.

AIEA, « Specific Safety Guides », *Site de l'AIEA (blog)*, disponible en ligne : <https://www.iaea.org/publications/search/type/safety-standards-series>, consulté le 16 septembre 2022.

AMAZON WEB SERVICES, « FAQ sur la confidentialité des données », *Site d'AWS (blog)*, disponible en ligne : <https://aws.amazon.com/fr/compliance/data-privacy-faq/>, consulté le 16 janvier 2023.

AMAZON WEB SERVICES, « Your Guide to AI and machine learning at re:Invent 2018", *Site d'AWS (blog)*, 27 septembre 2018, disponible en ligne : < https://aws.amazon.com/fr/blogs/machine-learning/your-guide-to-ai-and-machine-learning-at-reinvent-2018/>, consulté le 8 avril 2021.

Amodei D., Hernandez D., « Al and Compute », *Open AI* (blog), 16 mai 2018, disponible en ligne : <https://openai.com/blog/ai-and-compute/#fn2>, consulté le 7 avril 2021.

Angwin J., Larson J., Mattu S., Kirchner L., « Machine Bias : There's software used across the country to predict future criminals, and it's biased against blacks », *ProPublica*, 2016, disponible en ligne : <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>, consulté le 17 mars 2020.

ANSSI, « Les règles de sécurité », *Site de l'ANSSI (blog)*, disponible en ligne : <https://www.ssi.gouv.fr/administration/protection-des-oiv/les-regles-de-securite/>, consulté le 14 septembre 2022, concernant les SIIV.

ARBORUS, ORANGE, *Charte internationale pour une IA inclusive*, disponible en ligne : <https://charteia.arborus.org/>, consulté le 25 janvier 2022.

ARCEP, « Bac à sable règlementaire », *Site officiel de l'ARCEP*, 28 juin 2022, disponible en ligne : <https://www.arcep.fr/professionnels/startups-entrepreneurs/bac-a-sable-reglementaire.html>, consulté le 22 septembre 2022.

Balenieri R., « L'Arcep, en passe d'obtenir de nouveaux pouvoirs en matière d'environnement », *Les Echos*, 2 avril 2021, disponible en ligne :

<https://www.lesechos.fr/tech-medias/hightech/larcep-en-passe-dobtenir-de-nouveaux-pouvoirs-en-matiere-denvironnement-1303980>, consulté le 1er mai 2021.

Bertuzzi L., « Les institutions européennes donnent le coup d'envoi des négociations sur une loi visant les Big Tech », *Euractiv (blog)*, 11 janvier 2022, disponible en ligne : <https://www.euractiv.fr/section/economie/news/eu-institutions-kick-off-negotiations-on-law-targeting-big-tech/>, consulté le 6 juin 2022.

Boston Consulting Group, « Des experts de haut niveau en IA lancent CodeCarbon, un outil révolutionnaire de réduction des émissions de CO2 liées à l'informatique », *BCG (blog)*, 1 décembre 2020, disponible en ligne : <https://www.bcg.com/fr-fr/press/1december2020-top-ai-experts-launch-codecarbon>, consulté le 7 mai 2021.

Bruna G.B., « Du risque et de sa perception », *Variances* (blog), 2 septembre 2020 disponible en ligne : <http://variances.eu/?p=5246>, consulté le 14 septembre 2020.

Charrier T., « Agregio : du post-it à la construction d'un vertical », *Capgemini (blog)*, disponible en ligne : <https://www.capgemini.com/fr-fr/cas-client/agregio-cloud/>, consulté le 1er avril 2022.

Cheminat J., « Le gouvernement réoriente sa stratégie sur le cloud de confiance », *Le Monde de l'Informatique*, 17 mai 2021, disponible en ligne : <https://www.lemondeinformatique.fr/actualites/lire-le-gouvernement-reoriente-sa-strategie-sur-le-cloud-de-confiance-82937.html>, consulté le 16 janvier 2023.

CNIL, « « Bac à sable » données personnelles de la CNIL : appel à projets 2021 », *Site officiel de la CNIL*, 15 février 2021, disponible en ligne : <https://www.cnil.fr/fr/bac-a-sable-2021>, consulté le 5 janvier 2022.

CNIL, « Conformité des systèmes d'IA : les autres guides, outils et bonnes pratiques », *Site de la CNIL (blog)*, 5 avril 2022, disponible en ligne : <https://www.cnil.fr/fr/intelligence-artificielle/guide/conformite-des-systemes-dia-les-autres-guides-outils-et-bonnes-pratiques>, consulté le 18 septembre 2022.

CNIL, « La CNIL propose un nouveau « bac à sable » pour accompagner l'innovation numérique dans le domaine de l'éducation », *Site officiel de la CNIL*, 18 janvier 2022, disponible en ligne : <https://www.cnil.fr/fr/la-cnil-propose-un-nouveau-bac-sable-pour-accompagner-linnovation-numerique-dans-le-domaine-de>, consulté le 22 septembre 2022.

COMMISSION DE L'ÉTHIQUE EN SCIENCE ET EN TECHNOLOGIE (CEST), « Quelle est la différence entre éthique et morale ? », *Site gouvernemental du CEST* (blog), disponible en ligne : <https://www.ethique.gouv.qc.ca/fr/ethique/qu-est-ce-que-l-ethique/quelle-est-la-difference-entre-ethique-et-morale/>, consulté le 29 septembre 2021.

COMMISSION DE RÉGULATION DE L'ÉNERGIE (CRE), « Bac à sable réglementaire : la CRE accorde des dérogations à 9 projets innovants », *Site officiel de la CRE*, 24 mars 2021, disponible en ligne : <https://www.cre.fr/Actualites/bac-a-sable-reglementaire-la-cre-accorde-des-derogations-a-9-projets-innovants>, consulté le 4 janvier 2022.

COMMISSION DE RÉGULATION DE L'ÉNERGIE (CRE), « Dispositif d'expérimentation réglementaire », *Site officiel (blog)*, 29 juillet 2021, disponible en ligne : < https://www.cre.fr/Transition-energetique-et-innovation-technologique/dispositif-d-experimentation-reglementaire>, consulté le 28 avril 2022.

COMMISSION NATIONALE DU DÉBAT PUBLIC (CNDP), « La participation du public et la CNDP », *Site officiel de la CNDP*, 31 mars 2021, disponible en ligne : <https://www.debatpublic.fr/participation-et-environnement-692>, consulté le 29 novembre 2021.

COMMISSION NATIONALE DU DÉBAT PUBLIC (CNDP), « Méthodes et outils », *Site officiel de la CNDP*, 31 mars 2021, disponible en ligne : <https://www.debatpublic.fr/methodes-et-outils-665>, consulté le 30 novembre 2021.

Conrad J., Knight W., « China Is About to Regulate AI—and the World Is Watching », *Forbes (blog)*, 22 février 2022, disponible en ligne : <https://www.wired.com/story/china-regulate-ai-world-watching/>, consulté le 18 juillet 2022.

D'Agrain H., « Du Green IT au Green by IT : exemples d'applications dans les grandes entreprises », *CIGREF*, janvier 2017, disponible en ligne : <https://www.cigref.fr/wp/wp-content/uploads/2017/01/CIGREF-Du-Green-IT-au-Green-by-IT-2017.pdf>, consulté le 11 mars 2020.

De Matharel L., « EDF dégaine Sowee, un hub pour piloter la smart home et une marque IoT », *Le Journal du Net*, 13 octobre 2016, disponible en ligne : < https://www.journaldunet.com/economie/energie/1186520-edf-sowee-station-connectee-smart-home/>, consulté le 3 avril 2022.

EDF, « AI for Humanity : EDF, Thales et Total ouvrent le premier laboratoire industriel commun en Intelligence Artificielle », *Communiqué de presse*, 6 février 2020, disponible en ligne : <https://www.edf.fr/groupe-edf/espaces-dedies/journalistes/tous-les-communiques-de-presse/ai-for-humanity-edf-thales-et-total-ouvrent-le-premier-laboratoire-industriel-commun-en-intelligence-artificielle>, consulté le 7 août 2021.

EDF, « EDF lance Metroscope, la solution d'intelligence artificielle au service de l'excellence opérationnelle de ses clients industriels », Communiqué de presse, 29 mars 2018, disponible en

ligne : <https://www.edf.fr/sites/groupe/files/contrib/groupe-edf/espaces-dedies/espace-medias/cp/2018/cp-20180328-metroscope-vf.pdf>, consulté le 1er avril 2022.

EDF, « Le groupe EDF lance sa plateforme d'open data pour faciliter la compréhension de son action environnementale, industrielle et sociale », *Communiqué de presse*, 7 décembre 2020, disponible en ligne : <https://www.edf.fr/groupe-edf/espaces-dedies/journalistes/tous-les-communiques-de-presse/le-groupe-edf-lance-sa-plateforme-d-open-data-pour-faciliter-la-comprehension-de-son-action-environnementale-industrielle-et-sociale>, consulté le 29 avril 2022.

EDF, « Pour EDF l'IA doit être au service de l'humain », *Le Parisien*, 26 juin 2020, disponible en ligne : <https://www.leparisien.fr/societe/pour-edf-l-ia-doit-etre-au-service-de-l-humain-26-06-2020-8337975.php>, consulté le 3 avril 2022.

EDF, « Thales et Total forment un trio dans l'intelligence artificielle », *Les Echos*, 6 février 2020, disponible en ligne : <https://www.lesechos.fr/industrie-services/energie-environnement/edf-thales-et-total-forment-un-trio-dans-lintelligence-artificielle-1169776>, consulté le 17 juin 2022.

ENDESA, « Artificial intelligence to improve our services », *Site officiel d'Endesa (blog)*, 2019, disponible en ligne : <https://www.endesa.com/en/projects/a201904-artificial-intelligence-improve-services.html>, consulté le 12 juillet 2021.

ENISA, « Cybersecurity Standards and Certification », *Site officiel de l'ENISA*, disponible en ligne : < https://www.enisa.europa.eu/topics/standards?tab=details>, consulté le 12 février 2022.

Evans R., Gao J., « DeepMind AI reduces Google data centre cooling bill by 40% », *DeepMind website (blog)*, 20 juillet 2016, disponible en ligne : <https://www.deepmind.com/blog/deepmind-ai-reduces-google-data-centre-cooling-bill-by-40>, consulté le 12 février 2019.

Gavois S., « Une étude pointe les possibles effets pervers et dangers de l'intelligence artificielle », *Nextimpact (blog)*, 26 février 2018, disponible en ligne : <https://www.nextinpact.com/article/28064/106188-une-etude-pointe-possibles-effets-pervers-et-dangers-intelligence-artificielle>, consulté le 22 août 2021.

Gérot M., Maxwell W., « Le RGPD pourrait freiner les ambitions de l'Europe en matière d'intelligence artificielle », *ecommercemag (blog)*, 30 mars 2020, disponible en ligne : < https://www.ecommercemag.fr/Thematique/management-1225/Breves/Tribune-RGPD-pourrait-freiner-ambitions-Europe-matiere-intelligence-artificielle-348359.htm> consulté le 15 avril 2020.

Goldet B., « Il est urgent de mesurer les impacts énergétiques et environnementaux de nos solutions numériques », *Le monde de l'énergie*, disponible en ligne : <https://www.lemondedelenergie.com/impacts-energetiques-environnementaux-numerique/2019/10/22/>, consulté le 16 juin 2020.

GOUVERNEMENT FRANÇAIS, « COP 21 : Les engagements de la France », *Site officiel du gouvernement*, 1er décembre 2015, disponible en ligne : < https://www.gouvernement.fr/cop21-les-engagements-nationaux-de-la-france-3403>, consulté le 1er mai 2021.

GOVERNMENT OF CANADA, « Algorithmic Impact Assessment », *Site gouvernemental (blog)*, disponible en ligne : <https://open.canada.ca/aia-eia-js/?lang=en.>, consulté le 16 octobre 2019.

Goyal K., « AI Computing Emits CO_2. We Started Measuring How Much », *Medium (blog)*, 30 novembre 2020, disponible en ligne : <https://medium.com/bcggamma/ai-computing-emits-co%E2%82%82-we-started-measuring-how-much-807dec8c35e3>, consulté le 7 mai 2021.

Hern A., « WhatsApp loses millions of users after terms update », *The Guardian*, 24 janvier 2021, disponible en ligne : <https://www.theguardian.com/technology/2021/jan/24/whatsapp-loses-millions-of-users-after-terms-update>, consulté le 25 novembre 2021.

IBERDROLA, « Digital transformation and innovation plans », *Site official d'Iberdrola (blog)*, 2019, disponible en ligne : <https://www.iberdrola.com/about-us/utility-of-the-future/digital-transformation>, consulté le 12 juillet 2021.

IBM, *IBM's Principles for trust and transparency,* 30 mai 2018, disponible en ligne : <https://www.ibm.com/blogs/policy/wp-content/uploads/2018/06/IBM_Principles_SHORT.V4.3.pdf>, consulté le 20 octobre 2019.

IEA, « Greenhouse gas emissions from energy », *Site officiel de l'IEA*, chiffres de 2020, disponible en ligne : <https://www.iea.org/data-and-statistics/data-browser/?country=WORLD&fuel=CO2%20emissions&indicator=CO2BySector>, consulté le 20 janvier 2022.

IEA, *Data and Statistics : CO2 emissions by energy source*, 2018, disponible en ligne : <https://www.iea.org/data-and-statistics/data-browser/?country=WORLD&fuel=CO2%20emissions&indicator=CO2BySector>, consulté le 6 mai 2021.

IEA, *Global EV Outlook 2019*, 2019.

IEA, *Net Zero by 2050 : A Roadmap for the Global Energy Sector*, mai 2021, disponible en ligne : <https://iea.li/nzeroadmap>, consulté le 22 mai 2021.

INSEE, « Enjeux du numérique », *INSEE Références*, édition 2019, Fiche 4.2.

IRSN, « Démarche générale de prévention des accidents – La défense en profondeur », *Site de l'IRSN (blog)*, disponible en ligne : <https://www.irsn.fr/FR/connaissances/Installations_nucleaires/La_surete_Nucleaire/risque-nucleaire/demarche-prevention/Pages/1-defense-profondeur.aspx#.YySu7N8682w>, consulté le 16 septembre 2022.

Jeffrey J., « 8 Companies Using AI to Tackle Climate Change », *Entrepreneur*, 27 Septembre 2019, disponible en ligne : <https://www.entrepreneur.com/article/340002>, consulté le 5 avril 2021.

Kuziemski M., « Un principe de précaution face à l'intelligence artificielle », *Media24* (blog), 2018, disponible en ligne : <ttps://www.medias24.com/chro18268403052018Un-principe-de-precaution-face-a-l-intelligence-artificielle.html>, consulté le 14 avril 2021.

LABORATOIRE NATIONAL DE METROLOGIE ET D'ESSAIS (LNE), « Certification de processus pour l'IA », *Site du LNE (blog)*, disponible en ligne : <https://www.lne.fr/fr/service/certification/certification-processus-ia>, consulté le 18 septembre 2022.

Landa Y., « How Artificial Intelligence Will Incredibly Lower Your Energy Bill », *Medium (blog)*, 29 avril 2019, disponible en ligne : <https://medium.com/datadriveninvestor/how-artificial-intelligence-will-incredibly-lower-your-energy-bill-33914791eala>, consulté le 5 avril 2021.

Levy A., Kolodny L., « Tesla shares drop after report says its Autopilot system was engaged during a fatal crash », *CNBC (blog)*, 17 mai 2019, disponible en ligne : <https://www.cnbc.com/2019/05/17/tesla-shares-fall-on-report-autopilot-system-was-engaged-during-crash.html>, consulté le 25 novembre 2021.

Loiseau G., « Un droit de l'intelligence artificielle à l'échelle européenne en construction : analyse des deux récentes propositions de directives adaptant les règles de la responsabilité civile au développement de l'IA », *Le Club des Juristes (blog)*, 14 décembre 2022, disponible en ligne : <https://blog.leclubdesjuristes.com/un-droit-de-lintelligence-artificielle-a-lechelle-europeenne-en-construction-analyse-des-deux-recentes-propositions-de-directives-adaptant-les-regles-de-la-responsabilite-civile-au/>, consulté le 15 janvier 2023.

McCarthy J., « What is artificial intelligence ? », 24 novembre 2004, disponible en ligne : <https://homes.di.unimi.it/borghese/Teaching/AdvancedIntelligentSystems/Old/IntelligentSystems_2008_2009/Old/IntelligentSystems_2005_2006/Documents/Symbolic/04_McCarthy_whatisai.pdf>, consulté le 22 mai 2021.

Meneceur Y., « Analyse des principaux cadres supranationaux de régulation de l'intelligence artificielle », 31 mai 2021, disponible en ligne : <https://lestempselectriques.net/ANALYSE_IA.pdf>, consulté le 20 janvier 2022.

Meneceur Y., « Proposition de règlement de l'IA de la Commission européenne : entre le trop et le trop peu ? », *Les Temps électriques* (blog), 22 avril 2021, disponible en ligne : <https://lestempselectriques.net/index.php/2021/04/22/proposition-de-reglement-de-lia-de-la-commission-europeenne-entre-le-trop-et-le-trop-peu/#more-1762>, consulté le 28 février 2022.

MINISTÈRE DE LA TRANSITION ÉCOLOGIQUE ET SOLIDAIRE, « Consultation des projets soumis à étude d'impacts », *Site gouvernemental*, disponible en ligne : <https://www.projets-environnement.gouv.fr/pages/home/>, consulté le 4 décembre 2021.

MINISTÈRE DE LA TRANSITION ÉCOLOGIQUE ET SOLIDAIRE, « Le cadre de la participation du public au titre du code de l'environnement », *Site officiel du Ministère de la transition écologique*, 7 février 2019, disponible en ligne : <https://www.ecologie.gouv.fr/cadre-participation-du-public-au-titre-du-code-lenvironnement>, consulté le 29 novembre 2021.

MINISTÈRE DE LA TRANSITION ÉCOLOGIQUE, « Éviter, réduire et compenser les impacts sur l'environnement », *Site gouvernemental*, 14 juin 2021, disponible en ligne : <https://www.ecologie.gouv.fr/eviter-reduire-et-compenser-impacts-sur-lenvironnement>, consulté le 22 août 2021.

MINISTÈRE DE LA TRANSITION ÉCOLOGIQUE, *Cadre général des filières à responsabilité élargie des producteurs*, 6 avril 2021, disponible en ligne : <https://www.ecologie.gouv.fr/cadre-general-des-filieres-responsabilite-elargie-des-producteurs>, consulté le 6 mai 2021.

MINISTÈRE DE LA TRANSITION ÉCOLOGIQUE, *La production d'électricité*, site officiel, 17 février 2017, disponible en ligne : <https://www.ecologie.gouv.fr/production-delectricite>, consulté le 30 mars 2022.

Nichols M., « How Will AI Improve Microgrid Energy Efficiency ? », *Schooled By Science*, 25 avril 2019, disponible en ligne : <https://schooledbyscience.com/how-will-ai-improve-microgrid-energy-efficiency>, consulté le 5 avril 2021.

Nirwan P., « Le secret d'affaires : le droit de propriété intellectuelle caché sous le boisseau », *OMPI Magazine (blog)*, décembre 2017, disponible en ligne : <https://www.wipo.int/wipo_magazine/fr/2017/06/article_0006.html>, consulté le 3 août 2021.

NOYB, « Déclaration sur la décision d'adéquation de la Commission européenne concernant les États-Unis », *NOYB (blog)*, 13 décembre 2022, disponible en ligne : < https://noyb.eu/fr/declaration-sur-la-decision-dadequation-de-la-commission-europeenne-concernant-les-etats-unis>, consulté le 16 janvier 2023.

NOYB, « Le nouveau décret américain a peu de chances de satisfaire à la législation européenne », *NOYB (blog)*, 7 octobre 2022, disponible en ligne : <https://noyb.eu/fr/le-nouveau-decret-americain-peu-de-chances-de-satisfaire-la-legislation-europeenne>, consulté le 16 janvier 2023

Nunez C., « How artificial intelligence could lower nuclear energy costs », *Argonne National Laboratory (blog)*, disponible en ligne : <https://www.anl.gov/article/how-artificial-intelligence-could-lower-nuclear-energy-costs>, consulté le 22 septembre 2022.

OCDE, « Quarante-deux pays adoptent les nouveaux Principes de l'OCDE sur l'intelligence artificielle », *Site institutionnel de l'OCDE*, 22 mai 2019, disponible en ligne : <https://www.oecd.org/fr/presse/quarante-deux-pays-adoptent-les-nouveaux-principes-de-l-ocde-sur-l-intelligence-artificielle.htm>, consulté le 29 avril 2021.

Ochigame R., « How Big Tech Manipulates Academia to Avoid Regulation », *The Intercept* (blog), 20 décembre 2019, disponible en ligne : <https://theintercept.com/2019/12/20/mit-ethical-ai-artificial-intelligence/>, consulté le 23 janvier 2020.

Pehlivan C., Church P., « EU: The ePrivacy Regulation - Let the trilogue begin! », *Linklaters (blog)*, 12 février 2021, disponible en ligne : <https://www.linklaters.com/fr-fr/insights/blogs/digilinks/2021/february/eu---the-eprivacy-regulation---let-the-trilogue-begin>, consulté le 6 juin 2022.

Pierce F., « Energy Hogs: Can World's Huge Data Centers Be Made More Efficient ? », *Yale Environment*, 3 avril 2018, 360, disponible en ligne : <https://e360.yale.edu/features/energy-hogs-can-huge-data-centers-be-made-more-efficient>, consulté le 5 avril 2021.

Piguet V., « Sobriété numérique : les premiers outils juridiques mis en place », *Village de la Justice*, 15 février 2021, disponible en ligne : <https://www.village-justice.com/articles/numerique-pollution,38079.html>, consulté le 9 avril 2021.

Pollet M., « PFUE : la France attendue au tournant sur le volet numérique », *Euractiv*, 14 décembre 2021, disponible en ligne : <https://www.euractiv.fr/section/economie/news/pfue-la-france-attendue-au-tournant-sur-le-volet-numerique/>, consulté le 24 janvier 2022.

Radclyffe C., « The Four 'R's Of Sustainable Tech », *Forbes*, 17 août 2021, disponible en ligne : <https://www.forbes.com/sites/charlesradclyffe/2021/08/17/the-four-rs-of-sustainable-tech/?sh=6811c604243c>, consulté le 22 août 2021.

Rejcek P., « AI Is an Energy-Guzzler. We Need to Re-Think Its Design, and Soon », *Singularity Hub*, 28 février 2020, disponible en ligne : <https://singularityhub.com/2020/02/28/ai-is-an-energy-guzzler-we-need-to-re-think-its-design-and-soon/>, consulté le 7 avril 2021.

Revello S., « Yann LeCun : 'Les machines manquent de bon sens' », *Le Temps*, 4 octobre 2018, disponible en ligne : <https://www.letemps.ch/economie/yann-lecun-machines-manquent-sens>, consulté le 11 mars 2022.

Ritchie H., « Number of People in the World Without Electricity Falls Below One Billion », *Our World Data*, 18 janvier 2019, disponible en ligne : <https://ourworldindata.org/number-of-people-in-the-world-without-electricity-access-falls-below-one-billion>, consulté le 5 avril 2021.

RTE, DREEV, « Pour la première fois en France des véhicules électriques pourront participer à l'équilibrage en temps-réel du système électrique », *Communiqué de presse*, 1er février 2022, disponible en ligne : <https://assets.rte-france.com/prod/public/2022-02/CP_vehicules%20electriques_RTE_Dreev_V2G.pdf>, consulté le 3 avril 2022.

RTE, *Synthèse du bilan électrique 2020*, disponible en ligne : <https://bilan-electrique-2020.rte-france.com/synthese-les-faits-marquants-de-2020/#>, consulté le 6 mai 2021.

Sandler R., « Google Halts AI Tools For Oil Extraction », *Forbes (blog)*, 19 mai 2020, disponible en ligne : <https://www.forbes.com/sites/rachelsandler/2020/05/19/google-halts-ai-tools-for-oil-industry-after-greenpeace-report/>, consulté le 8 avril 2021.

SOWEE, « Avec Sowee & Alexa, votre maison vous obéit au doigt et à la voix ! », *Site officiel (blog)*, 27 juin 2018, disponible en ligne : <https://www.sowee.fr/conseils/autour-de-sowee/avec-sowee-alexa-votre-maison-vous-obeit-au-doigt-et-a-la-voix/>, consulté le 3 avril 2022.

STATISTA RESEARCH DEPARTMENT, *Amount of data created, consumed, and stored 2010-2025, Statista Research Department*, site officiel, 18 mars 2022, disponible en ligne : <https://www.statista.com/statistics/871513/worldwide-data-created/>, consulté le 25 avril 2022.

Swaton E., Neboyan V., Lederman L., « Human factors in the operation of nuclear power plants : Improving the way man and machines work together », *IAEA Bulletin*, 4/1987, disponible en ligne <https://www.iaea.org/sites/default/files/publications/magazines/bulletin/bull29-4/29405042730.pdf>, consulté le 9 septembre 2022.

THALES, « Les innovations de Thales en matière de gestion des clés de chiffrement aident les organisations à atteindre la souveraineté numérique dans les environnements hybrides et

multicloud », *Site de Thalès Group (blog)*, 12 octobre 2022, disponible en ligne : <https://www.thalesgroup.com/fr/monde/groupe/press_release/innovations-thales-matiere-gestion-des-cles-chiffrement-aident>, consulté le 16 janvier 2023.

The Guardian, *The Cambridge analytica files : a year long investigating into Facebook, data, and influencing elections in the digital age*, dossier thématique, 2018, disponible en ligne : <https://www.theguardian.com/news/series/cambridge-analytica-files>, consulté le 10 septembre 2019.

U.S. OFFICE OF SCIENCE, « Department of Energy Announces $5.7 Million for Research on Artificial Intelligence and Machine Learning (AI/ML) for Nuclear Physics Accelerators and Detectors », *Office of Science (blog)*, 2 décembre 2021, disponible en ligne : <https://www.energy.gov/science/articles/department-energy-announces-57-million-research-artificial-intelligence-and>, consulté le 22 septembre 2022.

Victor D., « Microsoft Created a Twitter Bot to Learn From Users. It Quickly Became a Racist Jerk », *New York Times*, 24 mars 2016, disponible en ligne : <https://www.nytimes.com/2016/03/25/technology/microsoft-created-a-twitter-bot-to-learn-from-users-it-quickly-became-a-racist-jerk.html>, consulté le 10 septembre 2019.

Viennot M., « Intelligence artificielle et éthique ne font pas (encore) bon ménage », *France Culture (blog)*, 25 mai 2019, disponible en ligne : <https://www.franceculture.fr/emissions/la-bulle-economique/intelligence-artificielle-et-ethique-ne-font-pas-encore-bon-menage>, consulté le 22 août 2021.

Vitard A., « Dans la perspective de la nouvelle réglementation sur l'IA, la Cnil se prépare à auditer les algorithmes », *L'usine digitale (blog)*, 22 septembre 2022, disponible en ligne : < https://www.usine-digitale.fr/article/dans-la-perspective-de-la-nouvelle-reglementation-sur-l-ia-la-cnil-se-prepare-a-auditer-les-algorithmes.N2046717>, consulté le 24 septembre 2022.

Wakabayashi D., « Self-Driving Uber Car Kills Pedestrian in Arizona, Where Robots Roam », *New York Times*, 19 mars 2018, disponible en ligne : <https://www.nytimes.com/2018/03/19/technology/uber-driverless-fatality.html>, consulté le 10 septembre 2019.

Whittaker M., Dobbe R., « AI and Climate Change: How they're connected, and what we can do about it », *Medium (blog)*, 17 octobre 2019, disponible en ligne : <https://medium.com/@AINowInstitute/ai-and-climate-change-how-theyre-connected-and-what-we-can-do-about-it-6aa8d0f5b32c>, consulté le 7 avril 2021.

Williams J., « Does Al Have the Power to Refine Oil and Gas Efficiency? », *EY* (blog), 4 juin 2019, disponible en ligne : <https://www.ey.com/en us/oil-gas/does-ai-have-the-power-to-refine-oil-and-gas-efficience>, consulté le 8 avril 2021.

Williams M., « Computer problems hit three nuclear plants in Japan », *CNN (blog)*, 3 janvier 2000, disponible en ligne : <https://edition.cnn.com/2000/TECH/computing/01/03/japan.nukes.y2k.idg/index.html>, consulté le 9 septembre 2022.

Zhang M., « Google Photos Tags Two African-Americans As Gorillas Through Facial Recognition Software », *Forbes*, 1er juillet 2015, disponible en ligne : <https://www.forbes.com/sites/mzhang/2015/07/01/google-photos-tags-two-african-americans-as-gorillas-through-facial-recognition-software/>, consulté le 10 septembre 2019.

§VI : Actes juridiques, lois, règlements, directives et conventions internationales

Accord de Paris sur le Climat, adopté le 12 décembre 2015, lors de la 21ème session de la Conférence des Parties à la Convention-cadre des Nations Unies sur les changements climatiques.

Accord interinstitutionnel 2003/C 321/01 du 13 avril 2016 entre le Parlement européen, le Conseil de l'Union européenne et la Commission européenne « Mieux légiférer », publiée au JOUE n°L123/1 le 12 mai 2016.

Algorithmic Accountability Act of 2019, proposition de loi par les sénateurs C. Booker et R. Wyden devant le Congrès américain, 4 octobre 2019, H.R.2231, 116th U.S. Congress (2019-2020).

Arrêté du 10 mars 2017 fixant les règles de sécurité et les modalités de déclaration des systèmes d'information d'importance vitale et des incidents de sécurité relatives au sous-secteur d'activités d'importance vitale « Nucléaire » et pris en application des articles R. 1332-41-1, R. 1332-41-2 et R. 1332-41-10 du code de la défense, publié au JORF n°0065 du 17 mars 2017, texte n° 2

Arrêté du 11 août 2016 fixant les règles de sécurité et les modalités de déclaration des systèmes d'information d'importance vitale et des incidents de sécurité relatives au sous-secteur d'activités d'importance vitale « Approvisionnement en énergie électrique » et pris en application des articles R. 1332-41-1, R. 1332-41-2 et R. 1332-41-10 du code de la défense, publié au JORF n°0197 du 25 août 2016, texte n°4.

Arrêté du 11 janvier 2016 portant homologation de la décision n°2015-DC-0532 de l'Autorité de sûreté nucléaire du 17 novembre 2015 relative au rapport de sûreté des installations nucléaires de base, publié au JORF n°0012 du 15 janvier 2016, texte n° 7.

Arrêté du 2 juin 2006 fixant la liste des secteurs d'activités d'importance vitale et désignant les ministres coordonnateurs desdits secteurs, publié au JORF n°129 du 4 juin 2006, texte n° 1.

Arrêté du 3 novembre 2014, relatif au contrôle interne des entreprises du secteur de la banque, des services de paiement et des services d'investissement soumises au contrôle de l'Autorité de contrôle prudentiel et de résolution, publié au JORF n°0256 du 5 novembre 2015.

Arrêté du 7 février 2012 fixant les règles générales relatives aux installations nucléaires de base, publié au JORF n°0033 du 8 février 2012, texte n°12.

Charte des droits fondamentaux de l'Union européenne, 2000/C 364/01, publiée au JOCE n°C364/3 du 18 décembre.

Circulaire du 10 mai 2010 récapitulant les règles méthodologiques applicables aux études de dangers, à l'appréciation de la démarche de réduction du risque à la source et aux plans de prévention des risques technologiques (PPRT) dans les installations classées en application de la loi du 30 juillet 2003, publié au BO du MEEDDM n° 2010/12 du 10 juillet 2010.

COMMISSION EUROPÉENNE, *Proposition de règlement du Parlement européen et du Conseil établissant des règles harmonisées concernant l'intelligence artificielle (législation sur l'intelligence artificielle) et modifiant certains actes législatifs de l'Union*, 21 avril 2021, 2021/0106 (COD).

COMMISSION EUROPÉENNE, *Proposition de règlement du Parlement européen et du Conseil sur des règles harmonisées relatives à l'accès équitable aux données et à leur utilisation équitable (Data Act)*, 23 février 2022, COM (2022) 68 final.

COMMISSION EUROPÉENNE, *Proposition de règlement n°2021/0240 instituant l'Autorité de lutte contre le blanchiment de capitaux et le financement du terrorisme*, COM (2021) 421 final, 20 juillet 2021.

COMMISSION EUROPÉENNE, *Proposition de règlement relatif à un marché intérieur des services numériques (Digital Services Act) et modifiant la directive 2000/31/CE*, 15 décembre 2020, COM (2020) 825 final.

COMMISSION EUROPÉENNE, *Proposition de règlement relatif aux marchés contestables et équitables dans le secteur numérique (Digital Markets Act)*, 15 décembre 2020, COM (2020) 842 final.

COMMISSION EUROPÉENNE, *Proposition de règlement sur la gouvernance européenne des données (Data Governance Act)*, 25 novembre 2020, COM (2020) 767.

Convention d'Aarhus du 25 juin 1998, ratifiée par la France le 8 juillet 2002 par la Loi n° 2002-285 du 28 février 2002 autorisant l'approbation de la Convention d'Aarhus, publiée au JORF le 1er mars 2002, p. 3904.

Convention sur la sûreté nucléaire, adoptée le 17 juin 1994, INFCIRC/449.

Déclaration de Rio sur l'environnement et le développement adoptée par la conférence des Nations unies sur l'environnement et le développement, Rio de Janeiro, 1992, disponible en ligne : <https://www.un.org/french/events/rio92/rio-fp.htm>, consulté le 30 avril 2021.

Décret n° 2016-1504 du 8 novembre 2016 portant publication de l'accord de Paris adopté le 12 décembre 2015, signé par la France à New York le 22 avril 2016, publié au JORF n°0262 du 10 novembre 2016, texte n°1.

Décret n° 2016-973 du 18 juillet 2016 relatif à la mise à disposition des personnes publiques de données relatives au transport, à la distribution et à la production d'électricité, de gaz naturel et de biométhane, de produits pétroliers et de chaleur et de froid, publié au JORF n°0167 du 20 juillet 2016, texte n° 2.

Décret n° 2020-456 du 21 avril 2020 relatif à la programmation pluriannuelle de l'énergie, publié au JORF n°0099 du 23 avril 2020, texte n°3.

Décret n° 2020-457 du 21 avril 2020 relatif aux budgets carbone nationaux et à la stratégie nationale bas-carbone, publié au JORF n°0099 du 23 avril 2020, texte n°4.

Décret n°2006-212 du 23 février 2006 relatif à la sécurité des activités d'importance vitale, publié au JORF n°47 du 24 février 2006, texte n° 1.

Décret n°2010-1022 du 31 août 2010 relatif aux dispositifs de comptage sur les réseaux publics d'électricité, publié au JORF n°0203 du 2 septembre 2010 ; A*rrêté du 4 janvier 2012 pris en application de l'article 4 du décret n° 2010-1022 du 31 août 2010 relatif aux dispositifs de comptage sur les réseaux publics d'électricité*, publié au JORF n°0008 du 10 janvier 2012.

Directive (UE) 2009/72/CE du Parlement européen et du Conseil du 13 juillet 2009 concernant des règles communes pour le marché intérieur de l'électricité et abrogeant la directive 2003/54/CE, publiée au JOUE n°L211/55 du 14 août 2009.

Directive (UE) 2016/1148 du Parlement européen et du Conseil du 6 juillet 2016 concernant des mesures destinées à assurer un niveau élevé commun de sécurité des réseaux et des systèmes d'information dans l'Union, publiée au JOUE n°L194/1 du 19 juillet 2016.

Directive (UE) 2016/680 du Parlement européen et du Conseil du 27 avril 2016 relative à la protection des personnes physiques à l'égard du traitement des données à caractère personnel par les autorités compétentes à des fins de prévention et de détection des infractions pénales, d'enquêtes et de poursuites en la matière ou d'exécution de sanctions pénales, et à la libre

circulation de ces données, et abrogeant la décision-cadre 2008/977/JAI du Conseil, publiée au JOUE n°L119/89 le 4 mai 2016.

Directive (UE) 2018/2001 du Parlement européen et du Conseil du 11 décembre 2018 relative à la promotion de l'utilisation de l'énergie produite à partir de sources renouvelables, publiée au JOUE n°L328/82 du 21 décembre 2018.

Directive (UE) 2019/770 du Parlement européen et du Conseil du 20 mai 2019 relative à certains aspects concernant les contrats de fourniture de contenus numériques et de services numériques, publiée au JOUE n°L136/1 du 22 mai 2019.

Directive (UE) 2019/771 du Parlement européen et du Conseil du 20 mai 2019 relative à certains aspects concernant les contrats de vente de biens, modifiant le règlement (UE) 2017/2394 et la directive 2009/22/CE et abrogeant la directive 1999/44/CE, publiée au JOUE n°L136/28 du 22 mai 2019.

Directive (UE) 2019/944 du Parlement européen et du Conseil du 5 juin 2019 concernant des règles communes pour le marché intérieur de l'électricité et modifiant la directive 2012/27/UE, publiée au JOUE n°l158/125 du 14 juin 2019.

Directive 2000/43/CE du Conseil du 29 juin 2000 sur l'égalité de traitement entre les personnes sans distinction de race ou d'origine ethnique, publiées au JOUE n°L180 du 19 juillet 2000.

Directive 2000/78/CE du Conseil du 27 novembre 2000 sur l'égalité de traitement en matière d'emploi et de travail, publiée au JOUE n°L303 du 2 décembre 2000.

Directive 2001/95/CE du Parlement européen et du Conseil du 3 décembre 2001 relative à la sécurité générale des produits, publié au JOCE n°L11/4 du 15 janvier 2002.

Directive 2002/95/CE du 27 janvier 2003, relative aux substances dangereuses dans ces équipements ; et *Directive 2002/96/CE, relative aux déchets d'équipements électriques et électroniques*.

Directive 2012/18/UE du Parlement européen et du Conseil du 4 juillet 2012 concernant la maîtrise des dangers liés aux accidents majeurs impliquant des substances dangereuses, modifiant puis abrogeant la directive 96/82/CE du Conseil, publiée au JOUE n°L197 du 24 juillet 2012.

Directive 2014/65/UE du Parlement européen et du Conseil du 15 mai 2014 concernant les marchés d'instruments financiers et modifiant la directive 2002/92/CE et la directive 2011/61/UE, publiée au JOUE n°L173/349 le 12 juin 2014.

Directive 95/46/CE *du Parlement et du Conseil du 24 octobre 1995 relative à la protection des personnes physiques à l'égard du traitement des données à caractère personnel et à la libre circulation de ces données*, publiée au JOUE n°L281 du 23 novembre 1995.

Directive on Automated Decision-Making, 1er avril 2019, disponible en ligne : <https://www.tbs-sct.gc.ca/pol/doc-eng.aspx?id=32592>, consulté le 16 octobre 2019.

GROUPE DE TRAVAIL « ARTICLE 29 » SUR LA PROTECTION DES DONNÉES, *Avis 05/2014 sur les Techniques d'anonymisation*, 0829/14/FR, WP216.

Loi constitutionnelle n°2005-205 du 1er mars 2005 relative à la Charte de l'environnement, publiée au JORF n°0051 du 2 mars 2005, 3697.

Loi n° 2000-108 du 10 février 2000 relative à la modernisation et au développement du service public de l'électricité, publiée au JORF n°35 du 11 février 2000, texte n° 1.

Loi n° 2002-285 du 28 février 2002 autorisant l'approbation de la convention d'Aarhus, publiée au JORF le 1er mars 2002, p. 3904.

Loi n° 2003-699 du 30 juillet 2003 relative à la prévention des risques technologiques et naturels et à la réparation des dommages, publiée au JORF du7 janv. 2004.

Loi n° 2005-1550 du 12 décembre 2005 modifiant diverses dispositions relatives à la défense créant les articles L1332-1 et suivants du Code de la défense, publiée au JORF n°289 du 13 décembre 2005, texte n° 2.

Loi n° 2005-1550 du 12 décembre 2005 modifiant diverses dispositions relatives à la défense créant les articles L1332-1 et suivants du Code de la défense, publiée au JORF n°289 du 13 décembre 2005, texte n° 2.

Loi n° 2006-686 du 13 juin 2006 relative à la transparence et à la sécurité en matière nucléaire, publiée au JORF n°136 du 14 juin 2006, texte n° 2.

Loi n° 2010-1488 du 7 décembre 2010 portant nouvelle organisation du marché de l'électricité, publiée au JORF n°0284 du 8 décembre 2010.

Loi n° 2013-1168 du 18 décembre 2013 relative à la programmation militaire pour les années 2014 à 2019 et portant diverses dispositions concernant la défense et la sécurité nationale, publiée au JORF n°0294 du 19 décembre 2013.

Loi n° 2015-992 du 17 août 2015 relative à la transition énergétique pour la croissance verte, publiée au JORF n°0189 du 18 août 2015, texte n° 1.

Loi n° 2016-1321 du 7 octobre 2016 pour une République numérique, publiée au JORF n°0235 du 8 octobre 2016, texte n° 1.

Loi n° 2018-133 du 26 février 2018 portant diverses dispositions d'adaptation au droit de l'Union européenne dans le domaine de la sécurité, publiée au JORF n°0048 du 27 février 2018, texte n° 2.

Loi n° 2019-1147 du 8 novembre 2019 relative à l'énergie et au climat, publiée au JORF n°0261 du 9 novembre 2019, dite « Loi énergie climat ».

Loi n° 2020-105 du 10 février 2020 relative à la lutte contre le gaspillage et à l'économie circulaire, publiée au JORF n°0035 du 11 février 2020.

Loi n° 2021-1485 du 15 novembre 2021 visant à réduire l'empreinte environnementale du numérique en France, publiée au JORF n°0266 du 16 novembre 2021.

Loi n° 2021-1755 du 23 décembre 2021 visant à renforcer la régulation environnementale du numérique par l'Autorité de régulation des communications électroniques, des postes et de la distribution de la presse, publiée au JORF n°0299 du 24 décembre 2021.

Loi n° 78-17 du 6 janvier 1978 relative à l'informatique, aux fichiers et aux libertés, dit « Loi Informatique et Libertés » ou « LIL », publiée au JORF du 7 janvier 1978.

Loi n°2000-108 du 10 février 2000 relative à la modernisation et au développement du service public de l'électricité, publiée au JORF n°35 du 11 février 2000, texte n°1.

Loi n°2021-1485 du 15 novembre 2021 visant à réduire l'empreinte environnementale du numérique en France, publiée au JORF n°0266 du 16 novembre 2021.

Loi n°88-19 du 5 janvier 1988 relative à la fraude informatique, publiée au JORF du 6 janvier 1988.

Loi n°95-101 du 2 février 1995 relative au renforcement de la protection de l'environnement, publiée au JORF le 3 février 1995.

OCDE, *Recommandation du Conseil de l'OCDE sur l'Intelligence Artificielle,* 22 mai 2019, Recueil des instruments juridiques de l'OCDE, OECD/LEGAL/0449.

Ordonnance n° 2016-1057 du 3 août 2016 relative à l'expérimentation de véhicules à délégation de conduite sur les voies publiques, publiée au JORF n°0181 du 5 août 2016 ; *Loi n° 2019-486 du 22 mai 2019 relative à la croissance et la transformation des entreprises*, publiée au JORF n°0119 du 23 mai 2019.

Règlement (CE) 401/2009 du Parlement et du Conseil du 23 avril 2009 relatif à l'Agence européenne pour l'environnement et au réseau européen d'information et d'observation pour l'environnement, publié au JOUE n°L126/13 du 21 mai 2009.

Règlement (CEE) 1210/90 *du Conseil du 7 mai 1990 relatif à la création de l'agence européenne pour l'environnement et du réseau européen d'information et d'observation pour l'environnement*, publié au JOCE n°L120/33 du 11 mai 1990.

Règlement (UE) 2016/679 *du Parlement européen et du Conseil du 27 avril 2016, relatif à la protection des personnes physiques à l'égard du traitement des données à caractère personnel et à la libre circulation de ces données*, publié au JOUE n°L119/1 du 4 mai 2016.

Règlement (UE) 2018/1807 du Parlement européen et du Conseil du 14 novembre 2018 établissant un cadre applicable au libre flux des données à caractère non personnel dans l'Union européenne, publié au JOUE n°L303/59 le 28 novembre 2018.

Règlement (UE) 2020/852 du Parlement européen et du Conseil, du 18 juin 2020 sur l'établissement d'un cadre visant à favoriser les investissements durables et modifiant le règlement (UE) 2019/2088, publié au JOUE n°L198/13 le 22 juin 2020.

Règlement (UE) n°1321/2014 de la Commission du 26 novembre 2014 relatif au maintien de la navigabilité des aéronefs et des produits, pièces et équipements aéronautiques, et relatif à l'agrément des organismes et des personnels participant à ces tâches, publié au JOUE n°L363/2 du 17 décembre 2014.

Règlement (UE) n°2019/881 du Parlement européen et du Conseil du 17 avril 2019 relatif à l'ENISA (Agence de l'Union européenne pour la cybersécurité) et à la certification de cybersécurité des technologies de l'information et des communications, et abrogeant le Règlement (UE) n°526/201, publié au JOUE n°L 151/15 le 7 juin 2019.

Règlement (UE) n°748/2012 de la Commission du 3 aout 2012 établissant des règles d'application pour la certification de navigabilité et environnementale des aéronefs et produits, pièces et équipements associés, ainsi que pour la certification des organismes de conception et de production, publié au JOUE n°L224/1 du 21 août 2012.

Réponse ministérielle n°15677 sur l'applicabilité aux logiciels de la loi n°98-389 du 19 juin 1998 relative à la responsabilité du fait des produits défectueux, publiée au JO du 24 août 1998, p. 4728.

Résolution 2015/2103(INL) du Parlement européen du 16 février 2017 contenant des recommandations à la Commission concernant des règles de droit civil sur la robotique, publiée au JOUE n°C252/239 le 18 juillet 2018.

Résolution 2020/2012(INL) *du Parlement européen du 20 octobre 2020 portant recommandations à la Commission sur un cadre aux aspects éthiques de l'intelligence artificielle, la robotique et autres technologies*.

Résolution 2020/2014(INL) *du Parlement européen du 20 octobre 2020 portant recommandations à la Commission sur un régime de responsabilité civile pour l'intelligence artificielle*.

Résolution 2020/2015(INL) du Parlement européen du 20 octobre 2020 sur les droits de propriété intellectuelle pour le développement des technologies d'intelligence artificielle.

Résolution 2020/2266(INI) du Parlement européen du 3 mai 2022 sur l'intelligence artificielle à l'ère du numérique.

The City Code on Takeovers and Mergers, 13ème édition, 5 juillet 2021, disponible en ligne : <https://www.thetakeoverpanel.org.uk/the-code/download-code>, consulté le 1er février 2022.

UNESCO, *Avant-projet de Recommandation sur l'éthique de l'intelligence artificielle, Bibliothèque numérique de l'UNESCO*, 7 septembre 2020, disponible en ligne : <https://unesdoc.unesco.org/ark:/48223/pf0000373434_fre>, consulté le 01/04/2021.

UNESCO, *Recommandation sur l'éthique de l'intelligence artificielle*, 23 novembre 2021, SHS/BIO/PI/2021/1, disponible en ligne : <https://unesdoc.unesco.org/ark:/48223/pf0000381137_fre>, consulté le 18 juillet 2022.

Version consolidée du Traité instituant la communauté européenne de l'énergie atomique (Euratom), 2010/C 84/01, publiée au JOUE n°C84, 30 mars 2010.

§VII : Jurisprudences, avis, délibérations, décisions et arrêts

CAA Bordeaux, 26 févr. 2015, n° 13BX00856, Société Notre famille.com : JurisData n° 2015-006245, JCP A 2015, 2239.

CADA, avis 20161147, séance du 09/06/2016, *Communication des tableaux statistiques (de type sondage au 100ème, répartition par centile/décile des clients TRV (tarifs réglementés de vente), base de données des clients TRV « à température normale », etc.), transmis à la Commission de régulation de l'énergie (CRE) depuis le 1er janvier 2010, dans le cadre des discussions relatives aux TRV en niveau et en structure.*

CADA, avis n° 20144578, 8 janvier 2015.

CADA, avis n°20141099, séance du 10/04/2014, *Copie du plan du réseau électrique afin de connaître les points de raccordement au réseau possibles pour le projet de construction de sa cliente sur le territoire de la commune de La Croix-Valmer.*

CADA, avis n°20164200, séance du 03/11/2016, *Copie, par voie électronique, ou par dépôt électronique sur un serveur, des résultats des modèles journaliers de l'outil « APOGEE » pour la période couvrant le 1er janvier 2010 au 15 juillet 2016.*

Cass. Civ. 1ère, 17 janvier 1995, n°93-13.075.

Cass. Civ. 2ème, 10 février 1982, n° 81-40.495, *JCP*, 1983, II, 20069, obs. A. Cœuret.

Cass. Civ. 2ème, 15 mars 2018, nº 16-15.791.

Cass. Civ. 2ème, 5 janvier 1956, GAJC, t. II, 13e éd., 2015, n°205, D. 1957, 261, note R. Rodière, JCP 1956. II. 9095, note R. Savatier.

Cass. Civ. 3ème, 20 octobre 1971, n°70-13.035, n°505, D. 1972, 414, obs. C. Lapoyade-Deschamps.

Cass. Civ. 3ème, 5 décembre 1972, D. 1973, 401, note J. Mazeaud.

Cass. Com., 30 mai 1967, Gaz. Pal., 1967, 2, 79.

Cass. Crim. 10 avr. 1997, n°96-82.183 P, *JCP*, 1997, IV, 1780.

Cass. Crim. 2 avr. 2007, *Dr. pénal*, 2008, 1, obs. L. Jean.

Cass. Crim. 22 juin 1994, n°93-83.900.

Cass. Crim. 29 mai 1996, *Gaz. Pal.*, 1996, 2, 152, obs. J.P. Doucet.

CE, 12 mars 2021, *Doctolib*, ord. réf., n° 450163.

CE, 19 novembre 2020, *Commune de Grande Synthe*, Recueil Lebon, n° 2020:427301.

CE, 2 septembre 2009, n° 318584 ; AJDA, 2009, 1522.

CE, 25 septembre 1998, *Association Greenpeace France*, Rec. CE 1998, p. 343 ; JurisData n° 1998-973428.

CE, 29 juillet 1983, *Docteur Cloarec*, n°32172 ; H. Maisl note ss. CE, 29 juillet 1983, *Recueil Dalloz*, 1985, 49.

CEPD, *Avis 4/2020, 29 juin 2020 sur le livre blanc de la Commission sur l'intelligence artificielle*.

CJCE, 12 janvier 2006, aff. C504/04, Agrarproduktion Staebelow.

CJCE, 26 mai 2005, aff. C132/03, Codacons et Federconsumatori.

CJCE, 9 novembre 2004, *The british Horseracing Board Ltd c/ W. Hill Organization Ltd*, aff. C-203/02, obs. F. Pollaud-Dulian, RTD Com., 2005, p. 90.

CJCE, 9 septembre 2003, aff. C236/01, Monsanto e.a.

CJCE, *Fixtures Marketing Ltd c/ OPAP*, aff. C-444/02, obs. F. Pollaud-Dulian, RTD Com., 2005, p. 90.

CJCE, *Fixtures Marketing Ltd c/ Oy Veikkaus Ltd*, aff. C-46/02, obs. F. Pollaud-Dulian, RTD Com., 2005, p. 90.

CJCE, *Fixtures Marketing Ltd c/ Svenska Spel AB*, aff. C-338/02, obs. F. Pollaud-Dulian, RTD Com., 2005, p. 90.

CJUE 8 juillet 2010, n° C-343/09, Afton Chemicals.

CJUE, 16 juillet 2020, aff. C-311/18, *Data Protection Commissioner c/ Facebook Ireland*, CCE 2020. Comm. 4, obs. N. Metallinos.

CJUE, 21 juin 2022, aff. C-817/19, *Ligue des droits humains*.

CJUE, 28 janvier 2010, aff. C-333/08, Commission c/ France.

CJUE, 29 avril 2010, aff. C-446/08, Solgar Vitamin's France.

CJUE, 8 juillet 2010, aff. C- 343/09, Afton.

CNIL, *Décision MED 2019-035 du 31 décembre 2019*.

CNIL, *Décision n°2016-058 du 30 juin 2016 mettant en demeure la société Microsoft Corporation*.

CNIL, *Délibération du 30 août 2017 prise à l'encontre du Ministère de l'Enseignement supérieur, de la Recherche et de l'Innovation, n°2017-053.*

CNIL, *Délibération n° 2012-404 du 15 novembre 2012 portant recommandation relative au traitement des données de consommation détaillées collectées par les compteurs communicants.*

CNIL, *Délibération n°2017-299 du 30 novembre 2017 portant avis sur un projet de loi d'adaptation au droit de l'Union européenne de la loi n°78-17 du 6 janvier 1978.*

CNIL, *Délibération n°2018-326 du 11 octobre 2018 portant adoption de lignes directrices sur les analyses d'impact relatives à la protection des données (AIPD) prévues par le règlement général sur la protection des données (RGPD),* publiée au JORF n°0256 du 6 novembre 2018, texte n° 81.

CNIL, *Délibération n°2019-118 du 12 septembre 2019 portant adoption de la liste des types d'opérations de traitement pour lesquelles une analyse d'impact relative à la protection des données n'est pas requise*, publiée au JORF n°0246 du 22 octobre 2019, texte n°90.

COMITÉ EUROPÉEN DE LA PROTECTION DES DONNÉES ET CONTRÔLEUR EUROPÉEN DE LA PROTECTION DES DONNÉES, *Avis conjoint 5/2021 sur la proposition de règlement établissant des règles harmonisées sur l'intelligence artificielle*, 18 juin 2021.

COMMISSION DE RÉGULATION DE L'ÉNERGIE (CRE), *Délibération n° 2021-59 du 11 mars 2021 portant décision sur l'octroi des dérogations des dossiers soumis à la CRE dans le cadre du premier guichet du dispositif d'expérimentation réglementaire prévu par la loi relative à l'énergie et au climat*.

Décision 2005/370/CE du Conseil du 17 février 2005 relative à la conclusion, au nom de la Communauté européenne, de la convention sur l'accès à l'information, la participation du public au processus décisionnel et l'accès à la justice en matière d'environnement, publiée au JOUE L124/1 du 17 mai 2005.

Décision d'exécution (UE)2021/915 de la Commission du 4 juin 2021 relative aux clauses contractuelles types entre les responsables du traitement et les sous-traitants au titre de l'article 28, paragraphe 7, du règlement (UE) 2016/679 du Parlement européen et du Conseil et de l'article 29, paragraphe 7, du règlement (UE) 2018/1725 du Parlement européen et du Conseil, publiée au JOUE n° L199/18 du 7 juin 2021.

TA Limoges E20000066/87 EP EOL BEAULIEU, 19 avril 2021, *Conclusions et avis de la commission d'enquête publique, sur la demande d'autorisation unique présentée par la société*

d'exploitation éolienne de Beaulieu en vue d'obtenir l'autorisation d'exploiter un parc éolien de quatre aérogénérateurs et d'un poste de livraison sur le territoire de la commune de Beaulieu (Indre).
TPICE, 30 juin 1999, aff. T-13/99 R.

INDEX ALPHABÉTIQUE

TABLE DES MATIÈRES