Règlement sur l'intelligence artificielle. Premiers éléments d'analyse

Cécile Crichton

Règlement sur l'intelligence artificielle. Premiers éléments d'analyse

Cécile Crichton

Doctorante à l'Université Paris Cité, Institut Droit et Santé

Enseignante à l'Institut Catholique de Paris

1. Calendrier législatif. – Premier texte régulant de manière générale les usages des systèmes d'intelligence artificielle, l'Artificial Intelligence Act (AI Act) ou Règlement sur l'intelligence artificielle (RIA)[1] a été adopté dans un laps de temps conventionnel eu égard à ses enjeux importants. A titre comparatif, le Règlement général sur la protection des données[2] a connu une période de gestation de près de trois ans entre la proposition de la Commission et son adoption définitive. En l'occurrence, la Commission européenne a publié sa proposition initiale le 21 avril 2021[3], qui a fait l'objet d'une orientation générale par le Conseil le 6 décembre 2022[4], suivie d'un premier vote en plénière au Parlement européen le 14 juin 2023[5]. A l'issue d'un nouveau trilogue, le texte définitif a été voté le 13 mars 2024[6], et publié au Journal Officiel le 13 juin 2024.

2. Présentation. – Bien entendu, le Règlement sur l'intelligence artificielle a déjà fait l'objet d'excellentes analyses[7]. Cette présentation se contentera d'apporter quelques commentaires rapides sur ce qui compose l'essentiel du corpus réglementaire, soit le champ d'application organisé en fonction des risques (**Sect. 1re**), le régime de conformité des

[1] Règlement (UE) 2024/1689 du Parlement européen et du Conseil du 13 juin 2024 établissant des règles harmonisées concernant l'intelligence artificielle.

[2] Règl. (UE) 2016/679 du 27 avril 2016.

[3] COM(2021) 206 final ; v. notre présentation *in* Dalloz actualité, 3 mai 2021 et 4 mai 2021.

[4] 15698/22 ; Céline CASTETS-RENARD, Proposition de règlement sur l'intelligence artificielle (derniers développements), D. 2023. 680.

[5] Juliette SENECHAL, Vote des parlementaires européens sur l'AI Act : vers une réglementation accrue des IA, des modèles de fondation et des IA génératives, s'inspirant du DSA, du Data Act et du RGPD ?, Dalloz actualité, 22 juin 2023.

[6] Anticipant le vote, v. Juliette SENECHAL, L'AI Act dans sa version finale – provisoire –, une hydre à trois têtes, Dalloz actualité, 11 mars 2024.

[7] V., par ex., le dossier « Quel futur droit de l'intelligence artificielle ? Analyses choisies de la proposition de règlement de la Commission européenne », *in* Dalloz IP/IT 2022. 66 s. – V., égal., sur l'inscription du texte dans le cadre européen : Chloé PLEDEL, Diane GALBOIS-LEHALLE et Bertrand CASSAR, L'articulation du projet de règlement sur l'intelligence artificielle avec le droit du numérique européen, Dalloz actualité, 17 juin 2023.

systèmes d'intelligence artificielle à haut risque (**Sect. 2**), ainsi que les sanctions encourues (**Sect. 3**).

Section 1re. – L'approche fondée sur les risques

3. Propos liminaires. – Rappelons quelques généralités qui – nul doute – sont à ce jour bien intégrées par le spécialiste. L'AI Act vise à réguler par un champ d'application large les systèmes d'intelligence artificielle en fonction de leurs risques : les applications qui présentent un risque inacceptable sont prohibées (**§ II**) ; celles qui présentent un risque élevé sont soumises à un régime de mise en conformité (**§ III**) ; et celles qui présentent un risque faible prévoient un devoir de transparence (**§ IV**). Cette construction sera naturellement suivie après l'étude de quelques généralités (**§ I**), et avant d'aborder le cas spécifique des modèles d'IA à usage général (**§ V**).

§ I. – Le champ d'application de l'AI Act (art. 1er à 4)

4. Présentation. – Les champs d'application matériel (**A**) et personnel (**B**) seront présentés avant de relever quelques exclusions notables (**C**).

A. – Champ d'application matériel : définition du système d'intelligence artificielle

5. Objectif de souplesse et de sécurité. – Trouver la définition idoine d'un système d'intelligence artificielle relève d'une gageure, ce d'autant plus qu'elle doit viser l'universalité législative, intégrée notamment à la proposition de texte relatif à la responsabilité extracontractuelle de l'IA. Lors de son livre blanc « Intelligence artificielle. Une approche européenne axée sur l'excellence et la confiance » du 19 février 2020[8], la Commission expliquait que « *la définition de l'IA devra être suffisamment souple pour tenir compte des progrès techniques tout en étant suffisamment précise pour garantir la sécurité juridique nécessaire* »[9].

[8] COM(2020) 65 final.

[9] Cette volonté est rappelée au considérant 12 du texte final, y ajoutant l'impératif de correspondre avec l'acception des autres organisations internationales : « *La notion de «système d'IA» figurant dans le présent règlement devrait être clairement définie et devrait être étroitement alignée sur les travaux des organisations internationales œuvrant dans le domaine de l'IA afin de garantir la sécurité juridique, et de faciliter la convergence internationale et une large acceptation, tout en offrant la souplesse nécessaire pour tenir compte des évolutions technologiques rapides dans ce domaine* ».

6. Evolutions. – Initialement, la Commission européenne proposait une définition large et procédait par renvoi à une annexe listant les différentes techniques et approches afférentes aux systèmes d'IA, susceptible d'être modifiée par acte délégué. A grands traits, l'annexe visait les systèmes d'IA utilisant des techniques d'apprentissage automatique (*machine learning*), des systèmes à base de règle (comme les systèmes experts) ou des systèmes statistiques. Relativement critiquable, l'approche a été abandonnée au profit d'une définition centrée sur l'autonomie du système. En effet, adopter une définition fondée sur une liste de techniques – quand bien même peut-elle être complétée par procédure accélérée – inscrit le texte dans une période quasi-figée sans anticiper les évolutions technologiques. De plus, le renvoi aux techniques symboliques et statistiques soulevait la difficulté d'une application, *in fine*, à la plupart des systèmes logiciels, imposant ainsi une lourde mise en conformité à des logiciels utilisés depuis des dizaines d'années et étrangers aux risques couverts par l'AI Act.

7. Définition finale. – L'objectif est-il accompli ? Assurément, en termes de souplesse : l'article 3, 1) du texte dispose effectivement qu'un système d'IA se définit comme « *un système automatisé conçu pour fonctionner à différents niveaux d'autonomie et peut faire preuve d'une capacité d'adaptation après son déploiement, et qui, pour des objectifs explicites ou implicites, déduit, à partir des entrées qu'il reçoit, la manière de générer des sorties telles que des prédictions, du contenu, des recommandations ou des décisions qui peuvent influencer les environnements physiques ou virtuels* ».

8. Difficultés. – En ôtant les exemples, un système d'IA serait ainsi « *un système automatisé qui est conçu pour fonctionner à différents niveaux d'autonomie et peut faire preuve d'une capacité d'adaptation après son déploiement, et qui […] déduit, à partir des entrées qu'il reçoit, la manière de générer des sorties […]* ». Sont alors mises en évidence la capacité de raisonnement, via l'*input* et l'*output*, ainsi que l'autonomie. Hormis ces aspects, rien n'est spécifié, interrogeant dès lors le second objectif poursuivi par la Commission : la sécurité juridique. Seul le considérant 12 éclaire le lecteur sur la définition. Excluant explicitement « *les systèmes fondés sur les règles définies uniquement par les personnes physiques pour exécuter automatiquement des opérations* », le texte se concentre sur la « capacité d'inférence » du système, qui « *va au-delà du traitement de données de base en ce qu'elle permet l'apprentissage, le raisonnement ou la modélisation* ». Dès lors, l'AI Act exclut de nombreux systèmes logiciels qui peuvent être considérés comme de simples exécutants fondés par exemple sur une logique « if… then… else ».

B. – Champ d'application personnel : fournisseur, déployeur, etc.

9. Extraterritorialité. – Suivant la logique d'échappatoire aux lois d'extraterritorialité, l'AI Act s'applique aux systèmes d'IA mis à disposition sur le marché de l'Union, peu important le lieu d'établissement de « l'opérateur », incluant « *un fournisseur, fabricant de produits, déployeur, mandataire, importateur ou distributeur* » (art. 3, 8).

10. Fournisseur. – La majorité des obligations pèse sur le fournisseur, entendu comme la personne qui met sur le marché ou met le système d'IA en service sous son propre nom ou sa propre marque (art. 3, 3), ce qui reste habituel en droit de l'Union[10].

11. Déployeur. – Le cas échéant, les obligations pèsent sur le « déployeur » qui utilise « *sous sa propre autorité un système d'IA* » (art. 3, 4). Initialement, le déployeur était désigné comme l'utilisateur. Cette modification est bienvenue en termes de légistique dans la mesure où la notion d'utilisateur s'apparente à l'utilisateur final, soit bien souvent le consommateur, alors que le règlement exclut explicitement de son champ le déployeur « *utilisant des systèmes d'IA dans le cadre d'une activité strictement personnelle à caractère non professionnel* » (art. 2, § 10).

12. Autres opérateurs. – Enfin, quelques obligations similaires s'appliquent sur les autres opérateurs.

C. – Quelques exclusions notables

13. Exception scientifique. – Deux exceptions – parmi tout un panel – méritent d'être relevées. D'une part, l'AI Act ne s'applique pas aux systèmes « *spécifiquement développés et mis en service uniquement à des fins de recherche et développement scientifiques, ni à leurs sorties* » (art. 2, § 6). Cette exemption prend fin dès lors que le système est mis sur le marché ou mis en service à la suite de l'activité de recherche et de développement (consid. 25). Relevons que l'AI Act s'applique sans préjudice d'autres régimes, comme celui des données à caractère personnel qui, bien que prévoyant un régime simplifié, reste applicable aux solutions d'IA procédant à un traitement de données personnelles.

14. Compétence exclusive des Etats membres. – D'autre part, et naturellement, le règlement ne s'applique pas en matière de sécurité nationale qui ressort de la compétence exclusive de chaque Etat membre, y incluant l'usage exclusivement réservé à des fins

[10] V. par ex. la notion de « producteur » au sens du régime des produits défectueux.

militaires, de défense ou de sécurité nationale (art. 2, § 3). Ainsi, il n'est pas à exclure que certaines pratiques considérées par principe comme interdites puissent être autorisées par l'Etat à ces fins, comme la reconnaissance faciale en temps réel dans l'espace public.

§ II. – Les pratiques interdites (art. 5)

15. Présentation. – La présentation qui suit dresse une liste exhaustive des pratiques inacceptables et de leurs évolutions au fil du processus législatif.

A. – Techniques subliminales (§ 1, a)

16. Intérêt incertain. – Prévue depuis la proposition initiale de la Commission, l'interdiction du recours à des techniques subliminales par IA n'a pas substantiellement été modifiée. Malgré la mention de l'objet et de l'effet, impliquant une prohibition en présence du seul élément d'intentionnalité, il ne s'agit pas d'une interdiction *per se*, car la technique doit causer ou être susceptible de causer un préjudice important. Or, à ce jour, de telles techniques ne semblent pas exister. Relevons également la suppression d'une exception introduite initialement par le Parlement, autorisant le recours à de telles techniques à des fins thérapeutiques approuvées et sur la base du consentement.

B. – Exploitation des vulnérabilités (§ 1, b)

17. Principe. – L'interdiction d'un système d'IA qui exploite les éventuelles vulnérabilités d'une personne ou d'un groupe de personnes ne soulève que peu de difficultés. Les vulnérabilités se limitent aux seuls âge ou handicap, situation sociale ou économique. En juin 2023, le Parlement proposait l'ajout des « *traits de personnalité connue ou prévisible* », ce qui n'a pas été retenu. Comme pour les techniques subliminales, l'exploitation des vulnérabilité est interdite si celle-ci est, *a minima*, « *raisonnablement susceptible de causer un préjudice important* ».

C. – Notation des personnes (§ 1, c)

18. Interdiction de principe. – En réaction au crédit social chinois et aux inquiétudes qu'il suscite, il est naturel que l'AI Act ait considéré la pratique comme inacceptable. Plus encore, l'interdiction a été étendue. Initialement destinée aux seuls « *pouvoirs publics ou pour leurs comptes* », la suppression de cette mention dans le texte final implique une extension de l'interdiction à tout opérateur, qu'il agisse pour le compte d'entités publiques ou privées.

L'interdiction concerne autant la notation d'une personne que d'un groupe de personnes « *au cours d'une période donnée en fonction de leur comportement social ou de caractéristiques personnelles ou de personnalité connues, déduites ou prédites* ». Elle est toutefois restreinte à deux conditions alternatives reproduites ci-dessous. Est ainsi interdit :

- « *le traitement préjudiciable ou défavorable de certaines personnes physiques ou de groupes de personnes dans des contextes sociaux dissociés du contexte dans lequel les données ont été générées ou collectées à l'origine* » et/ou ;
- *« le traitement préjudiciable ou défavorable de certaines personnes ou de groupes de personnes, qui est injustifié ou disproportionné par rapport à leur comportement social ou à la gravité de celui-ci* ».

19. Autorisation conditionnée. – Il résulte de ces éléments que les pratiques de notations des personnes perdureront, sous réserve de ne pas entrer dans les conditions prévues par le Règlement[11]. Par exemple, une société sera autorisée à noter ses clients pourvu qu'elle ne dissocie pas les données de leur contexte ou qu'elle n'emploie pas, en réaction, des moyens injustifiés ou disproportionnés par rapport à ces données.

D. – Police prédictive (§ 1, d)

20. Principe. – Toujours en réaction aux faiblesses des systèmes d'IA de police prédictive, notamment aux discriminations qu'ils génèrent[12], l'AI Act interdit l'utilisation de systèmes « *pour mener des évaluations des risques des personnes physiques visant à évaluer ou à prédire la probabilité qu'une personne physique commette une infraction pénale, uniquement sur la base du profilage d'une personne physique ou de l'évaluation de ses traits de personnalités ou caractéristiques* ». Toute analyse de risque qui ne serait pas fondée sur du profilage reste, *a contrario*, autorisée[13].

[11] V. égal. consid. 31 : « *Cette interdiction ne devrait pas avoir d'incidence sur les évaluations licites des personnes physiques qui sont pratiquées dans un but précis, dans le respect du droit de l'Union et du droit national* ».

[12] V. pour des usages pratiques : Céline CASTETS-RENARD, L'IA en pratique : la police prédictive aux États-Unis, Dalloz IP/IT 2019. 314.

[13] V. consid. 42, qui énumère les exemples suivants : « *les systèmes d'IA utilisant l'analyse des risques pour évaluer la probabilité de fraude financière de la part d'entreprises sur la base de transactions suspectes ou d'outils d'analyse des risques permettant de prédire la probabilité de la localisation de stupéfiants ou de marchandises illicites par les autorités douanières, par exemple sur la base de voies de trafic connues* ».

21. Difficulté. – Si l'AI Act ne s'applique pas au domaine militaire, il n'est pas à exclure que de telles technologies soient tout de même développées pour cette matière spécifique, qui relève de la compétence exclusive des Etats membres.

E. – Reconnaissance faciale par moissonnage de données (§ 1, e)

22. Principe. – Est interdite « *la mise sur le marché, la mise en service à cette fin spécifique ou l'utilisation de systèmes d'IA qui créent ou développent des bases de données de reconnaissance faciale par le moissonnage non ciblé d'images faciales provenant de l'internet ou de la vidéosurveillance* ».

23. Historique. – Création parlementaire, l'interdiction de la reconnaissance faciale par moissonnage de données semble directement faire échos au scandale Clearview AI, par lequel la société procédait à des extractions massives et automatisées d'images sur Internet (*web scraping*) afin de constituer une base de données de reconnaissance faciale. La société commercialisait un moteur de recherche destiné à corréler l'image d'une personne avec les images de la base de données. En France, la CNIL a adressé à Clearview AI une mise en demeure[14], suivie d'une sanction[15] dépourvue de réponse[16]. En Europe, quelques autorités nationales de contrôle en matière de protection des données se sont également saisies de l'affaire[17].

24. Vidéosurveillance. – Cette interdiction appelle à trois remarques. Premièrement, nous nous contenterons de relever que le Parlement étend l'interdiction aux images provenant de la vidéosurveillance.

25. Limitation à la reconnaissance faciale. – Deuxièmement, il est légitime de se demander pourquoi l'AI Act restreint l'interdiction à la reconnaissance faciale. Si le texte a vocation à s'inscrire dans la durée, l'extension à toute reconnaissance biométrique aurait été plus appropriée. Lors de leur avis conjoint sur la proposition de l'AI Act, le Comité européen de la protection des données et le Contrôleur européen de la protection des données exprimaient explicitement le souhait de prohiber toute reconnaissance biométrique ou comportementale, citant à ce titre des exemples qui n'ont pas encore été éprouvés comme la

[14] 26 nov. 2021, délib. n° MED-2021-134, Dalloz IP/IT 2022. 9, nos obs. ; *ibid*. 220, obs. C. Lequesne-Roth.
[15] 17 oct. 2022, délib. n° SAN-2022-019, Dalloz actualité, 9 nov. 2022, nos obs.
[16] 17 avr. 2023, délib. n° SAN-2023-005.
[17] V. par ex., en Italie : GPDP, *Facial recognition: Italian SA fines Clearview AI eur 20 million Bans use of biometric data and monitoring of Italian data subjects*, 9 mars 2022.

reconnaissance de la manière de frapper au clavier[18]. En effet, nul ne sait si, à l'avenir, la démarche d'une personne ou autre technique pourrait ou non l'identifier de manière fiable.

26. Effectivité du droit de la protection des données. – Troisièmement, l'opportunité d'une telle disposition interroge. Théoriquement, le droit de la protection des données personnelles suffirait à couvrir l'interdiction, ne serait-ce qu'en mobilisant le fondement de la base légale. Le moissonnage non ciblé ne permet pas en soi le recueil d'un consentement individuel valable puisqu'il nécessiterait, *a minima* et en ligne, de solliciter l'intermédiation de toutes les plateformes qui mettent les données à destination du public. L'intérêt légitime est contestable du fait de ses conditions plus que restrictives[19]. Quant aux autres bases légales, elles sont inopérantes. Par exemple, il est impossible de se fonder sur la nécessaire exécution du contrat en cas de collecte non ciblée de données en l'absence de relation contractuelle. Dans l'hypothèse de l'existence d'une relation contractuelle, comme la reconnaissance faciale par vidéosurveillance sur le lieu de travail, le traitement ne respecterait pas le principe de minimisation eu égard à l'existence de techniques alternatives moins attentatoires aux droits et libertés des personnes. Dès lors, la disposition de l'AI Act interroge. Cette prohibition ainsi consacrée témoigne-t-elle d'un aveu sur l'ineffectivité du droit de la protection des données ?

F. – Reconnaissance émotionnelle (§ 1, f)

27. D'un risque faible à un risque inacceptable. – Alors qu'auteurs et institutions appellent à la plus grande prudence vis-à-vis du recours aux systèmes d'IA de reconnaissance des émotions[20], l'AI Act a saisi progressivement la mesure d'un tel enjeu. Initialement, la Commission avait classé la technique comme étant à faible risque. Le Conseil s'est ensuite accordé pour la considérer comme à haut risque. Enfin, introduite par le Parlement en juin 2023 et conservée dans la mouture finale, la reconnaissance émotionnelle fait désormais partie des pratiques inacceptables[21].

28. Limites. – Néanmoins, le texte reste timide, limitant la prohibition « *sur le lieu de travail et dans les établissements d'enseignement, sauf lorsque l'utilisation du système d'IA est destinée à être mise en place ou mise sur le marché pour des raisons médicales ou de*

[18] 18 juin 2021, 5/2021, Dalloz actualité, 2 juillet 2021, nos obs.

[19] V. la sanction Clearview AI, préc., qui développe cet aspect. – V., égal., CEPD, Avis 06/2014 sur la notion d'intérêt légitime, 9 avr. 2014, WP217.

[20] V. not. Judith ROCHFELD et Célia ZOLYNSKI, Quelles limites aux traitements des émotions ?, Dalloz IP/IT 2023. 617 ; et, dès la parution de la proposition de l'AI Act, Pierre SIRINELLI et Stéphane PREVOST, Reconnaissance émotionnelle, connaissance irrationnelle ?, Dalloz IP/IT 2021. 237.

[21] V. égal. consid. 44.

sécurité ». Lors du premier vote des députés, en juin 2023, l'interdiction s'appliquait également « *dans les domaines des activités répressives et de la gestion des frontières* » : domaine fréquemment dérogatoire dans le texte final.

G. – Catégorisation biométrique (§ 1, g)

29. Principe. – Introduite par le Parlement en juin 2023, l'interdiction de la catégorisation biométrique exécutée par système d'IA suit naturellement les préoccupations relatives aux biais discriminatoires. Cette interdiction se heurte toutefois à deux limites. D'une part, seule est concernée la catégorisation qui mène à « *des déductions ou des inférences* » concernant les données personnelles sensibles des personnes[22]. D'autre part, sont autorisés « *l'étiquetage ou le filtrage d'ensembles de données biométriques acquis légalement* » ainsi que la « *catégorisation de données biométriques dans le domaine répressif* », ce dernier bénéficiant à nouveau d'une dérogation.

H. – Identification biométrique à distance en temps réel dans des espaces accessibles au public (§ 1, h à § 7)

30. Méfiances. – La reconnaissance biométrique dans l'espace public, en particulier la reconnaissance faciale, incarne l'une des inquiétudes prédominantes en matière d'intelligence artificielle[23]. De nombreux rapports font état des risques relatifs à l'atteinte aux droits et libertés fondamentaux ainsi qu'à un basculement vers une société de surveillance. L'avis conjoint 5/2021 des CEPDs sur l'AI Act, par exemple, y consacre plusieurs paragraphes en préconisant une interdiction totale, autant pour le compte d'acteurs publics ou privés (pts 30 à 34). Par extension, les acteurs impliqués dans le droit de la protection des données appellent à la plus grande prudence sur l'usage de ces technologies[24].

31. Protection des données personnelles. – A ce jour, le droit de la protection des données reste garante principale en la matière. Les juges administratifs ont à plusieurs reprises censuré le déploiement de dispositifs de reconnaissance faciale dans l'espace public. Par exemple, l'utilisation du logiciel Briefcam fait actuellement l'objet d'études par la CNIL et le

---

[22] « *leur race, leurs opinions politiques, leur affiliation à une organisation syndicale, leurs convictions religieuses ou philosophiques, leur vie sexuelle ou leur orientation sexuelle* ».

[23] V., recensant les usages au sein de l'Europe, Caroline LEQUESNE ROTH, Mehdi KIMRI et Pierre LEGROS, Rapport, La Reconnaissance Faciale dans l'espace public - Une cartographie européenne, Université Côte d'Azur, Nice, 2020, hal-0313313.

[24] V. not. Convention 108, Lignes directrices sur la reconnaissance faciale, 28 janv. 2021, T-PD(2020)03rev4, Dalloz IP/IT 2021. 361, obs. C. Lequesne Roth ; CNIL, Reconnaissance faciale : pour un débat à la hauteur des enjeux, 15 nov. 2019, Dalloz actualité, 22 nov. 2019, nos obs.

juge administratif[25]. De même, a été jugée contraire au droit de la protection des données l'utilisation d'un dispositif de reconnaissance faciale à l'entrée de deux lycées[26]. Récemment encore, le Tribunal administratif d'Orléans a censuré une convention conclue entre la ville éponyme et une société, qui avait pour objet d'instaurer un dispositif d'audio-surveillance algorithmique[27] sur le fondement de l'obligation de traitement licite des données personnelles.

32. Prudence du législateur. – En France, les débats se sont particulièrement resserrés sur les dispositifs de reconnaissance faciale via la captation d'images par drones mis en œuvre pour le compte des autorités policières[28]. La loi interdit désormais les « traitements automatisés de reconnaissance faciale »[29], qu'ils soient en temps réel ou différé[30] ; limites qui ont été reproduites au sein de la loi d'expérimentation relative aux Jeux Olympiques[31]. La prudence du législateur français s'explique en partie par anticipation du vote de l'AI Act, lequel était discuté sur la question de la surveillance par les autorités policières et judiciaires dans l'espace public. Initialement, la Commission européenne proposait l'interdiction de « *l'utilisation de systèmes d'identification biométrique à distance en temps réel dans des espaces accessibles au public à des fins répressives* », sauf trois exceptions. Le Parlement a souvent exprimé sa réticence vis-à-vis de ces dispositifs. Par exemple, il s'était opposé à l'utilisation de tels dispositifs à des fins générales et exigeait de solides garanties pour les utilisations ciblées lors de sa résolution 2020/2016(INI) du 6 octobre 2021[32]. Naturellement, le premier vote au Parlement a mené à une suppression totale des exceptions et à une extension de l'interdiction à tout système d'identification en temps réel dans les espaces accessibles au public, en supprimant la mention de l'objectif (« *à des fins répressives* »). Toutefois, le Conseil est revenu sur cette position par la réintroduction des exceptions et de la mention de l'utilisation à des fins répressives.

33. Principe. – La mouture finale a rejoint cette dernière position. Ainsi, l'article 5, § 1, h) interdit désormais « *l'utilisation de systèmes d'identification biométrique à distance en temps*

[25] Ant., en référé, v., TA Caen, 22 nov. 2023, n° 2303004 ; annulé, en l'absence de condition d'urgence, par CE 21 déc. 2023, n° 489990.

[26] Pour la CNIL, v. Jean-Marc PASTOR, La CNIL recadre les projets sécuritaires de deux collectivités, Dalloz actualité, 5 nov. 2019 ; pour le juge administratif, v. TA Marseille, 27 févr. 2020, req. n° 1901249.

[27] TA Orléans, 12 juill. 2024, n0 2104478.

[28] Pour un historique, v. notre résumé : Encadrement législatif de la surveillance par drones, Dalloz IP/IT 2022. 63.

[29] CSI, art. L. 242-4.

[30] Cons. const. 20 janv. 2022, décis. n° 2021-834 DC, pt 30, Dalloz actualité, 27 janv. 2022, obs. E. Maupin.

[31] L. n° 2023-380 du 19 mai 2023, art. 10 ; v. notre analyse : Vidéosurveillance intelligente aux JO : validation sous réserve par le Conseil constitutionnel, Dalloz actualité, 24 mai 2023.

[32] Dalloz IP/IT 2021. 538, nos obs.

réel dans des espaces accessibles au public à des fins répressives, sauf si et dans la mesure où cette utilisation est strictement nécessaire eu égard à l'un des objectifs suivants :

- *i) la recherche ciblée de victimes spécifiques d'enlèvement, de la traite et de l'exploitation sexuelle d'êtres humains, ainsi que la recherche de personnes disparues ;*
- *ii) la prévention d'une menace spécifique, substantielle et imminente pour la vie ou la sécurité physique de personnes physiques ou d'une menace réelle et actuelle ou réelle et prévisible d'attaque terroriste ;*
- *iii) la localisation ou l'identification d'une personne soupçonnée d'avoir commis une infraction pénale, aux fins de mener une enquête pénale, d'engager des poursuites ou d'exécuter une sanction pénale pour des infractions visées à l'annexe II et punissables dans l'État membre concerné d'une peine ou d'une mesure de sûreté privatives de liberté d'une durée maximale d'au moins quatre ans* », l'annexe II visant les infractions suivantes : terrorisme ; traite des êtres humains ; exploitation sexuelle des enfants et pédopornographie ; trafic de stupéfiants ou de substances psychotropes ; trafic d'armes, de munitions ou d'explosifs ; homicide volontaire, coups et blessures graves ; trafic d'organes ou de tissus humains ; trafic de matières nucléaires ou radioactives ; enlèvement, séquestration ou prise d'otages ; crimes relevant de la compétence de la Cour pénale internationale ; détournement d'avion ou de navire ; viol ; criminalité environnementale ; vol organisé ou à main armée ; sabotage ; ou participation à une organisation criminelle impliquée dans une ou plusieurs des infractions énumérées ci-dessus[33].

34. Encadrement des pratiques autorisées. – Les paragraphes qui suivent encadrent l'utilisation sous exception qui vient d'être citée. Dès lors et pour l'heure, le droit français semble conforme à l'AI Act.

35. Transition. – Les pratiques interdites ayant été énumérées, il convient de s'intéresser désormais au second niveau relatif aux pratiques à haut risque.

[33] Ces infractions sont fondées sur les infractions énumérées dans la décision-cadre 2002/584/JAI du Conseil (consid. 33).

§ III. – Les pratiques soumises à une évaluation de conformité (art. 6 à 49)

36. Présentation. – Nous nous contenterons de présenter l'article 6 de l'AI Act relatif aux pratiques à haut risque soumises à une évaluation de conformité ; le régime plus détaillé étant développé ci-après[34].

A. – Pratiques visées à l'annexe I

37. Liste de l'annexe I. – Hormis la nouvelle numérotation (initialement annexe II), l'annexe I reste inchangée. Celle-ci renvoi à une liste de textes européens qui concerne :

- Les machines ;
- Les jouets ;
- Les bateaux de plaisance et les véhicules nautiques à vapeur ;
- Les ascenseurs et leurs composants de sécurité ;
- Les appareils et systèmes de protection destinés à être utilisés en atmosphères explosibles ;
- Les équipements radioélectriques ;
- Les équipements sous pression ;
- Les installations à câble ;
- Les équipements de protection individuelle ;
- Les appareils brûlants des combustibles gazeux ;
- Les dispositifs médicaux et dispositifs de diagnostic *in vitro* ;
- L'aviation civile ;
- Les véhicules à deux ou trois roues et les quadricycles ;
- Les véhicules agricoles et forestiers ;
- Les équipements marins ;
- Le système ferroviaire ;
- Les véhicules à moteur et leurs remorques, ainsi que les systèmes, composants et entités techniques distinctes destinés à ces véhicules, en ce qui concerne leur sécurité générale et la protection des occupants des véhicules et des usagers vulnérables de la route.

[34] V. *infra*, sect. 2.

38. Conditions. – Pour entrer dans la catégorie de systèmes d'IA à haut risque, l'article 6, § 1 dresse deux conditions cumulatives. Le système, qu'il soit indépendant ou non des produits visés ci-dessus, doit, d'une part, constituer lui-même le produit ou être composant de sécurité du produit et, d'autre part, être soumis à une évaluation de conformité avant mise sur le marché ou mise en service. Ainsi, un dispositif d'IA d'imagerie médicale entrerait dans cette catégorie, comme un système d'IA destiné à prévenir la maintenance d'ascenseurs.

B. – Pratiques visées par l'annexe III

39. Liste de l'annexe III. – Plus discuté, l'article 6, § 2 renvoi à une annexe III qui liste les domaines pour lesquels les systèmes d'IA sont considérés comme à haut risque, modifiable par acte délégué de la Commission (art. 7). L'annexe étant longue, nous nous contenterons de présenter le premier point relatif à la biométrie.

40. Biométrie. – Lors de sa proposition de 2021, la Commission proposait de limiter les pratiques à haut risques pour les systèmes « *destinés à être utilisés pour l'identification biométrique à distance en temps réel et* a posteriori *des personnes physiques* ». Rappelons que l'identification en temps réel faisait partie des pratiques interdites sauf exceptions à des fins répressives. Le Parlement a ensuite introduit toute identification biométrique portant sur des personnes physiques ainsi que des caractéristiques dérivées comme la reconnaissance émotionnelle. Finalement, les systèmes d'IA utilisés à des fins de biométrie sont considérés comme à haut risque en présence de l'une des conditions suivantes :

a) ils sont utilisés « *à distance* », sauf si l'unique finalité est de confirmer l'identité de la personne, comme l'utilisation d'une empreinte digitale pour déverrouiller un téléphone, et sauf moissonnage non ciblé d'images pour reconnaissance faciale qui constitue une pratique interdite (art. 5, § 1, e) ;

b) ils sont « *utilisés à des fins de catégorisation biométrique, en fonction d'attributs ou de caractéristiques sensibles ou protégés, sur la base de la déduction de ces attributs ou de ces caractéristiques* » ;

c) ils sont « *utilisés pour la reconnaissance des émotions* », sauf sur le lieu de travail et dans les établissements d'enseignement qui entrent dans les pratiques interdites (art. 5, § 1, f).

41. Conditions. – Pour être considérée comme à haut risque, la pratique doit présenter un « *risque important de préjudice pour la santé, la sécurité ou les droits fondamentaux des*

personnes physiques, y compris en n'ayant pas d'incidence significative sur le résultat de la décision » (art. 6, § 3). En tout état de cause, les pratiques listées à l'annexe III sont toujours considérées comme à haut risque en présence d'un « profilage de personnes physiques » (art. 6, § 3, *in fine*).

42. Transition. – Présentant un risque moindre, les pratiques à risque faible seront désormais étudiées.

§ IV. – Les pratiques soumises à un devoir de transparence (art. 50)

43. Les pratiques considérées par l'AI Act comme étant à faible risque ont été considérablement modifiées depuis la proposition de la Commission. Figurant à l'article 50 du Règlement, elles concernent les agents conversationnels (**A**), les IA génératives (**B**), la reconnaissance émotionnelle (**C**) et les *deep fake* (**D**).

A. – Interaction avec des personnes

44. Principe. – Pour ces premiers systèmes, sont visés les systèmes d'IA destinés à « *interagir directement avec des personnes physiques* ». Sont néanmoins exclues les situations dans lesquelles une personne « *raisonnablement attentive et avisée* » comprend clairement qu'elle interagit avec une machine, ainsi – à nouveau – en cas d'utilisation à des fins de prévention ou de détection des infractions pénales, d'enquêtes ou de poursuites en la matière.

B. – IA générative

45. Principe. – L'AI Act introduit les systèmes d'IA générative de « *contenus de synthèse de type audio, image, vidéo ou texte* », dont la technologie a sensiblement augmenté depuis la publication de la première proposition en avril 2021. Introduits depuis l'orientation générale du Conseil et déjà commenté[35], nous nous contenterons d'apporter les deux remarques suivantes.

46. Manque de cohérence du texte, nature de l'obligation. – D'une part, la disposition outrepasse la seule obligation de transparence – titre de l'article 50 – en obligeant le fournisseur à développer une solution efficace, interopérable, solide et fiable ; interrogeant dès lors le calendrier de l'AI Act et laissant l'impression d'une version textuelle insuffisamment aboutie.

[35] Juliette SENECHAL, L'IA Act déjà obsolète face aux IA de nouvelle génération ?, Dalloz actualité, 1er févr. 2023.

47. Manque de cohérence du texte, débiteur de l'obligation. – D'autre part, l'obligation de transparence consiste à marquer les résultats générés par IA « *dans un format lisible par machine et identifiables comme ayant été générés ou manipulés par une IA* ». En présence d'un contenu textuel généré par IA et repris par une personne pour son propre compte, l'obligation ne s'applique pas puisqu'elle pèse exclusivement sur un fournisseur. Ainsi est-il possible d'imaginer qu'un éditeur de site web puisse publier des contenus exclusivement générés par IA sans que son public n'en soit informé, les dispositions de droit de la consommation relatives en la matière n'étant pas forcément adaptées. En présence d'un contenu visuel, le fournisseur pourrait appliquer des filigranes ou signatures. Le contenu téléversé sur Internet ne cesse s'artificialiser, risquant à terme une forme d'épuisement culturel. Il est ainsi heureux de contraindre les fournisseurs à un marquage du contenu. Toutefois, la disposition reste limitée. Comme pour le contenu textuel, une personne peut tout à fait s'accaparer un contenu qui, sans filigrane suffisamment robuste, pourrait s'apparenter à une contenu produit naturellement. En outre, reste la question des moteurs de recherche ou autres plateformes intermédiaires, qui n'ont pas la maîtrise sur le contenu publié et sur qui il serait impossible d'imposer une obligation générale de surveillance de modération d'un contenu généré par IA[36]. A nouveau, la mouture finale de l'AI Act laisse une impression de précipitation du législateur sur un sujet qui aurait mérité de plus amples réflexions. Les enjeux relatifs à l'IA génératives restent effectivement particuliers, ce qui nécessite des débats hors cadre du Règlement sur l'IA.

C. – Reconnaissance émotionnelle

48. Principe. – Le paragraphe 3 de l'article 50 porte sur les systèmes de reconnaissance émotionnelle ou de catégorisation biométrique lorsque le traitement constitue un traitement de données à caractère personnel. Contrairement aux applications à risque inacceptable et à haut risque, l'obligation de transparence pèse cette fois-ci sur le déployeur, non le fournisseur. En outre et à nouveau, cette obligation ne s'applique pas pour les systèmes dont la loi autorise l'utilisation à des fins de prévention ou de détection des infractions pénales ou d'enquêtes en la matière.

[36] V. les dispositions du règl. « DSA » (UE) 2022/2065 du 19 oct. 2022.

D. – Deep fake

49. Principe. – Enfin, le paragraphe 4 impose une obligation de transparence sur les déployeurs utilisant des systèmes d'IA qui génèrent ou manipulent « *des images ou des contenus audio ou vidéo constituant un hypertrucage* ». Deux exceptions sont prévues :

- A nouveau, si la loi autorise une telle utilisation « *à des fins de prévention ou de détection des infractions pénales, d'enquêtes ou de poursuite en la matière* » ;
- Ou lorsque le *deep fake* « *fait partie d'une œuvre ou d'un programme manifestement artistique, créatif, satirique, fictif ou analogue* ». Dans cette hypothèse, il suffit de divulguer l'existence du *deep fake* d'une manière « *qui n'entrave pas l'affichage ou la jouissance de l'œuvre* ».

50. Information du public sur des questions d'intérêt public. – Un second alinéa a été introduit : « *Les déployeurs d'un système d'IA qui génère ou manipule des textes publiés dans le but d'informer le public sur des questions d'intérêt public indiquent que le texte a été généré ou manipulé par une IA* ». Cette obligation ne s'applique pas :

- A nouveau, si la loi autorise une telle utilisation « *à des fins de prévention ou de détection des infractions pénales, d'enquêtes ou de poursuite en la matière* » ;
- Ou, assurant un contrôle, lorsqu'une personne assume la responsabilité éditoriale du contenu.

51. Transition. – Outre la pyramide des risques, le Règlement sur l'IA introduit un corpus de règles portant sur les modèles d'IA à usage général.

§ V. – Les modèles d'IA à usage général (art. 51 à 56)

52. Définitions. – Défini par l'article 3, 63) de l'AI Act, le modèle d'IA à usage général est « *un modèle d'IA, y compris lorsque ce modèle d'IA est entraîné à l'aide d'un grand nombre de données utilisant l'auto-supervision à grande échelle, qui présente une généralité significative et est capable d'exécuter de manière compétente un large éventail de tâches distinctes, indépendamment de la manière dont le modèle est mis sur le marché, et qui peut être intégré dans une variété de systèmes ou d'applications en aval, à l'exception des modèles d'IA utilisés pour des activités de recherche, de développement ou de prototypage avant leur mise sur le marché* ». Le système d'IA à usage général, fondé sur ce modèle, « *a la capacité*

de répondre à diverses finalités, tant pour une utilisation directe que pour une intégration dans d'autres systèmes d'IA » (art. 3, 66).

53. Principes. – A grands traits, les articles 53 et 54 imposent aux acteurs impliqués dans la chaîne de production de modèles d'IA à usage général de constituer une documentation. En outre, sur le modèle du régime des très grandes plateformes imposé par l'article 34 du règlement « Digital Services Act »[37], les fournisseurs de modèles d'IA à usage général présentant un risque systémique sont tenus d'effectuer une analyse des risques systémiques (art. 55). Hors présomption prévue au paragraphe 2, et selon l'article 51, § 1 de l'AI Act, présente un risque systémique le modèle d'IA à usage général qui :

- « *Dispose de capacités à fort impact évaluées sur la base de méthodologies et d'outils techniques appropriés […] ;*
- *Ou est désigné comme tel par une décision de la Commission* ».

54. Renvoi. – L'annexe XIII de l'AI Act dresse une liste de critères qui doivent être tenus par la Commission pour déterminer si le modèle a des capacités ou un impact équivalent.

55. Transition. – Après avoir défini le champ d'application de l'Artificial Intelligence Act, l'étude du régime applicable aux systèmes d'intelligence artificielle à haut risque mérite une présentation détaillée.

Section 2. – Les règles relatives aux systèmes d'IA à haut risque

56. Champ d'application. – Pour rappel, les systèmes d'IA à risque élevé se déclinent en deux grandes catégories, parmi lesquelles les systèmes qui sont composants de sécurité d'un produit ou eux-mêmes un produit visé par l'annexe I (art. 6, § 1). Cette annexe dresse une liste de réglementations applicables à des produits, comme le règlement (UE) 2017/745 du 5 avril 2017 relatif aux dispositifs médicaux. Si le produit intégrant un système d'IA est soumis à une évaluation de mise en conformité préalable à la mise en service ou sur le marché, il entre dans le régime des systèmes d'IA à risque élevé prévu aux article 8 à 49 de l'AI Act. En outre, sont considérés comme à haut risque les systèmes visés à l'annexe III (art. 6, § 2) sous réserve de présenter un « *risque important de préjudice pour la santé, la sécurité ou les droits fondamentaux des personnes physique, y compris en n'ayant pas d'incidence significative sur*

[37] Règl. (UE) 2022/2065 du 19 oct. 2022.

le résultat de la prise de décision » ou lorsque le système effectue un profilage de personnes physiques (art. 6, § 3).

57. Présentation. – Seront dressées de manière synthétique les nombreuses obligations pesant sur les opérateurs concernées (**§ I**) avant d'étudier la procédure de mise en conformité (**§ II**).

§ I. – Obligations pesant sur les opérateurs des systèmes à haut risque

58. Présentation. – Une grande partie du Règlement étant réservée au fournisseur du système d'IA, ses obligations seront présentées (**A**), avant de se préoccuper du déployeur (**B**) et des autres acteurs (**C**).

A. – Obligations du fournisseur du système d'IA

59. Présentation. – L'AI Act organise essentiellement des obligations pesant sur le fournisseur du système d'IA, entendu comme la personne qui le met sur le marché ou le met en service « *sous son propre nom ou sa propre marque, à titre onéreux ou gratuit* » (art. 3, 3). Redondantes et parfois indigestes, ces obligations seront regroupés en huit étapes à suivre que nous avons arbitrairement choisies en raison de leur cohérence chronologique.

1. – Etablissement d'une documentation

60. Principe. – Afin d'assurer une transparence adéquate, le fournisseur du système d'IA à haut risque est tenu en premier lieu d'établir une documentation complète à destination notamment des autorités de contrôle. L'article 18, § 1 de l'AI Act dresse une liste des documents à conserver dix ans après la mise en circulation du système. Il s'agit d'une étape essentielle qui nous semble devoir être établie avant toute autre obligation.

61. Documentation technique complète. – Le point a) concerne la documentation technique visée à l'article 11 qui relate « *sous une forme claire et intelligible* » (al. 2) les éléments nécessaires à la preuve du respect des exigences liées « *à la présente section* » qui comportent : le système de gestion des risques (art. 9), la gouvernance des données (art. 10), la journalisation (art. 12), la transparence vis-à-vis du déployeur (art. 13), le contrôle humain (art. 14), ainsi que l'exactitude, la robustesse et la cybersécurité (art. 15). La documentation technique doit comprendre au moins les éléments listés à l'annexe IV qui décrivent la plupart des obligations auxquelles est soumis le fournisseur.

62. Documentation technique simplifiée. – Deux situations atténuent les nombreux éléments à préciser. D'une part, les PME bénéficient d'une documentation simplifiée (§ 1, al. 2). D'autre part, il est possible de regrouper en un document unique les documentations des systèmes liés à un produit couvert par l'un des textes énuméré à la section A de l'annexe I et celle des produits eux-mêmes. Ainsi, et par exemple, un dispositif médical embarquant un système d'intelligence artificielle pourrait ne comporter qu'une seule documentation technique qui regrouperait les impératifs du Règlement sur l'intelligence artificielle et ceux portant sur les dispositifs médicaux. Cette recherche de cohérence est rappelée à l'article 8, § 2 de l'AI Act afin « *d'éviter les doubles emplois et de réduire au minimum les charges supplémentaires* ». Relevons que cette intégration des documentations et des procédures est possible pour tous les produits visés à la section A de l'annexe I.

63. Système de gestion de la qualité. – Le point b) de l'article 18 renvoie au système de gestion de la qualité visé à l'article 17. Cette disposition établit une liste d'éléments qui doivent *a minima* apparaître. Parmi ceux-ci figurent des informations qui paraissent évidentes comme les procédures d'examen, de test et de validation effectuées tout au long du développement du système et postérieurement à sa mise en circulation (§ 1, b). Figurent également quelques renvois à d'autres dispositions, comme l'établissement d'un système de gestion des risques visé à l'article 9 (§ 1, g). Il convient de relever que cet élément doit déjà figurer dans la documentation technique de l'article 11, laquelle fait, comme le système de gestion de qualité, également partie des documents à conserver au titre de l'article 18. L'AI Act présente en effet un certain nombre de redondances regrettables provenant sans doute d'une volonté d'adoption précipitée.

64. Autres documents. – Enfin, les points c) à e) visent les différents documents émis et approuvés par les organismes notifiés ainsi que la déclaration de conformité.

65. Journalisation. – Outre des différents documents, le fournisseur du système d'IA doit conserver les journaux générés automatiquement durant une période adaptée (art. 19) ; journalisation qui doit s'intégrer au système d'IA tout au long de sa durée de vie (art. 12, § 1).

2. – Gouvernance des données

66. Appréciation suivant la destination du système. – Imposant un système de contrôle de la qualité des données d'entraînement, de validation et de test, l'article 10 de l'AI Act établit un certain nombre de critères qui entrent en résonnance avec les principes promus par l'article 5 du RGPD. En effet, les dispositions de l'article 10 de l'AI Act s'appliquent

systématiquement suivant la destination des systèmes d'IA, comme il est nécessaire d'apprécier la conformité d'un traitement de données à caractère personnel suivant la finalité déterminée.

67. Bonnes pratiques. – Outre cette évaluation contextuelle, le deuxième paragraphe de l'article 10, § 2 impose *a minima* un certain nombre de bonnes pratiques qui restent conventionnelles en matière d'IA, comme « *les opérations de traitement pertinentes pour la préparation des données, telles que l'annotation, l'étiquetage, le nettoyage, la mise à jour, l'enrichissement et l'agrégation* » (c). Il convient de souligner qu'un dispositif de détection (f) et de correction (g) des biais doit être mis en œuvre. Pour se faire, le cinquième paragraphe organise une exception à l'interdiction du traitement de données à caractères personnelles sensibles.

68. Importance contextuelle. – Par extension, l'article 10 exige un jeu de données suffisamment représentatif (§ 3) en tenant compte d'éléments contextuels liés à la destination du système (§ 4), étant précisé que sont présumés conformes les systèmes qui tiennent compte « *du cadre géographique, comportemental, contextuel ou fonctionnel spécifiques dans lequel ils sont destinés à être utilisés* » (art. 42, § 1).

69. Annexes. – Relevons que l'annexe IV, § 2, d) exige que soit inscrite dans la documentation technique une description détaillée des méthodes et techniques d'entraînement ainsi que les jeux de données utilisées et la manière dont ils ont été obtenus et travaillés. Il s'agit, selon nous, d'un aspect novateur. En effet, l'AI Act ne porte pas seulement sur le système d'intelligence en lui-même, mais aussi sur toutes les données utiles à son développement. En d'autres termes, les règles de mise en conformité ne se limitent pas au logiciel *per se*, ce qui constitue un réel changement de paradigme.

3. – Système de gestion de la qualité, des risques et des incidents

70. Principe. – Suivant la logique de documentation à des fins de pré-constitution de preuve, l'article 17 de l'AI Act impose au fournisseur d'établir une documentation « *de manière méthodique et ordonnée sous la forme de politiques, de procédures et d'instruction écrites* » (§ 1) relative au système de gestion de qualité, et proportionnés à la taille de l'organisation du fournisseur (§ 2). Outre les aspects techniques portant sur le respect du règlement, la documentation inclut le système de gestion des risques (§ 1, g), le système de surveillance après commercialisation (§ 1, h) et les procédures relatives à la notification d'un incident grave (§ 1, i). Ces aspects seront successivement présentés.

71. Système de gestion des risques. – L'article 9 de l'AI Act impose au fournisseur d'établir un système de gestion des risques « *documenté et tenu à jour* » (§ 1) qui se déroule sur « *l'ensemble du cycle de vie [du] système d'IA à haut risque* » (§ 2). Il tient compte des risques « *pour la santé, la sécurité ou les droits fondamentaux* » (§ 2, a). La description détaillée du système de gestion des risques doit être reproduite dans la documentation technique, conformément à l'annexe IV, § 5).

72. Obligation de moyens. – L'article 9 qui vient d'être exposé organise un ensemble d'obligations de moyens. En effet, et comme les dispositions relatives à la gouvernance des données, l'établissement du système de gestion des risques s'apprécie conformément à la destination normale du système (§ 2, a) ou en cas d'une mauvaise utilisation raisonnablement prévisible (§ 2, b). Par extension, la notion de risque se restreint aux seuls risques « *qui peuvent être raisonnablement atténués ou éliminés* » (§ 3). Ainsi le fournisseur peut-il s'exempter d'une éventuelle faute en prétendant l'imprévisibilité du risque survenu ou la mauvaise utilisation de son système. A notre sens, cette mesure fera sans doute l'objet d'une jurisprudence fournie car il est déjà possible d'observer que le public multiplie les tentatives de déstabilisation de systèmes d'IA, notamment génératifs, pour les forcer à dysfonctionner. Par exemple, de nombreux internautes se sont amusés à manipuler le système ChatGPT pour obtenir des propos outranciers ou des conseils dangereux. Dans ces hypothèses, se pose la question de l'usage normal du système et de la potentielle exonération de responsabilité de son fournisseur.

73. Autres mesures. – L'article prévoit ensuite un certain nombre de dispositions relatives aux essais effectués pour détecter les risques (§ 6 à 8) ainsi que les mesures de gestion des risques (§ 4 et 5) qui n'appellent pas de commentaire particulier.

74. Surveillance après commercialisation. – Puisque la gestion des risques doit s'appréhender tout au long de la vie du système, le paragraphe 2, c) de l'article 9 impose aux fournisseurs leur surveillance après commercialisation, détaillée à l'article 72 de l'AI Act. Cette obligation s'exécute en accord avec le déployeur ou toute autre source permettant de recueillir les données pertinentes (art. 72, § 2). Ce plan de surveillance après commercialisation doit être intégré dans la documentation technique (art. 72, § 3 et annexe IV, § 9).

75. Notification. – Enfin, en cas de survenance d'un incident grave, une obligation de notification aux autorités de contrôle est imposée au fournisseur par l'article 73 de l'AI Act.

4. – Supervision humaine

76. Principe. – L'article 14 de l'AI Act impose une forme de garantie humaine, matérialisée par des interfaces permettant un « *contrôle effectif par des personnes physiques pendant leur période d'utilisation* » (§ 1). Les dispositions de l'article sont à lire en parallèle de l'article 9 relatif au système de gestion des risques dans la mesure où le contrôle « *vise à prévenir ou à réduire au minimum les risques pour la santé, la sécurité ou les droits fondamentaux qui peuvent apparaître lorsqu'un système d'IA à haut risque est utilisé conformément à sa destination ou dans des conditions de mauvaise utilisation raisonnablement prévisible* » (§ 2). L'interface doit pouvoir être aisément utilisée par le déployeur (§ 4).

77. Documentation technique. – Il convient de relever que l'évaluation des mesures de contrôle humain doit figurer dans la documentation technique conformément à l'annexe IV § 2, e).

5. – Exactitude, robustesse et sécurité

78. Principe. – Il résulte de l'article 15, § 1 de l'AI Act que « *La conception et le développement des systèmes d'IA à haut risque sont tels qu'ils leur permettent d'atteindre un niveau approprié d'exactitude, de robustesse et de cybersécurité, et de fonctionner de façon constante à cet égard tout au long de leur cycle de vie* ». En d'autres termes, cette obligation de moyens s'envisage *by design*. Le contenu de l'article reste classique, hormis quelques propositions de solutions techniques inhérentes aux systèmes d'IA comme la lutte contre l'empoisonnement de données (*data poisoning*). Les performances du système doivent être reproduites dans la documentation technique (annexe IV).

79. Présomption. – Il résulte des termes de l'article 42, § 2 de l'AI Act que sont présumés conformes aux exigences de cybersécurité énoncées à l'article 15 les systèmes d'IA « *qui ont été certifiés ou pour lesquels une déclaration de conformité a été délivrée dans le cadre d'un schéma de cybersécurité conformément au règlement (UE) 2019/881*[38] *et dont les références ont été publiées au Journal officiel de l'Union européenne* ».

[38] 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.

6. – Transparence

80. Règles de fond. – Il ressort de l'article 13 de l'AI Act deux obligations : l'une relative à l'interprétabilité des algorithmes du système d'IA à haut risque, et l'autre relative à l'information. En effet, le paragraphe 1 impose de manière subtile un seuil minimal d'interprétabilité, faisant échec à toute tentative de commercialisation de systèmes entièrement opaques. Celui-ci dispose que « *La conception et le développement des systèmes d'IA à haut risque sont tels que le fonctionnement de ces systèmes est suffisamment transparent pour permettre aux déployeurs d'interpréter les sorties d'un système et de les utiliser de manière appropriée* ». Par conséquent, est prohibée la mise en circulation d'un système d'IA à haut risque qui serait une boîte noire : les paramètres sur système doivent être compréhensibles.

81. Règles de forme. – L'obligation d'information se manifeste par une notice d'utilisation, laquelle contient au minimum un certain nombre d'informations listées par le paragraphe 3 de l'article.

7. – Mesures correctives

82. Principe. – Dès lors que le fournisseur constate une non-conformité de son système d'IA à haut risque, celui-ci doit appliquer les « *mesures correctives nécessaires pour le mettre en conformité, le retirer, le désactiver ou le rappeler, selon le cas* » (art. 20, § 1). Ces mesures s'accompagnent d'une information adressée aux acteurs concernés.

8. – Coopération avec les autorités compétentes

83. Mise à disposition de la documentation. – La constitution de toute la documentation décrite précédemment est utile pour démontrer la conformité du système d'IA auprès des autorités compétentes qui, conformément à l'article 21 de l'AI Act, sont autorisées à y accéder sur demande motivée.

84. Notifications le cas échéant. – En outre, l'article 20, § 2 impose de prévenir les autorités de tout risque, entendu comme un risque « *pour la santé, la sécurité ou les droits fondamentaux* » (art. 79, § 1), et des éventuelles mesures correctives prises, le cas échéant en collaboration avec le déployeur.

B. – Obligations du déployeur du système d'IA

85. Définition. – Entendu comme une personne « *utilisant sous sa propre autorité un système d'IA* » (art. 3, 4), le déployeur représente généralement l'intermédiaire entre le fournisseur et l'utilisateur final. A ce titre, un certain nombre d'obligations pèse sur lui. Bien entendu, n'est pas considéré comme déployeur la personne qui utilise le système « *dans le cadre d'une activité personnelle à caractère non professionnel* » (*ibid.*).

86. Contrôle des obligations du fournisseur. – Premièrement, pèsent sur le déployeur des obligations de contrôle et de respect des obligations qui incombent au fournisseur : respect de la notice d'utilisation (art. 26, § 1) ; contrôle humain (art. 26, § 2 et 3) ; contrôle des données d'entrées (art. 26, § 4) ; notification des incidents, et le cas échéant, suspension de l'utilisation du système (art. 26, § 5) ; journalisation (art. 26, § 6) ; et coopération avec les autorités compétentes (art. 26, § 12).

87. Dispositions spéciales. – Deuxièmement, quelques dispositions spéciales sont prévues à l'article 26, § 7 à 11, appliquées selon les usages, par exemple, lorsque le déployeur est un employeur qui met en service un système d'IA sur le lieu de travail.

88. Analyse d'impact. – Troisièmement, l'article 27 impose au déployeur d'effectuer une analyse d'impact relative aux droit fondamentaux en présence d'un système d'IA à haut risque visé par l'article 6, § 2, qui renvoi à l'annexe III. Cette analyse peut compléter l'analyse d'impact relative à la protection des données prévue à l'article 35 du RGPD.

89. Requalification. – Il convient de préciser, enfin, que l'article 25 prévoit une requalification du déployeur en fournisseur dès lors qu'il a) commercialise le système sous son propre nom ou sous sa propre marque, b) apporte une modification substantielle au système, ou c) en modifie sa destination.

C. – Obligations pesant sur les autres acteurs

90. Qualifications. – Outre le fournisseur et le déployeur, l'AI Act impose quelques obligations aux autres opérateurs impliqués dans la chaine de production : mandataires, importateurs et distributeurs. Il convient de préciser qu'ils peuvent être requalifiés en fournisseur dans les mêmes conditions que pour le déployeur, selon les termes de l'article 25.

91. Mandataire. – Le mandataire s'entend comme une personne située sur le territoire de l'Union ayant reçu un mandat écrit du fournisseur « *pour s'acquitter en son nom des*

obligations et des procédures établies par [l'AI Act] » (art. 3, 5). En tant qu'interlocuteur, il est tenu de conserver une copie du mandat à destination des autorités compétentes (art. 22, § 3) ; de tenir à leur disposition les coordonnées du fournisseurs, « *une copie de la déclaration UE de conformité, la documentation technique et, le cas échéant, le certificat délivré par l'organisme notifié* » (art. 22, § 3, b) ; et de coopérer avec les autorités (art. 22, § 3, d). Si le mandataire considère ou a des raisons de considérer qu'il existe un manquement commis par le fournisseur, il est tenu de mettre fin au mandat et d'en informer les autorités (art. 23, § 4).

92. Importateur. – L'importateur est une personne située ou établie sur le territoire de l'Union qui met le marché un système d'IA dont le nom ou la marque appartient à un opérateur établi hors du territoire (art. 3, 6). Il est tenu de vérifier la bonne conformité du système (art. 23, § 1) et de vérifier que les conditions de stockage ou de transport ne compromettent pas sa conformité (art. 23, § 4). Le cas échéant, il ne peut mettre le système sur le marché qu'après sa mise en conformité (art. 23, § 2). Comme pour le mandataire, l'importateur est tenu de conserver « *une copie du certificat délivré par l'organisme notifié, selon le cas, de la notice d'utilisation et de la déclaration UE de conformité* » (art. 23, § 5) ; de communiquer tout document utile aux autorités de contrôle (art. 23, § 6) ; et de coopérer avec elles (art. 23, § 7). Enfin, il doit apposer au système son nom ou sa marque (art. 23, § 3).

93. Distributeur. – Le distributeur met un système d'IA à disposition sur le marché de l'Union, sans être qualifié de fournisseur ou importateur (art. 3, 7). Comme tout opérateur, il doit vérifier la conformité du système (art. 24, § 1) et, le cas échéant, ne le mettre sur le marché qu'après mise en conformité (art. 24, § 2) ou de prendre les mesures correctives nécessaires qui peuvent aller jusqu'au rappel (art. 24, § 4). Comme pour l'importateur, il doit s'assurer le cas échéant des bonnes conditions de stockage ou de transport (art. 24, § 3). Enfin, il doit communiquer tout document aux autorités compétentes (art. 24, § 5) et coopérer avec elles (art. 24, § 6).

§ II. – Procédure d'évaluation de la conformité

94. Propos liminaires. – Afin de démontrer la conformité du système d'IA à haut risque, le fournisseur est soumis à une procédure de mise en conformité qui implique les « organismes notifiés » chargés d'évaluer la mise en conformité. Pour rappel, les systèmes d'IA à haut risque renvoient, soit à l'un des produits visés par l'une des réglementations couverte par l'annexe I et qui font déjà l'objet d'une procédure de mise en conformité

spécifique à leur nature (art. 6, § 1), soit à l'une des pratique listée par l'annexe III (art. 6, § 2).

95. Présentation. – Comme pour la plupart des produits couverts par une procédure de mise en conformité, l'opérateur a le choix entre les procédures simplifiées (**A**) et la procédure de droit commun (**B**) pour obtenir sa certification (**C**).

A. – Présomptions de conformité

96. Principe. – Sont présumés conformes les systèmes d'IA à haut risque ayant respecté les normes harmonisées ou les spécifications communes, visées respectivement aux articles 40 et 41 de l'AI Act.

97. Normes harmonisées. – Les normes harmonisées renvoient au règlement (UE) 1025/2012 du 25 octobre 2012 relatif à la normalisation européenne, présentées par la Commission européenne aux organisations européennes de normalisation.

98. Spécifications communes. – Les spécifications communes sont des actes d'exécution adoptés par la Commission en présence d'une carence des normes harmonisées.

B. – Evaluation de la conformité

99. Procédure selon la qualification du système. – Il ressort de l'article 43 de l'AI Act des procédures d'évaluations distinctes selon que le système d'IA à haut risque concerne l'annexe I ou III. Celles-ci seront successivement présentées.

100. Systèmes visés par l'annexe I, principe. – Concernant l'annexe I, et en l'absence d'application ou d'existence de norme harmonisée ou de spécification commune, le fournisseur est tenu de suivre la procédure d'évaluation de la conformité prévue à l'annexe VII, étant précisé qu'il est libre de choisir l'organisme notifié (art. 47, § 1, al. 2 et al. 3). La procédure comporte trois phases principales : un contrôle du système de gestion de la qualité, un contrôle de la documentation technique, ainsi qu'une surveillance du système de gestion de la qualité approuvé.

101. Systèmes visés par l'annexe I, renouvellement. – Une nouvelle procédure d'évaluation de conformité doit être suivie en cas de modification substantielle du système d'IA, étant précisé qu'un apprentissage continu – après mise en circulation – ne constitue pas une modification substantielle (art. 47, § 4).

102. Systèmes visés par l'annexe I, dérogations. – Il convient de préciser que le fournisseur peut, à titre dérogatoire, mettre en circulation le système d'IA durant la demande d'autorisation et pendant une période limitée, d'une part « *pour des raisons exceptionnelles de sécurité publique ou pour assurer la protection de la vie et de la santé humaines, la protection de l'environnement ou la protection d'actifs industriels et d'infrastructures d'importance majeure* » (art. 46, § 1) ou, d'autre part, « *dans une situation d'urgence dûment justifiée pour des raisons exceptionnelles de sécurité publique ou en cas de menace spécifique, substantielle et imminente pour la vie ou la sécurité physique des personnes physiques* » (art. 46, § 2). Cette dérogation est autorisée par l'autorité de surveillance du marché et contrôlée par la Commission. Enfin, pour les systèmes couverts par l'annexe I, section A, « *seules les dérogations à l'évaluation de la conformité établies dans ces actes législatifs d'harmonisation de l'Union s'appliquent* » (art. 46, § 7).

103. Systèmes visés par l'annexe I, déclaration de conformité. – A l'issue de la procédure, le fournisseur établit une déclaration UE de conformité selon les exigences de l'article 47 de l'AI Act ainsi que de son annexe V. Afin d'éviter toute redondance, l'article 8, § 2 autorise l'intégration de la procédure de mise en conformité du système d'IA à la procédure existante le cas échéant. Cette possibilité est rappelée à l'article 47, § 3. Enfin, et selon l'article 18, § 1, e), la déclaration UE de conformité visée à l'article 47 doit être conservée dix ans après la mise sur le marché ou la mise en service du système. Elle doit également figurer dans la documentation technique conformément à l'annexe IV, § 8.

104. Systèmes visés par l'annexe III. – S'agissant des régimes liés aux systèmes d'IA visés par l'annexe III, deux régimes distincts sont prévus par l'article 43. En présence d'un système de biométrie visé au premier point de l'annexe III, le fournisseur doit d'abord appliquer les normes harmonisées ou, le cas échéant, les spécifications communes avant de choisir l'une des procédure d'évaluation de conformité suivante : le contrôle interne visé à l'annexe VI ou la procédure visée à l'annexe VII (art. 43, § 1, al. 1er). Seule une procédure de contrôle interne est exigée pour les systèmes destinés aux infrastructures critiques ; à l'éducation et la formation professionnelle ; à l'emploi, la gestion de la main d'œuvre et l'accès à l'emploi indépendant ; à l'accès et au droit aux services privés essentiels et aux services publics et prestations sociales essentiels ; à la répression ; à la migration, l'asile et la gestion des contrôles aux frontières ; et à l'administration de la justice et aux processus démocratiques (§ 2, renvoyant à l'annexe III, pts 2 à 8).

C. – Certificat, marquage CE et enregistrement

105. Durée. – Les certificats délivrés par les organismes notifiés n'excèdent pas cinq ans pour les systèmes d'IA relevant de l'annexe I et quatre ans pour ceux relevant de l'annexe III (art. 44, § 2). Ils peuvent être suspendus ou retirés (art. 44, § 3). Un marquage CE est apposé « *de façon visible, lisible et indélébile* » sur le système d'IA (art. 48).

106. Enregistrement. – S'agissant des systèmes d'IA visés à l'annexe III, une procédure d'enregistrement doit en outre être effectuée conformément aux dispositions de l'article 49 de l'AI Act, étant précisé que les modalités diffèrent suivant la destination du système.

Section 3. – Sanctions

107. « Effective, proportionnée et dissuasive ». – Les Etats membres sont libres d'organiser le régime des sanctions pourvue qu'elles soient effectives, proportionnées et dissuasives. Si la sanction est une amende administrative, l'article 99, § 7 dresse une liste de critères à prendre en compte pour le calcul de son montant. Un régime calqué sur celui du droit des données à caractère personnel pourrait ainsi être envisagé[39].

108. Systèmes d'IA à risque inacceptable. – Il résulte de l'article 99, § 3 du Règlement que l'usage de systèmes d'IA à risque inacceptable fait l'objet d'une amende administrative d'un montant de 35.000.000 d'euros maximum ou, si, l'auteur est une entreprise, 7 % du chiffre d'affaires annuel mondial, le montant le plus élevé étant retenu.

109. Systèmes d'IA à risque élevé. – Conformément à l'article 99, § 4, les fournisseurs, mandataires, importateurs, distributeurs ou déployeurs risquent une amende d'un maximum de 15.000.000 d'euros ou de 3 % du chiffre d'affaires annuel mondial, le montant le plus élevé étant retenu, en cas de manquement à l'une de leurs obligations. En outre, « *la fourniture d'informations inexactes, incomplètes ou trompeuses aux organismes notifiés ou aux autorités nationales compétente en réponse à une demande fait l'objet d'une amende administrative pouvant aller jusqu'à 7 500 000 EUR ou, si l'auteur de l'infraction est une entreprise, jusqu'à 1 % de son chiffre d'affaires annuel mondial total réalisé au cours de l'exercice précédent, le montant le plus élevé étant retenu* » (art. 99, § 5). Le montant le plus faible est retenu en présence d'une PME (art. 99, § 6).

[39] RGPD, art. 84, § 1 : « *Ces sanctions sont effectives, proportionnées et dissuasives* ». – Egal., art. 83, § 2 sur la liste des critères à prendre en compte en cas d'amende. – Sur les mesures correctrices et les sanctions en France, v. L. n° 78-17 du 6 janv. 1978, art. 20 à 23.

110. Systèmes d'IA à risque faible. – Le manquement aux obligations de transparence par le fournisseur ou le déployeur est soumis à la même sanction que celle concernant les systèmes d'IA à haut risque (art. 99, § 4, g).