



**HAL**  
open science

## Normes techniques et éthique de l'IA

Mélanie Gornet, Winston Maxwell

► **To cite this version:**

Mélanie Gornet, Winston Maxwell. Normes techniques et éthique de l'IA. CNIA 2023 - Conférence Nationale en Intelligence Artificielle, Jul 2023, Strasbourg, France. hal-04121843

**HAL Id: hal-04121843**

**<https://hal.science/hal-04121843>**

Submitted on 8 Jun 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Normes techniques et éthique de l'IA

M. Gornet<sup>1</sup>, W. Maxwell<sup>1</sup>

<sup>1</sup> Télécom Paris, IP Paris - Institut Polytechnique de Paris  
NOS - Numérique, Organisation et Société  
i3 - institut interdisciplinaire de l'innovation

melanie.gornet@telecom-paris.fr

## Résumé

Les progrès technologiques de l'IA engendrent des avancées majeures, mais ont aussi des conséquences sociales importantes qui nécessitent de réguler les pratiques. Le projet de règlement européen sur l'IA, également connu sous le nom d'AI Act, imposera aux fournisseurs de systèmes d'IA à haut risque un nombre d'obligations pouvant être qualifiées d'« éthiques »<sup>1</sup>, notamment le respect de l'équité et des droits fondamentaux. Pour s'assurer de la conformité de ces systèmes aux exigences éthiques, le projet d'AI Act prévoit le recours aux normes harmonisées, soulevant ainsi la question de la compatibilité entre les normes techniques et les enjeux éthiques. Dans cette étude, nous contribuons à ce débat en rappelant le rôle des normes et de la certification en Europe, avant de présenter les différents acteurs de la normalisation travaillant actuellement sur des normes « éthiques » pour l'IA. Nous montrons à travers cet inventaire la diversité de leurs travaux ainsi que la concurrence émergente entre différentes visions de l'éthique de l'IA. Enfin, nous discutons des risques que soulèvent ces normes, tels que la difficulté à définir des critères objectifs et la possibilité de confusion avec une garantie de l'éthique.

## Mots-clés

normes, certification, éthique de l'AI, droit de l'IA, régulation, AI Act

## Abstract

Technological advancements in AI are generating major breakthroughs, but they also have significant social consequences that require regulation. The proposed European regulation on AI, also known as the AI Act, will impose a number of obligations on providers of high-risk AI systems, that can be considered as “ethical” obligations, including respect for fairness and fundamental rights. To ensure compliance of these systems with ethical requirements, the proposed AI Act plans to use harmonized standards, raising the

1. L'adjectif « éthique » est défini dans le dictionnaire Larousse par : « Qui concerne la morale », <https://www.larousse.fr/dictionnaires/francais/%C3%A9thique/31388>; le dictionnaire Le Robert ajoute un second sens : « Qui intègre des critères moraux dans son fonctionnement », <https://dictionnaire.lerobert.com/definition/ethique>.

question of compatibility between technical standards and ethical issues. In this paper, we contribute to this debate by recalling the role of standards and certification in Europe, before presenting the actors currently working on “ethical” AI standards. We show through this inventory the diversity of their work and the competition that is emerging between different visions of AI ethics. Finally, we discuss the risks raised by these standards, such as the difficulty of defining objective criteria and the possibility that citizens may be misled.

## Keywords

standards, certification, AI ethics, AI law, regulation, AI Act

## 1 Introduction

Les systèmes dits d'« intelligence artificielle »<sup>2</sup> (IA), ont été l'objet de nombreuses controverses ces dernières années, renforcées par leur couverture médiatique. Entre autres, l'opacité des systèmes est fortement critiquée car elle serait un obstacle à la compréhension des prises de décision. Les nombreux biais indésirables des systèmes d'IA ont également été pointés du doigt, renforçant les discriminations dans le domaine de la sécurité<sup>3</sup>, de la justice<sup>4</sup>, dans l'accès à l'emploi<sup>5</sup> ou aux aides sociales<sup>6</sup>. De nombreuses initiatives ont ainsi vu le jour, appelant à prendre en considération un certain nombre de principes éthiques lors du cycle de vie des systèmes d'IA. Parmi ces principes, l'équité, l'explicabilité, la transparence ou encore la vie privée s'érigent en préceptes universels [33]. Pourtant, si ces principes font consensus, l'absence d'instructions claires,

2. Le terme « intelligence artificielle » est très controversé, notamment pour son caractère anthropomorphe, c'est-à-dire rappelant des caractéristiques normalement réservés aux êtres humains [56].

3. Par exemple l'arrestation à tort de personnes afro-américaines, mal reconnues par les algorithmes de reconnaissance faciale des forces de l'ordre aux États-Unis [29].

4. Par exemple le logiciel COMPAS utilisé par la justice américaine pour prédire le taux de récidive de criminels, qui semble prédire un risque plus élevé pour les personnes afro-américaines [5].

5. Par exemple le logiciel de tri des CVs d'Amazon qui rejetait plus facilement les candidates femmes [53].

6. Par exemple l'algorithme de détection de fraudes aux allocations familiales de l'autorité fiscale néerlandaise, signalant davantage les personnes issues de l'immigration [55].

concrètes et opérationnelles sur la manière d’atteindre réellement ces objectifs rend leur adoption complexe.

En avril 2021, la Commission Européenne a proposé un projet de règlement établissant des règles harmonisées en matière d’intelligence artificielle, nommé *AI Act*, et inspiré pour partie de ses travaux préliminaires sur l’éthique de l’IA. Le texte prévoit le recours à des normes harmonisées pour respecter les exigences essentielles [30]. Alors que ces normes ont jusque-là été utilisées pour codifier des critères techniques, elles vont, dans le cadre de l’IA, devoir adresser des problèmes éthiques. Mais est-il possible de réguler des questions éthiques par des normes techniques ?

La question des normes pour traiter de l’éthique de l’IA est quasiment absente de la littérature. En 2017, [7] examinait les premières initiatives de normes éthiques, en particulier les initiatives de l’IEEE, mais ces initiatives se sont depuis démultipliées. D’autres travaux ont recensé les projets de normes pour l’IA, sans différencier les normes traitant des aspects éthiques [22, 64]<sup>7</sup>.

Dans cette étude, nous proposons d’examiner les différentes normes en préparation en matière d’éthique de l’IA, les acteurs qui les développent, et les dynamiques qui se dégagent. Cet inventaire permettra d’apporter des premières réponses sur le rôle des normes et de la certification en matière d’éthique de l’IA.

Nous commençons par rappeler dans la Section 2 le fonctionnement des procédures de normalisation et de certification, ainsi que la manière dont l’*AI Act* introduit le recours à des normes dites « éthiques ». Nous dressons dans la Section 3 une vue d’ensemble des activités des différents acteurs de la normalisation et de leurs visions de l’éthique, mettant en lumière leurs désaccords et leur rivalité. Nous discutons dans la Section 4 des problèmes posés par ces normes, notamment la difficulté de définir ce qu’est, ou n’est pas une norme « éthique », la subjectivité de leur contenu, et la certification vis à vis de ces normes qui s’érige insidieusement comme une garantie de l’éthique. Enfin, en conclusion, nous dégageons quelques pistes de réflexion pour contribuer au débat plus large sur la coexistence entre normes techniques et débats éthiques.

## 2 Rôle des normes techniques et de la certification

### 2.1 Normes techniques et organismes de normalisation

Les normes sont des documents techniques destinés à établir des solutions communes à des exigences données<sup>8</sup> [10]. Elles permettent notamment de « *définir un langage commun entre les acteurs, de clarifier, d’harmoniser les pratiques et de définir le niveau de qualité [...] des produits [et des] services* » [2]. Leur application, que ce soit dans le secteur privé ou public, est volontaire : il n’est jamais exigé légalement d’appliquer une norme, mais

cela peut faciliter la certification<sup>9</sup>. Toutefois, l’accès aux normes est souvent payant<sup>10</sup>.

Les normes peuvent remplir plusieurs rôles : elles aident les entreprises à optimiser les coûts et à augmenter leur efficacité [19], elles contribuent à dynamiser l’économie [25], à stimuler l’innovation [4], à encourager la concurrence entre les entreprises [6], et à assurer la protection des consommateurs [52].

Les plus connues sont les normes ISO, du nom de l’organisme les développant : l’Organisation internationale de normalisation, ou *International Organization for Standardization* en anglais. Les normes ISO sont couramment utilisées pour certifier la qualité des produits ou des services d’une entreprise<sup>11</sup>, sa performance environnementale<sup>12</sup>, ou encore la sécurité de ses systèmes informatiques<sup>13</sup>. Dans le cadre des systèmes numériques, l’ISO collabore souvent avec la Commission électrotechnique internationale (IEC)<sup>14</sup>. Enfin, une troisième organisation, nommé l’Union internationale des télécommunications (ITU)<sup>15</sup> qui est une agence des Nations Unies, participe également à établir des normes pour les technologies de l’information et de la communication. L’ISO, l’IEC et l’ITU sont les trois agences de normalisation principales participant à l’élaboration de normes pour l’IA. Mais d’autres organismes de normalisation du monde du numérique comme le World Wide Web Consortium (W3C)<sup>16</sup>, lancent également des groupes de travail sur l’IA<sup>17</sup>.

Au niveau européen, trois organismes se chargent du développement de normes : le Comité européen de normalisation (CEN), le Comité européen de normalisation électrotechnique (CENELEC)<sup>18</sup> et l’Institut européen des normes de télécommunication (ETSI)<sup>19</sup>. Ils sont parfois nommés ESOs, pour *European Standardisation Organisations*<sup>20</sup>.

Les normes destinées à soutenir la législation européenne sont publiées au Journal Officiel de l’Union Européenne<sup>21</sup>, elles sont alors appelées « normes harmonisées »<sup>22</sup>. Ces

9. Voir Section 2.2

10. C’est le cas notamment des normes ISO.

11. Par exemple, la norme ISO 9001 :2015 - Quality management systems et ses normes connexes : <https://www.iso.org/iso-9001-quality-management.html>

12. Par exemple, la norme ISO 14001 :2015 - Environmental management systems et ses normes connexes : <https://www.iso.org/iso-14001-environmental-management.html>

13. Par exemple, la norme ISO 27001 :2022 - Information security, cybersecurity and privacy protection et ses normes connexes : <https://www.iso.org/isoiec-27001-information-security.html>

14. <https://iec.ch/homepage>

15. <https://www.itu.int/en/Pages/default.aspx>

16. <https://www.w3.org/>

17. <https://www.w3.org/blog/2021/04/w3c-launches-the-web-machine-learning-working-group/>

18. CEN et CENELEC travaillent régulièrement ensemble pour le développement des normes, <https://www.cencenelec.eu/>

19. <https://www.etsi.org/>

20. Voir le site de la Commission Européenne recensant les acteurs clés de la normalisation : [https://single-market-economy.ec.europa.eu/single-market/european-standards/key-players-european-standardisation\\_en](https://single-market-economy.ec.europa.eu/single-market/european-standards/key-players-european-standardisation_en)

21. <https://eur-lex.europa.eu/oj/direct-access.html?locale=fr>

22. Voir le site de la Commission Européenne sur les normes harmonisées : <https://single-market-economy.ec.europa.eu/si>

7. Voir également l’initiative du AI Standards Hub : <https://aistandardshub.org/ai-standards-search/>

8. Traduction des auteurs.

normes harmonisées peuvent être des normes européennes développées par la CEN, la CENELEC ou l'ETSI, mais peuvent également être des normes internationales adoptées par les ESOs après vote des membres. Une fois adoptées au niveau européen, ces normes peuvent alors être adoptées, par ricochet, par les organismes nationaux.

En effet, chacun des pays membres de l'Union possède son propre organisme de normalisation. Par exemple, en France il s'agit de l'Association Française de Normalisation (AFNOR) et en Allemagne du *Deutsches Institut für Normung* (DIN). Ces organismes nationaux sont parfois nommés NSBs, pour *National standardisation bodies*<sup>23</sup>. Ils sont constitués d'experts réunis en différentes commissions, comme la Commission de Normalisation sur l'IA (CN IA) de l'AFNOR, qui votent pour l'adoption de normes, le lancement de projets de normes, et participent à la création de normes dans les instances européennes, comme la CEN-CENELEC, ou internationales, comme l'ISO. Ces experts sont souvent des acteurs industriels, représentant les intérêts de leurs entreprises d'origines, ces dernières ayant passé un contrat de prestation avec un NSB. Mais les experts peuvent également être issus d'instituts de recherche ou d'établissements publics par exemple. Tout le monde peut ainsi demander à rejoindre un NSB pour participer à l'élaboration de normes et aux votes des CN, en échange de frais d'adhésion.

## 2.2 Processus de certification

Les processus de certification interviennent indépendamment de la création de normes : ni l'ISO, ni les ESOs, ni les NSBs ne se chargent de les mener. Une fois une norme publiée, un organisme indépendant peut proposer une assurance écrite, appelée « certificat », attestant que le produit ou service répond aux exigences définies dans la norme<sup>24</sup>. Ces organismes de certification peuvent être des organismes officiels<sup>25</sup> ou non.

En Europe, il existe un type particulier de certification appelé « marquage CE », pour « Conformité Européenne ». Pour certains groupes de produits, cette marque est obligatoire pour pouvoir entrer sur le marché européen. C'est le cas, par exemple, des produits électriques et électroniques, des jouets, des dispositifs médicaux, des machines, etc<sup>27</sup>. Elle indique la conformité aux exigences essentielles définies dans les directives européennes et est appliquée directement par le constructeur. En effet, le fabricant est le seul responsable de la déclaration de conformité. Il peut choisir d'évaluer lui-même son produit, ou de faire appel à un orga-

---

[single-market/european-standards/harmonised-standards\\_en](https://single-market/european-standards/harmonised-standards_en)

23. Voir supra note 20

24. Voir le site de l'ISO sur la certification : <https://www.iso.org/certification.html>

25. Les organismes de certification peuvent être accrédités. L'accréditation consiste à recevoir une reconnaissance officielle par un organisme indépendant, appelé organisme d'accréditation, que l'organisme de certification fonctionne bien conformément aux normes internationales. Toutefois, l'accréditation n'est pas obligatoire pour devenir organisme de certification<sup>26</sup>.

27. Une exception notable est le cas des véhicules tels que les voitures, les bus, les camions ou les tracteurs qui suivent une procédure spécifique.

nisme notifié, i.e. une organisation désignée par un pays de l'UE pour évaluer la conformité de certains produits avant leur mise sur le marché<sup>28</sup>. Le recours à un organisme notifié est même obligatoire pour certaines catégories de produits. Dans le cas d'un recours à un organisme notifié, le fabricant doit constituer un dossier technique documentant la conformité, puis signer la déclaration de conformité européenne. Il doit être en mesure de fournir les documents justificatifs du marquage CE sur demande de l'autorité nationale compétente<sup>29</sup>. Toutefois, le marquage ne signifie pas que l'UE a approuvé un produit comme étant sûr ou conforme [15], son apposition sur le produit n'engage que le fabricant.

Bien que les normes soient volontaires, en Europe les produits fabriqués dans le respect des normes harmonisées bénéficient d'une présomption de conformité vis-à-vis de la législation. Les fabricants peuvent alors bénéficier de procédures d'évaluation de la conformité simplifiées. S'ils choisissent de ne pas appliquer les normes harmonisées, ils doivent démontrer par d'autres moyens que le produit satisfait aux exigences essentielles définies dans la directive ou dans le règlement européen correspondant [28].

En plus de ces certifications encadrées par la loi, certaines entreprises ou associations développent elles-mêmes leurs propres critères pour juger de la qualité d'un produit. Ces marques délivrées par des organismes privés sont appelées « labels » et ont l'avantage d'être plus rapides à mettre en place que la certification aux normes. Ils sont notamment très utilisés pour le respect de critères environnementaux<sup>30</sup>. Ainsi, la certification publique, basée sur la mise en conformité à partir normes harmonisées, est destinée à faciliter le commerce tout en garantissant le respect des exigences réglementaires. La certification privée quant à elle, qu'elle soit basée sur des normes européennes non-harmonisées, des normes internationales ou sur des cadres d'évaluation privés, est davantage présentée comme un repère pour le consommateur, qui sert à démontrer la qualité supérieure d'un produit et à améliorer son potentiel commercial.

## 2.3 AI Act et normes « éthiques »

L'*AI Act* [23] est aujourd'hui encore en discussion au sein des instances européennes, le texte n'est donc pas définitif et est amené à évoluer. Nous nous baserons ici sur la proposition de la Commission européenne datant d'avril 2021. Dans la proposition de règlement, les systèmes d'IA sont classés en plusieurs catégories, selon leur niveau de risques : risque minimal (art. 69), risque faible (art. 52), haut risque (art. 6 et suivants) et risque inacceptable (art. 5). Ces derniers seront interdits d'utilisation, tandis que les sys-

---

28. Voir le site de la Commission Européenne sur les organismes notifiés : [https://single-market-economy.ec.europa.eu/single-market/goods/building-blocks/notified-bodies\\_en](https://single-market-economy.ec.europa.eu/single-market/goods/building-blocks/notified-bodies_en)

29. Voir le site de la Commission Européenne sur le marquage CE : [https://europa.eu/youreurope/business/product-requirements/labels-markings/ce-marking/index\\_en.htm](https://europa.eu/youreurope/business/product-requirements/labels-markings/ce-marking/index_en.htm)

30. Voir le guide du ministère de la transition écologique sur les labels environnementaux : <https://www.ecologie.gouv.fr/label-s-environnementaux>

tèmes à risque minimal ou faible seront autorisés respectivement sans restriction, et simplement avec des obligations d'information et de transparence. Les systèmes soumis à une évaluation de conformité seront seulement les systèmes jugés à haut risque, qui devront être certifiés CE.

La Commission insiste sur sa volonté d'intégrer des considérations « éthiques » pour l'encadrement des systèmes d'IA. Ainsi, selon le considérant 5<sup>31</sup>, le règlement contribue à « faire de l'Union un acteur mondial de premier plan dans le développement d'une intelligence artificielle sûre, fiable et éthique, et il garantit la protection de principes éthiques ». Selon l'exposé des motifs, le règlement a pour but de renforcer « la contribution de l'Union à la définition de normes mondiales et à la promotion d'une IA digne de confiance qui soit conforme aux valeurs et aux intérêts de l'Union ». Il est également précisé que les exigences minimales proposées s'inspirent des lignes directrices en matière d'éthique du Groupe d'Experts de Haut Niveau sur l'IA de la Commission européenne (GEHN)<sup>32</sup> [49].

Ces considérations éthiques se reflètent dans l'attention particulière portée aux violations des droits individuels. Ainsi un système peut être considéré à haut risque s'il présente « un risque de préjudice pour la santé et la sécurité, ou un risque d'incidence négative sur les droits fondamentaux » (art. 7).

Pour réguler l'IA, la Commission a fait le choix du modèle de la conformité [9] : les systèmes d'IA devront respecter les critères définis par la réglementation avant de pouvoir être distribués sur le marché européen [30]. Ce choix est également celui de nombreux textes en droit du numérique [9]. La normalisation devrait alors « jouer un rôle essentiel [...] afin de garantir la conformité » (considérant 61). Cependant, les exigences minimales étant inspirées de critères éthiques, les normes conçues pour les évaluer touchent par là-même à des aspects éthiques.

Les normes dites « éthiques » se sont fortement développées ces dernières années. La première norme à traiter explicitement d'éthique en robotique date de 2016, elle a été suivie depuis par d'autres initiatives comme celle de l'IEEE [7]. Aujourd'hui, de nombreux acteurs développent des normes touchant de près ou de loin à des aspects éthiques. Nous étudions cet écosystème dans la Section 3.

Néanmoins, ces normes restent marginales. Par exemple, le Règlement n°1025/2012 relatif à la normalisation européenne liste les éléments pouvant être considérés comme des spécifications techniques (art. 2.4.a) [1], comme la protection de l'environnement, la santé, ou encore la sécurité, mais n'inclut pas les critères éthiques.

31. Selon le dictionnaire Larousse, un considérant est un alinéa dans un arrêt d'une cour ou dans une décision de juridiction administrative qui motive la décision. <https://www.larousse.fr/dictionnaires/francais/consid%C3%A9rant/18384>

32. Plus connu sous le nom de *High-Level Expert Group on Artificial Intelligence*, ou *HLEG*, en anglais.

## 3 Une multitude d'acteurs et d'initiatives en matière de normes éthiques pour l'IA

### 3.1 Une ruée vers les normes pour accompagner les progrès techniques

Il existe un décalage entre la vitesse de déploiement des produits et services basés sur l'IA et le développement des normes. Cela crée une pression sur les organismes de normalisation afin de publier des normes le plus rapidement possible. À cela s'ajoute les contraintes dues au calendrier de l'*AI Act* et de sa mise en application.

L'ISO est en avance dans la création de normes techniques pour l'IA : 17 standards publiés et 24 en développement aujourd'hui<sup>33</sup>, en comptant simplement les normes transverses<sup>34</sup>. Parmi ces normes pour l'IA, de nombreuses initiatives adressent des questions éthiques larges comme les impacts sociaux de la technologie<sup>35</sup>, tentent de définir ce qu'est une IA « de confiance »<sup>36</sup>, ou adaptent le management de risques à l'IA<sup>37</sup>. D'autres normes choisissent une approche plus pratique, définissant des méthodes de conception pour éviter le traitement des biais<sup>38</sup>, ou des critères de qualité des données<sup>39</sup>.

Au contraire, les européens CEN et CENELEC démarrent à peine leur campagne de création de normes relative à l'*AI Act*. À ce jour, aucune norme n'a encore été publiée dans ce cadre. Seule leur feuille de route [11] laisse transparaître les objectifs d'adoption de normes. Dans cette feuille de route, plusieurs thèmes clés sont identifiés pour les travaux de normalisation futurs : la terminologie, la confiance, l'éthique, la sécurité et la sûreté, la résilience et la souveraineté, le respect du droit<sup>40</sup>. De son côté, l'ETSI travaille également sur la création de normes dans le cadre de l'*AI Act*, par exemple sur l'évaluation des systèmes d'IA, la définition de paramètres de qualité, l'explicabilité et la transparence des traitements, ou encore la traçabilité des modèles [41].

33. Chiffres relevés en février 2023, selon le site de l'ISO : <https://www.iso.org/committee/6794475.html>

34. L'ISO comporte plusieurs sous-comités, comme l'ISO JTC1/SC42 travaillant sur des normes transverses à tous les systèmes d'IA, mais également des sous-comités dédiés à certaines applications tel que l'ISO JTC1/SC37 travaillant sur les systèmes biométriques. Nous décomptons ici simplement le travail du JTC1/SC42.

35. ISO/IEC TR 24368 :2022, Information technology — Artificial intelligence — Overview of ethical and societal concerns : <https://www.iso.org/standard/78507.html>

36. ISO/IEC TR 24028 :2020 Information technology — Artificial intelligence — Overview of trustworthiness in artificial intelligence : <https://www.iso.org/standard/77608.html>

37. ISO/IEC 23894 :2023 Information technology — Artificial intelligence — Guidance on risk management : <https://www.iso.org/standard/77304.html>

38. ISO/IEC TR 24027 :2021, Information technology — Artificial intelligence (AI) — Bias in AI systems and AI aided decision making : <https://www.iso.org/standard/77607.html>; ISO/IEC CD TS 12791 Information technology — Artificial intelligence — Treatment of unwanted bias in classification and regression machine learning tasks : <https://www.iso.org/standard/84110.html>

39. ISO 8000-1 :2022 Data quality — Part 1 : Overview : <https://www.iso.org/standard/81745.html>

40. Voir Annexe D de la feuille de route [11].

Le développement de normes harmonisées pour l'*AI Act*, comprenant la création de nouvelles normes et l'adoption par les organismes des pays membres de normes ISO déjà publiées, doit se poursuivre jusqu'en 2025. Les normes européennes jouent donc avec un calendrier très serré.

L'une des problématiques liées aux contraintes temporelles est que les normes publiées en premier ont tendance à être adoptées plus facilement que celles qui suivent. Toutefois, si les normes arrivent trop tôt, elles peuvent conduire à une sélection prématurée et inefficace de la technologie [6]. C'est le cas du clavier QWERTY qui, malgré ses nombreux inconvénients, s'est imposé sur le marché au détriment d'autres bien meilleures solutions [4]. Il y a un risque que cela se reproduise pour l'IA, en normalisant des pratiques encore imparfaites et empêchant de nouvelles pratiques, plus éthiques, de se mettre en place.

De plus, les normes et les processus de certification sont généralement longs à mettre en place. Ainsi, les labels indépendants ont plus de chances de se développer rapidement, avant la création de normes harmonisées.

### 3.2 La coexistence de normes européennes et internationales cache des enjeux géopolitiques

Les organismes européens et internationaux ont l'habitude de collaborer sur l'élaboration de normes. Notamment, les accords de Vienne et de Francfort, conclus respectivement entre la CEN et l'ISO, et entre la CENELEC et l'IEC, facilitent les échanges d'informations entre les organismes et évitent les doublons dans les travaux [32, 24]. Cette collaboration se poursuit jusque dans l'adoption de normes puisque les normes ISO et IEC peuvent intégrer le catalogue des normes européennes par ratification par le CEN-CENELEC. Actuellement, près de 33% des publications du CEN sont issues de l'ISO, et 73% de celles du CENELEC sont issues de l'IEC. En ce qui concerne les normes harmonisées, les normes ISO et IEC ont la priorité lorsqu'elles existent, à moins qu'il ne soit prouvé que la demande de la Commission ne peut être satisfaite par les normes issues de ces organismes internationaux [16].

Dans le cadre de l'IA, les normes européennes sont en retard par rapport aux normes internationales<sup>41</sup>. Ainsi, des appels sont lancés pour approfondir les liens avec l'ISO [44] et converger avec les normes internationales [31]. Cependant, le modèle de conformité présenté par la Commission dans l'*AI Act* renforce la distinction entre les normes européennes harmonisées, bénéficiant d'une présomption de conformité, et les normes internationales non-ratifiées [37]. Il existe donc une tension considérable quant à savoir qui des instances européennes ou internationales développera les normes qui façonneront l'IA en Europe. Cette tension est d'autant plus importante en ce qui concerne les normes relatives aux aspects éthiques des

systèmes, où les visions peuvent diverger entre l'Union européenne et d'autres régions du monde.

Ainsi, certaines parties prenantes remettent en question la confiance dans les normes internationales, car selon elles, rien ne garantit que ces normes soient conformes aux droits et valeurs de l'UE [21]. L'ANEC<sup>42</sup>, une organisation qui défend les intérêts des consommateurs européens dans les processus de normalisation et de certification, se dit ainsi préoccupée par l'adoption en Europe de normes auxquelles ont participé des pays ou des entreprises non européens [54]. Elle propose que la Commission précise si une norme harmonisée peut être confiée à l'ISO, ou si elle doit être développée au sein des organismes de normalisation européens afin de « *préserver les valeurs ou l'éthique européennes* »<sup>43</sup> [54]. De plus, elle insiste sur l'importance de ne pas mettre en péril les valeurs fondamentales européennes dans le seul but de réduire les délais de développement [54]. Cela implique de ne pas laisser aux organismes internationaux la souveraineté des normes et de prendre le temps de développer des normes européennes [13].

La composition de ces organismes de normalisation est également au cœur du débat. Si une grande partie des membres de l'ISO viennent d'Europe de l'ouest, près de la moitié viennent d'ailleurs dans le monde, particulièrement d'Asie et d'Amérique du Nord [40]. De plus, le plus grand groupe de parties prenantes de l'ISO est l'industrie [40]. Certains acteurs considèrent que cela permet à l'ISO de disposer d'une expertise industrielle supérieure à celle des ESOs qui serait bénéfique aux normes européennes [37]. D'autres, critiquent l'absence de représentation de certaines parties prenantes. De fait, les associations représentant les intérêts des consommateurs telles que l'ANEC, ainsi que celles représentant les travailleurs ou les petites entreprises, ne disposent pas officiellement du droit de participer aux travaux de l'ISO et de l'IEC. Par conséquent, elles n'ont pas de voix dans l'élaboration de ces normes internationales qui, à terme, seront adoptées comme normes européennes [16]. Si cette composition revêt une certaine importance dans le cadre général, l'enjeu est encore plus grand lorsqu'il s'agit des normes liées à l'éthique, afin de ne pas laisser l'industrie dicter les codes sociaux et éthiques.

### 3.3 La diversité des initiatives en dehors des instances européennes et internationales

En dehors du cadre bien défini des organismes de normalisation, certaines entités développent leurs propres initiatives pour normaliser ou certifier l'éthique de l'IA. C'est le cas des laboratoires nationaux de métrologie qui travaillent sur le développement de cadres d'évaluation pour l'IA, indépendamment des instances européennes et internationales de normalisation.

**Laboratoires nationaux** En France, le Laboratoire National de Métrologie et d'Essais (LNE), a développé son propre « *Référentiel de certification de processus pour l'IA* » [17]. Il définit un certain nombre d'exigences à res-

41. Voir le retour de DEKRA sur la stratégie de normalisation de la Commission européenne : [https://ec.europa.eu/info/law/better-regulation/have-your-say/initiatives/13099-Standardisation-strategy/F2662668\\_en](https://ec.europa.eu/info/law/better-regulation/have-your-say/initiatives/13099-Standardisation-strategy/F2662668_en)

42. <https://www.anec.eu/>

43. Traduction des auteurs.

pecter lors de la conception, le développement, l'évaluation et le maintien en conditions opérationnelles des systèmes d'IA. Le référentiel insiste notamment sur la transparence des processus, listant les éléments qui doivent être documentés et communiqués au client. Par ailleurs, il ne se base pas sur une technologie en particulier mais sur une manière de fonctionner au niveau de l'entreprise. S'il n'a pas une volonté particulière de traiter l'éthique, il témoigne néanmoins du souhait d'instaurer aux sein des entreprises d'IA un ensemble de bonnes pratiques. Il permet, en outre, la délivrance d'un certificat attestant que l'entreprise remplit bien les conditions indiquées. Quelques entreprises françaises ont déjà pu recevoir ce certificat<sup>44</sup>. Au sein du LNE, le laboratoire d'évaluation des systèmes d'IA<sup>45</sup> travaille actuellement sur d'autres normes pour encadrer les pratiques et auditer les systèmes. Le LNE pourrait, dans le cadre de l'*AI Act*, devenir un organisme notifié, chargé de vérifier la conformité. Sa façon d'aborder l'éthique est donc décisive pour les futurs cadres d'évaluation des systèmes d'IA.

Aux États-Unis, le *National Institute of Standards and Technology* (NIST) travaille également sur des normes pour l'IA [48]. Parmi ses travaux majeurs, les *Face Recognition Vendor Test*<sup>46</sup> posent les bases de l'évaluation des systèmes de reconnaissance faciale. Le NIST compare ainsi les performances de dizaines d'algorithmes provenant de différents fabricants partout dans le monde<sup>47</sup>. Ces comparaisons sont effectuées sur différents critères de précision et, en ce qui concerne l'équité algorithmique, sur des mesures mathématiques d'écarts démographiques [27, 20]. Des métriques comme le taux de divergence ou le ratio de cas d'erreurs entre deux populations permettent ainsi de calculer la gravité d'un biais. Les tests réalisés par le NIST sont très suivis par les industriels et sont déterminants comme argument de vente auprès des clients. Les mesures qu'ils sélectionnent pour évaluer les systèmes ont donc de fortes chances de devenir le standard de référence du domaine. L'ISO pourrait notamment s'en inspirer dans le cadre de son sous-comité 37 sur la biométrie et de sa norme *ISO/IEC WD 19795-10* relative à la quantification de la variation des performances des systèmes biométriques dans les groupes démographiques<sup>48</sup>.

Plus récemment, le NIST a publié un cadre de gestion des risques visant à « améliorer la capacité d'intégrer des considérations de fiabilité dans la conception, le dévelop-

pement, l'utilisation et l'évaluation des produits, services et systèmes d'IA »<sup>49</sup>. Il présente notamment les « caractéristiques des systèmes d'IA dignes de confiance », ainsi que des actions pour assurer leur mise en pratique [47].

**Associations de professionnels** En plus de ces laboratoires nationaux, certaines associations de professionnels développent leurs propres référentiels. Parmi les initiatives internationales sur l'éthique de l'IA, la plus importante est sans doute celle de l'*Institute of Electrical and Electronics Engineers* (IEEE). En 2019, l'IEEE publie un document, reprenant une liste de principes éthiques et construisant un cadre pour les rendre opérationnels [12]. Le rapport présente notamment des recommandations et lignes directrices pour les « normes, la certification, la réglementation ou la législation dans la conception, la fabrication et l'utilisation des systèmes [...] pour le bien-être social »<sup>50</sup>. Cette étape est la première de la *IEEE Global Initiative*<sup>51</sup> sur l'éthique des systèmes autonomes et intelligents. Elle est suivie d'un développement massif de normes relatives à l'éthique des systèmes. La première à être publiée en 2021 est la norme *IEEE Std 7000™-2021* qui traite la prise en compte des préoccupations éthiques lors de la conception des systèmes<sup>52</sup>. D'autres normes suivront, plus spécialisée sur un aspect éthique, comme la norme *IEEE 7001™-2021* sur la transparence<sup>53</sup>. Certaines sont encore en développement, comme la norme *IEEE P7003™* sur les biais algorithmiques [36]. Au total, le projet prévoit le développement de plus d'une quinzaine de normes relatives à l'éthique de l'IA<sup>54</sup>. En plus de ces normes « éthiques », l'IEEE développe également un programme de certification pour l'éthique de l'IA, nommé *CertifAIED*<sup>55</sup>, basé sur la détermination d'un profil de risques du système d'IA puis d'une évaluation selon une série de critères.

D'autres initiatives méritent également d'être soulignées, comme le *Verband der Elektrotechnik* (VDE) en Allemagne, qui a développé son propre cadre pour la caractérisation de l'IA de confiance [50]. Il reprend des valeurs comme la transparence, la responsabilité, la vie privée, l'équité et la fiabilité, et leur attribue des critères et des indicateurs mesurables. Un score peut ensuite être calculé, représentant le degré d'éthique du système.

**Entreprises privées** Certaines entreprises développent aussi leurs propres cadres d'évaluation. Ainsi, Microsoft a publié l'année dernière sa norme sur l'IA responsable [39] dans laquelle sont listées diverses exigences relatives à la

44. C'est le cas par exemple de la société Axionable : <https://www.axionable.com/ia-performante-et-ethique-axionable-decroche-la-lere-certification-delivree-par-le-lne-2/>

45. <https://www.lne.fr/fr/actualites/leia-plateforme-inedite-evaluation-intelligence-artificielle>

46. Tous les FRVT sont répertoriés sur le site du NIST : <https://www.nist.gov/programs-projects/face-recognition-vendor-test-frvt>

47. Voir leur dernier classement : <https://pages.nist.gov/frvt/html/frvt11.html>

48. *ISO/IEC WD 19795-10 Information technology — Biometric performance testing and reporting — Part 10 : Quantifying biometric system performance variation across demographic groups* : <https://www.iso.org/standard/81223.html>

49. Traduction des auteurs

50. Traduction des auteurs

51. <https://standards.ieee.org/industry-connections/ec/autonomous-systems/>

52. *IEEE Std 7000™-2021 : IEEE Standard Model Process for Addressing Ethical Concerns during System Design* : <https://standards.ieee.org/standard/7000-2021.html>

53. *IEEE Std 7001™-2021 : IEEE Standard for Transparency of Autonomous Systems* : <https://standards.ieee.org/ieee/7001/6929/>

54. Ces normes sont recensées sur le site de l'IEEE : <https://standards.ieee.org/initiatives/autonomous-intelligence-systems/standards/>

55. <https://engagestandards.ieee.org/ieeecertified.html>

responsabilité, la transparence, l'équité, la fiabilité, la vie privée et l'inclusivité. Des cadres plus larges sont parfois publiés sous la forme de politique d'entreprise en matière d'IA<sup>56</sup>. Néanmoins, ce ne sont souvent que de simples documents de recommandations et non de réels référentiels d'évaluation.

D'autres plus petites entreprises développent également leurs propres cadres d'évaluation des systèmes, mettant en place des labels d'IA éthiques, basés sur des questionnaires ou des audits algorithmiques. Notons par exemple en France, la société GoodAlgo qui propose un label nommé ADEL, pour évaluer l'éthique des systèmes d'IA basé sur le respect d'un certain nombre de critères « éthiques »<sup>57</sup>. Certains labels se centrent sur des critères plus précis, comme le label de « garantie humaine » pour l'IA en santé portée par le *Digital Medical Hub* et la société Ethik-IA<sup>58</sup>, ou le label GEEIS IA pour l'égalité des chances<sup>59</sup>. Ces labels sont plus faciles et plus rapides à mettre en place que les processus de certification basé sur les normes harmonisés et permettent aux consommateurs d'orienter leurs choix en attendant un examen officiel. Néanmoins les critères d'évaluation des systèmes et de délivrance du certificat ne sont souvent pas publics.

La diversité de voix qui s'expriment peut favoriser la démocratisation de l'éthique, mais elle peut également provoquer une collision des différentes visions. L'analyse de ces voix permet de détecter des dynamiques et de visualiser les différentes orientations que peut prendre l'éthique de l'IA : une approche par la gouvernance et le management d'entreprise, mettant en place des bonnes pratiques pour la planification, la conception ou encore la supervision des systèmes, la gestion des risques et l'anticipation des défis sociaux ; ou une approche par la mesure et l'évaluation de la performance technique des systèmes.

## 4 Les normes et certifications « éthiques » soulèvent des risques

### 4.1 La frontière entre l'éthique et la technique est difficile à tracer

Le développement de ces normes « éthiques » ne fait pas l'unanimité. Selon [57], l'adjectif « éthique » se rapporte à la morale et ne peut donc s'appliquer qu'à « *une démarche, une délibération, une réflexion, une question, un principe, une valeur* ». Ainsi la notion de « conformité éthique » ou de « norme éthique » est contestable car l'éthique est variable par essence [57]. D'autres acteurs considèrent que la normalisation s'éloigne à tort des procédures scientifiques et techniques pour englober des questions sociales qui né-

cessitent un consensus politique [46]. Les normes sont en effet souvent utilisées pour la capture réglementaire [8] et certains vont jusqu'à accuser les organismes de normalisation européens de faire de la politique au service des normes [37].

Par ailleurs, la délimitation du caractère éthique d'une norme n'est pas toujours simple à évaluer. Actuellement, peu de normes abordent activement les défis sociaux [41]. Cependant, même les normes les plus techniques ont des implications sociales. Définir des critères d'équité ou de transparence des processus peut déjà être considéré comme un choix éthique. Selon [63], toutes les normes peuvent donc être considérées comme des normes éthiques implicites.

Notre recensement des initiatives de normalisation permet de distinguer deux types de normes. Les premières sont les normes « éthiques », ou « normes de gouvernance », qui discutent d'aspects sociaux ou des mécanismes à mettre en place pour effectivement respecter les principes éthiques. Ce sont souvent plutôt des normes autour des processus et du cycle de vie de l'IA, considérant la façon dont le système s'intègre dans un contexte social. Un exemple d'une telle norme est l'*IEEE Std 7000™-2021*<sup>60</sup>. Elle inclut notamment la prise en compte des différentes parties prenantes et de leurs valeurs durant les phases d'exploration et de développement des systèmes. La deuxième catégorie de normes correspond aux normes « éthiques implicites », ou « normes de mesure », qui répondent à des principes éthiques par des critères techniques. Par exemple, la norme *ISO/IEC TR 24027 :2021*<sup>61</sup> donne de nombreuses définitions mathématiques de l'équité. Même si elle ne prend pas parti quant à la meilleure mesure à adopter, cette norme comporte inévitablement un aspect éthique en raison du sujet qu'elle aborde. Dans les deux cas, les normes ne dictent pas les résultats attendus d'un système : il n'existe pas pour l'instant de « norme de performance ». Les normes actuelles mettent en lumière des bonnes pratiques, que ce soit dans l'organisation des procédures entourant le système, ou les méthodes techniques de conception et d'évaluation. Dans ce sens, elles ne remplacent pas l'avis d'un juge en décidant de ce qui serait ou non acceptable en matière de droits humains. Leur utilisation n'est pourtant pas neutre par rapport au respect de ces droits.

### 4.2 La difficile définition de critères objectifs

L'identité de l'Europe s'est construite sur la définition d'un ensemble de valeurs démocratiques et économiques qui, dans le cas du numérique, s'allient aux critères techniques et complexifient leur application [34]. Ainsi, les normes techniques ont du mal à s'adapter aux aspects éthiques et sociaux auxquels sont confrontés les systèmes. Puisque l'*AI Act* délègue au processus de normalisation une grande partie des exigences quant à l'IA éthique, savoir quels garde-fous seront mis en place, et par qui, devient un enjeu de gouvernance et de démocratie.

Cet enjeu est renforcé par l'aspect contextuel et culturel de

56. Voir par exemple celle de la société SAP : <https://www.sap.com/documents/2022/01/a8431b91-117e-0010-bca6-c68f7e60039b.html>

57. <https://goodalgo.fr/labels-ethiquement-engages/>

58. <http://esante.gouv.fr/agenda/lancement-du-premier-label-de-garantie-humaine-de-lintelligence-artificielle>

59. <https://arborus.org/label/>

60. Voir supra note 52

61. Voir supra note 38

l'éthique : il n'y a pas une unique façon de l'aborder ou de la définir. Les décisions prises lors des processus de normalisation et de certification ne sont pas des choix neutres, mais plutôt le reflet de valeurs ayant des répercussions sociales. Les « bons » choix à faire ne sont pas universels. Il existe pourtant une volonté de converger vers des valeurs et principes communs, comme en témoignent les nombreux documents sur l'éthique de l'IA publiés au niveau européen [49] et international [59]. Toutefois, dans le cadre de la normalisation, ces choix sont rarement présentés comme des choix de valeurs, mais comme des choix purement techniques et donc « neutres ». Par exemple, l'adhésion aux normes harmonisées est supposée fournir un moyen « objectivement vérifiable » de se conformer aux exigences essentielles [51]. La certification « éthique » de l'IA se confronte alors aux mêmes obstacles que d'autres domaines comme le commerce équitable ou les normes environnementales, pour lesquels les labels et certifications servent de marque de qualité et déterminent implicitement les valeurs morales et sociales à suivre [42].

Les différentes lignes directrices et chartes en matière d'éthique édictées par les instances internationales ont déjà été critiquées pour le prisme qu'elles adoptent dans leur vision de l'éthique. Toutefois, cette critique revêt une importance accrue dans le cas des normes et des certifications qui non seulement codifient les principes éthiques, mais les figent également dans nos pratiques. La légitimité des organismes de normalisation quant à l'imposition de cette vision de l'éthique est donc remise en question.

De plus, l'éthique de l'IA est dynamique et un système ne saurait être perpétuellement aligné sur des valeurs éthiques [8]. En cela, une certification statique en matière d'éthique de l'IA ne peut saisir aucun enjeu pertinent. Une IA certifiée « éthique », le restera alors même que le système et son contexte évoluent. Bien que certains organismes souhaitent réévaluer régulièrement le système et garder un certificat à jour<sup>62</sup>, une telle marque ne pourra jamais être totalement adaptée à un contexte spécifique, sinon elle perdrait son caractère universellement applicable.

En plus d'instituer arbitrairement ces valeurs, de nombreuses entités se contentent d'adopter des normes de façon symbolique, sans changer leurs pratiques [14]. Le respect de ces normes devient alors un simple argument de vente et l'obtention d'une marque ou d'un label « éthique » ne fait que légitimer ces pratiques, et perpétue une culture de l'*ethics washing* [61]. Certains craignent alors que les normes développées au niveau européen ne soient que trop peu contraignantes, laissant ainsi l'industrie agir à sa guise [51].

Les choix effectués pour aboutir à une norme ou un référentiel d'évaluation impactent fortement l'évolution des systèmes d'IA. Par exemple, la diffusion généralisée du cadre de gestion des risques du NIST [47] sur les réseaux et par diverses instances pourrait favoriser son adoption à grande échelle, préemptant de fait l'émergence de référentiels spécifiquement européens. Bien que le référentiel NIST puisse

être compatible avec la vision de l'*AI Act* concernant l'IA « de confiance », le fait que cela soit une norme américaine constitue un obstacle pour l'Europe dans sa volonté de définir sa propre vision.

La question de l'équité illustre les difficultés à converger sur une vision commune de l'éthique. En effet, de nombreuses définitions techniques de l'équité existent [43], et sont même souvent contradictoires [35]. Le choix des approches mathématiques aura un impact sur les droits des personnes. Un système équitable selon une définition ne l'est pas forcément selon une autre. C'est le cas du logiciel COMPAS utilisé aux Etats-Unis pour prédire le taux de récidive de criminels, qui a été accusé de pénaliser les personnes afro-américaines selon un certain critère d'équité [5], alors qu'il respectait l'équité selon une autre méthode de mesure [45]. Cette diversité des mesures d'équité risque de mener à des choix de simplification stratégique [3] : les constructeurs affichent simplement la mesure de l'équité qui montre que leur système est exempt de biais et donc « juste » selon eux, et pas les autres mesures.

Instaurer une unique mesure, ou un ensemble limité de mesures dans les normes risquent d'accentuer cette tendance. De plus, ces choix sont intégrés dans les systèmes d'IA sans véritable débat public.

### 4.3 La certification comme garantie de l'éthique

La certification vis-à-vis des normes techniques est souvent perçue comme une garantie de sécurité [18]. C'est le cas notamment du marquage CE, compris à tort comme un gage de qualité alors même qu'il ne signifie que la conformité à la réglementation. De plus, il est généralement apposé par le fabricant lui-même. Or un produit marqué CE peut également avoir des failles de sécurité [62].

De la même façon que pour le marquage CE et les normes de sécurité, il est probable qu'une certification vis-à-vis de normes « éthiques » soit également considérée par le consommateur comme une garantie de l'éthique du produit ou procédé. Pourtant, respecter une notion mathématique de l'équité telle que définie dans une norme ne garantit pas que le système ne discriminer pas. De même, respecter une norme de conception éthique qui donnerait pour critère la réalisation d'une étude d'impacts permet de mesurer certaines conséquences du déploiement d'un système, mais cela ne signifie pas que tous les préjudices possibles ont été pris en compte, ou que les mesures de protection prises sont suffisantes. La marque de certification pourrait alors induire chez les utilisateurs des systèmes et des citoyens un faux sentiment de protection.

Les normes doivent alors être co-construites avec un régime de responsabilité [38]. Toutefois, cette responsabilité est plus dure à prouver une fois les produits et processus certifiés : lorsque la conformité est démontrée, il est difficile de s'y opposer. Certains considèrent qu'il y a une « juridification » du processus de normalisation [58]. Depuis l'affaire James Elliott Construction [60], les normes harmonisées peuvent même être considérées comme des dispositions du droit de l'Union européenne.

62. C'est le cas notamment du référentiel du LNE [17]

Utiliser des normes techniques pour assurer le respect de droits fondamentaux est risqué, car les normes rentreraient alors en concurrence avec le travail du législateur et des juges, qui sont les seuls compétents à évaluer, et mettre en équilibre, les ingérences dans différents droits fondamentaux [26]. Une norme technique visant à protéger la sécurité des personnes vise également à préserver un droit fondamental, notamment le droit à la vie. Mais une norme technique sur la résistance au feu, par exemple, rentrera moins en conflit avec le rôle premier des juges. Certes, un juge pourra toujours estimer que l'application d'une norme de sécurité dans un cas précis était insuffisant. Mais la norme technique en matière de sécurité ne sera pas perçue par les juges comme une ingérence dans leur travail, alors qu'une norme en matière de respect des droits fondamentaux pourrait l'être.

## 5 Conclusion

Notre cartographie des acteurs et initiatives offre une grille de lecture pour mieux comprendre les enjeux de la normalisation et de la certification de l'éthique de l'IA. Nous pensons qu'elle peut servir à la fois aux dirigeants d'entreprises, aux chercheurs et aux développeurs de systèmes d'IA dont les pratiques vont être impactées par ces normes, ainsi qu'aux juristes, sociologues et philosophes s'intéressant à l'évolution des représentations de l'éthique normative.

Nous avons vu que de plus en plus d'acteurs se lancent dans la mise en place de normes ou de processus de certification pour l'éthique de l'IA, alors que le concept même d'une norme technique pour l'éthique fait débat. Nous avons dressé un panorama de leurs activités et avons cherché à identifier les tensions qui peuvent émerger entre les différents organismes. Cet écosystème est notamment caractérisé par une rivalité de gouvernance opposant d'un côté les normes européennes, et de l'autre les normes internationales. L'émergence de cadres normatifs en dehors de ces structures officielles permet de multiplier les garde-fous de façon plus rapide que les processus habituels, mais risquent également de dévoyer l'éthique en diffusant des critères d'évaluation manquant toute légitimité démocratique. Toutefois, il est essentiel de garder à l'esprit que l'éthique est intrinsèquement contextuelle et, dans une certaine mesure, subjective. Une certification ne pourra garantir le respect absolu de principes éthiques et de droits individuels. La question de savoir qui est responsable du développement de ces cadres d'évaluation est cruciale, car ils façonneront le développement des futurs systèmes, avec parfois un impact mondial. Afin de ne pas donner aux utilisateurs et aux citoyens la fausse impression que leurs droits sont nécessairement préservés grâce au marquage CE, les limites de cette certification doivent toujours être mises en avant. Même si le marquage CE témoigne d'une certaine démarche de qualité, pour identifier et réduire les biais par exemple, cette démarche ne change rien en ce qui concerne la responsabilité de l'opérateur pour une décision algorithmique discriminatoire. Ainsi, la certification ne fonctionnera que si

elle est accompagnée d'un régime de responsabilité et de procédures de recours en cas de préjudice.

Enfin, notons que si les initiatives visant à normaliser et à certifier l'éthique cherchent à prévenir les dérives potentielles, les mesures de protection précises qu'elles définiront restent encore indéterminées. En effet, la plupart des normes dans ce domaine ne sont pas encore développées, et celles qui le sont, ne sont pas encore opérationnelles [22]. Enfin, l'interconnexion entre les instances de normalisation techniques et les instances étudiant le respect par l'IA des droits fondamentaux<sup>63</sup>, semble indispensable.

## Remerciements

Merci à Tiphaine Viard pour ses commentaires précieux. Cette recherche a été financée dans le cadre du projet LIMPID<sup>64</sup> (Projet ANR 20-CE23-0028).

## Références

- [1] Regulation (EU) No 1025/2012 of the European Parliament and of the Council of 25 October 2012 on European standardisation.
- [2] AFNOR. Parler normes couramment. L'essentiel, 2014.
- [3] Ulrich Aivodji, Hiromi Arai, Olivier Fortineau, Sébastien Gambis, Satoshi Hara, and Alain Tapp. Fairwashing : the risk of rationalization. In *Proceedings of the 36th International Conference on Machine Learning*, pages 161–170. PMLR, 2019.
- [4] Robert H Allen and Ram D Sriram. The Role of Standards in Innovation. *Technological Forecasting and Social Change*, 64(2) :171–181, 2000.
- [5] Julia Angwin, Jeff Larson, Surya Mattu, and Lauren Kirchner. Machine Bias. *ProPublica*, 2016.
- [6] Knut Blind. The impact of standardisation and standards on innovation. In *Handbook of Innovation Policy Impact*, pages 423–449. 2016.
- [7] Joanna Bryson and Alan Winfield. Standardizing Ethical Design for Artificial Intelligence and Autonomous Systems. *Computer*, 50(5) :116–119, 2017.
- [8] Joanna J Bryson. Belgian and Flemish Policy Makers' Guide to AI Regulation. 2022.
- [9] Céline Castets-Renard and Philippe Besse. Ex ante Accountability of the AI Act : Between Certification and Standardization, in *Pursuit of Fundamental Rights in the Country of Compliance*, 2022.
- [10] Tom Cellucci. Developing Operational Requirements - A Guide to the Cost-Effective and Efficient Communication of Needs. Technical report, U.S. Department of Homeland Security, 2008.
- [11] CEN-CENELEC. CEN-CENELEC Focus Group Report : Road Map on Artificial Intelligence (AI), 2020.

63. Voir par exemple le Comité sur l'Intelligence Artificielle (CAI) du Conseil de l'Europe : <https://www.coe.int/fr/web/artificial-intelligence/cai>

64. <https://limpid.telecom-paris.fr/>

- [12] Raja Chatila and John C. Havens. Ethically Aligned Design. A Vision for Prioritizing Human Well-being with Autonomous and Intelligent Systems. Version 2. Technical report, The IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems, 2019.
- [13] Julien Chiaroni and Arno Pons. IA de confiance - Opportunité stratégique pour une souveraineté industrielle et numérique. Technical report, Digital New Deal, 2022.
- [14] Peter Cihon, Moritz J. Kleinaltenkamp, Jonas Schuett, and Seth D. Baum. AI Certification : Advancing Ethical Practice by Reducing Information Asymmetries. *IEEE Transactions on Technology and Society*, 2(4) :200–209, 2021.
- [15] European Commission. Commission Notice - The 'Blue Guide' on the implementation of EU product rules 2022. Technical report, Information from European Union Institutions, Bodies, Offices and Agencies, 2022.
- [16] Pierluigi Cuccuru. Interest Representation in European Standardisation : The Case of CEN and CENELEC, 2019.
- [17] Laboratoire National de Métrologie et d'Essais (LNE). Référentiel de certification de processus pour l'IA - Conception, développement, évaluation et maintien en conditions opérationnelles, 2021.
- [18] Av de Tervueren. ANEC Position Paper on CE marking "Caveat Emptor - Buyer Beware", 2012.
- [19] Henk J. de Vries. *Standardization : A Business Approach to the Role of National Standardization Organizations*. Springer US, Boston, MA, 1999.
- [20] David L. Duewer. Face Recognition Vendor Test (FRVT) Part 8 : Summarizing Demographic Differentials. Technical Report NIST IR 8429, National Institute of Standards and Technology (NIST), 2022.
- [21] European Trade Union Confederation (ETUC). Feedback on the (roadmap) consultation of citizens and stakeholders on the forthcoming "EU Standardisation strategy", 2021.
- [22] European Commission. Joint Research Centre. AI Watch, AI standardisation landscape state of play and link to the EC proposal for an AI regulatory framework. Technical report, Publications Office, LU, 2021.
- [23] 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, 2021.
- [24] European Committee for Electrotechnical Standardization (CENELEC). CENELEC Guide 13 FAQ. Frequently Asked Questions on the Frankfurt Agreement. Edition 1, 2017.
- [25] DIN German Institute for Standardization e. V. *Economic benefits of standardization - Summary of results*. Beuth Verlag, DE, 2000.
- [26] Mélanie Gornet and Winston Maxwell. Intelligence artificielle : normes techniques et droits fondamentaux, un mélange risqué. *The Conversation*, 2022.
- [27] Patrick J. Grother, Mei L. Ngan, and Kayee K. Hanaka. Face Recognition Vendor Test (FRVT) Part 3 : Demographic Effects. Technical Report NIST IR 8280, National Institute of Standards and Technology (NIST), 2019.
- [28] Laurens Hernalsteen and Constant Kohler. Drafting Harmonized Standards in support of the Artificial Intelligence Act (AIA) - CEN-CENELEC, 2022.
- [29] Kashmir Hill. Wrongfully Accused by an Algorithm. *The New York Times*, 2020.
- [30] Marion Ho-Dac. La normalisation, clé de voûte de la réglementation européenne de l'intelligence artificielle (AI Act). *Dalloz IP/IT*, pages 228–236, 2023.
- [31] Japan Business Council in Europe (JBCE). The Roadmap of European Standardisation Strategy, 2021.
- [32] ISO and CEN. Foire aux questions relatives à l'Accord de Vienne, 2016.
- [33] Anna Jobin, Marcello Ienca, and Effy Vayena. The global landscape of AI ethics guidelines. *Nature Machine Intelligence*, 1(9) :389–399, 2019.
- [34] Jonathan Keller and Claire Levallois-Barth. La fragile définition de l'identité européenne par ses valeurs numériques. *Revue générale du droit. Chronique de droit de l'Union*, 2021.
- [35] Jon Kleinberg, Sendhil Mullainathan, and Manish Raghavan. Inherent Trade-Offs in the Fair Determination of Risk Scores. *arXiv :1609.05807*, 2016.
- [36] Ansgar Koene, Liz Dowthwaite, and Suchana Seth. IEEE P7003™ standard for algorithmic bias considerations : work in progress paper. In *Proceedings of the International Workshop on Software Fairness*, pages 38–41. ACM, 2018.
- [37] Mark McFadden, Kate Jones, Emily Taylor, and Georgia Osborn. Harmonising Artificial Intelligence : The role of standards in the EU AI Regulation. 2021.
- [38] Jacob Metcalf, Emanuel Moss, Elizabeth Anne Watkins, Ranjit Singh, and Madeleine Clare Elish. Algorithmic Impact Assessments and Accountability : The Co-construction of Impacts. In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, pages 735–746, New York, NY, USA, 2021. Association for Computing Machinery.
- [39] Microsoft. Microsoft Responsible AI Standard, v2, 2022.
- [40] Mari Morikawa and Jason Morrison. A Survey of Participation in ISO's International Standards Development Processes. 2004.

- [41] Markus Mueck, Scott Cadzow, Cadzow Communications, and Suno Wood. ETSI White Paper No. #52. ETSI Activities in the field of Artificial Intelligence Preparing the implementation of the European AI Act. 1st Edition. 2022.
- [42] Warwick Murray, Overton John, and Howson Kelle. *Ethical Value Networks in International Trade : Social Justice, Sustainability and Provenance in the Global South*. Edward Elgar Publishing, 2022.
- [43] Arvind Narayanan. Tutorial : 21 fairness definitions and their politics, 2019.
- [44] Executive Board Netherlands Standardization Institute (NEN). Statement on the European Commission’s Roadmap for a Standardization Strategy, 2021.
- [45] Northpointe. Practitioner’s Guide to COMPAS Core. Technical report, 2019.
- [46] European Council of Engineers Chambers. Statement on “Roadmap Standardization Strategy”, 2021.
- [47] National Institute of Standards and Technology (NIST). AI Risk Management Framework : AI RMF (1.0). Technical Report NIST AI 100-1, 2023.
- [48] National Institute of Standards and Technology (NIST). NIST AI Program. Artificial Intelligence : The Vitals, 2023.
- [49] Independent High-Level Expert Group on Artificial Intelligence set up by the European Commission (HLEG). Ethics guidelines for trustworthy AI. Technical report, European Commission, 2019.
- [50] Christoph Peylo, Dirk Slama, Sebastian Hallensleben, Andreas Hauschke, and Stephanie Hildebrandt. VCIO based description of systems for AI trustworthiness characterisation. Technical Report VDE SPEC 90012 V1.0 (en), Verband der Elektrotechnik (VDE), 2022.
- [51] Hadrien Pouget. The EU’s AI Act Is Barreling Toward AI Standards That Do Not Exist. *Lawfare*, 2023.
- [52] Thomas Reardon, Jean-Marie Codron, Lawrence Busch, R. James Bingen, and Craig Harris. Global change in agrifood grades and standards : agribusiness strategic responses in developing countries. *International Food and Agribusiness Management Review*, 02(3-4), 1999.
- [53] Reuters. Amazon ditched AI recruiting tool that favored men for technical jobs. *The Guardian*, 2018.
- [54] Stephen Russell. Roadmap for the standardisation strategy. ANEC response. 2021.
- [55] Björn ten Seldam and Alex Brenninkmeijer. The Dutch benefits scandal : a cautionary tale for algorithmic enforcement. *EU Law Enforcement*, 2021.
- [56] Catherine Tessier. Ethique et IA : analyse et discussion. In *Conférence Nationale en Intelligence Artificielle (CNIA)*, 2021.
- [57] Catherine Tessier. Parler du numérique et de son éthique : un questionnement... éthique. In *Pour une éthique du numérique*, pages 97–105. Puf edition, 2022.
- [58] Carlo Tovo. Judicial review of harmonized standards : Changing the paradigms of legality and legitimacy of private rulemaking under EU law. *Common Market Law Review*, 55(4), 2018.
- [59] United Nations Educational Scientific and Cultural Organization (UNESCO). Recommendation on the Ethics of Artificial Intelligence. Technical report, 2021.
- [60] Arnaud van Waeyenberge and David Restrepo. James Elliot construction : A “New(ish) approach” to judicial review of standardisation. *European Law Review*, 42 :882–893, 2017.
- [61] Ben Wagner. Ethics as an escape from regulation. From “ethics-washing” to ethics-shopping? In *BEING PROFILED : COGITAS ERGO SUM*. Amsterdam University Press, 2018.
- [62] I. M. E. Wentholt, J. B. L. Hoekstra, A. Zwart, and J. H. DeVries. Pendra goes Dutch : lessons for the CE mark in Europe. *Diabetologia*, 48(6) :1055–1058, 2005.
- [63] Alan Winfield. Ethical standards in robotics and AI. *Nature Electronics*, 2(2) :46–48, 2019.
- [64] Wolfgang Ziegler. A Landscape Analysis of Standardisation in the Field of Artificial Intelligence. *Journal of ICT Standardization*, pages 151–184, 2020.