



HAL
open science

Vers un référentiel universel pour un usage éthique de l'intelligence artificielle

Halima Brahmi, Saïda Belouali, Yves Demazeau, Toumi Bouchentouf, Nissrine Hassini Alaoui

► **To cite this version:**

Halima Brahmi, Saïda Belouali, Yves Demazeau, Toumi Bouchentouf, Nissrine Hassini Alaoui. Vers un référentiel universel pour un usage éthique de l'intelligence artificielle. *East African Journal of Information Technology*, 2023, 6 (1), pp.91-106. 10.37284/eajit.6.1.1226 . hal-04294053

HAL Id: hal-04294053

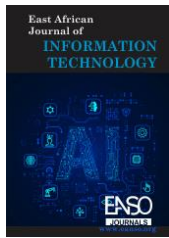
<https://hal.science/hal-04294053>

Submitted on 19 Nov 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.

Public Domain



East African Journal of Information Technology

eajit.eanso.org

Volume 6, Issue 1, 2023

Print ISSN: 2707-5346 | Online ISSN: 2707-5354

Title DOI: <https://doi.org/10.37284/2707-5354>

EASO

EAST AFRICAN
NATURE &
SCIENCE
ORGANIZATION

Original Article

Vers un référentiel universel pour un usage éthique de l'intelligence artificielle

Halima Brahmi¹, Saïda Belouali¹, Yves Demazeau², Toumi Bouchentouf^d & Nissrine Hassini Alaoui^{1*}

¹ Université Mohammed Premier, Oujda, B.P. 524 Oujda 60000 Maroc

² Université Grenoble Alpes, 621 Av. Centrale, 38400 Saint-Martin-d'Hères, France

* Correspondance ORCID ID: <https://orcid.org/0000-0002-9325-1129>; Email: n.hassinialaoui@ump.ac.ma

Article DOI: <https://doi.org/10.37284/eajit.6.1.1226>

Date Published: **ABSTRACT**

23 May 2023

Keywords :

Intelligence
Artificielle,
Usage Éthique,
Principes Éthiques,
Référentiel
Universel

Cet article met en question l'universalité des codes éthiques pour assurer un développement sain et une utilisation responsable de l'intelligence artificielle. Le but de cette étude est de proposer un référentiel commun qui catégorise un ensemble de codes éthiques destiné à guider l'usage éthique de l'intelligence artificielle. Afin d'atteindre cet objectif, une étude exploratoire a été menée sur les codes éthiques de onze entités internationales. Nous avons synthétisé ces codes à travers une méthode d'Analyse en Correspondances Multiples (ACM) de l'Analyse Factorielle (AF). Les résultats ont permis de regrouper 31 principes éthiques étudiés en six catégories communes, puis d'engendrer un référentiel universel par catégorie pour guider l'usage éthique de l'intelligence artificielle dans le monde.

APA CITATION

Brahmi, H., Belouali, S., Demazeau, Y., Bouchentouf, T. & Alaoui, N. H. (2023). Vers un référentiel universel pour un usage éthique de l'intelligence artificielle. *East African Journal of Information Technology*, 6(1), 91-106. <https://doi.org/10.37284/eajit.6.1.1226>

CHICAGO CITATION

Brahmi, Halima, Saïda Belouali, Yves Demazeau, Toumi Bouchentouf and Nissrine Hassini Alaoui. 2023. "Vers un référentiel universel pour un usage éthique de l'intelligence artificielle". *East African Journal of Information Technology* 6 (1), 91-106. <https://doi.org/10.37284/eajit.6.1.1226>.

HARVARD CITATION

Brahmi, H., Belouali, S., Demazeau, Y., Bouchentouf, T. & Alaoui, N. H. (2023) "Vers un référentiel universel pour un usage éthique de l'intelligence artificielle", *East African Journal of Information Technology*, 6(1), pp. 91-106. doi: 10.37284/eajit.6.1.1226.

IEEE CITATION

H. Brahmi, S. Belouali, Y. Demazeau, T. Bouchentouf, & N. H. Alaoui, "Vers un référentiel universel pour un usage éthique de l'intelligence artificielle", *EAJIT*, vol. 6, no. 1, pp. 91-106, May. 2023.

MLA CITATION

Brahmi, Halima, Saïda Belouali, Yves Demazeau, Toumi Bouchentouf & Nissrine Hassini Alaoui. "Vers un référentiel universel pour un usage éthique de l'intelligence artificielle". *East African Journal of Education Studies*, Vol. 6, no. 1, May. 2023, pp. 91-106, doi:10.37284/eajit.6.1.1226.

INTRODUCTION

Actuellement, l'Intelligence Artificielle (IA) ouvre de multiples possibilités de développement économique, social ou éducatif dans le monde (Thomasset, 2020). Face aux problématiques d'ordre éthique imposées par l'émergence de l'IA dans les différents secteurs de notre vie quotidienne, un usage éthique, sain et sûr de l'IA s'avère nécessaire pour garantir le respect des valeurs suprêmes érigées par les communautés internationales et admises comme consensuelles. Le questionnement sur l'universalité et l'unicité d'un code éthique (Belouali et al., 2022) nous a conduits à réfléchir sur la manière et la possibilité d'établir un code commun par catégorie qui servira de base pour parler un langage commun et unifié.

La légitimité du questionnement sur l'universalité de l'éthique dans le domaine de l'intelligence artificielle est prise en considération vu le caractère contingent de l'éthique et le relativisme lié au concept de valeur lui-même (Belouali, 2017). Ainsi, l'encouragement de l'usage de l'IA doit être accompagné des lignes directrices qui exigent l'usage sûr, responsable et contrôlable de cette technologie afin d'éviter son mauvais ciblage (Besse et al., 2018).

Dans ce contexte, le présent travail mène une étude exploratoire examinant les codes éthiques de l'usage de l'IA dans onze entités internationales, notamment l'ASILOMAR, l'Organisation de Coopération et de Développement Économiques (OCDE), l'Organisation des Nations Unies pour l'Éducation, la Science et la Culture (UNESCO), la Chine, les Émirats Arabes Unis (EAU), l'Union Européenne (UE), les États-Unis d'Amérique (USA), le Canada, le Japon, le Royaume-Uni et les GAFAM (Google, Apple, Facebook, Amazon et Microsoft).

Malgré l'hétérogénéité de ce corpus objet d'étude constitué par une conférence, des institutions étatiques et des organismes privés, nous avons estimé intéressant de procéder à un parallélisme dont l'objectif est de questionner leurs propres

codes éthiques édictés et publiés indépendamment de la qualité de l'entité ou de la procédure suivie pour établir ces chartes et codes. L'objectif est de les synthétiser par le biais de l'approche de l'Analyse en Correspondances Multiples (ACM) et d'explorer la possibilité de les voir s'aligner vers un référentiel universel des principes éthiques de l'IA.

Pour répondre à cet objectif, la première partie de ce papier présentera une exploration des codes éthiques de onze entités internationales objet de notre étude. La deuxième partie de ce travail synthétise ces codes éthiques via la méthode de l'Analyse en Correspondance Multiple (ACM) de l'analyse factorielle. Enfin, nous proposerons un référentiel de l'usage éthique de l'IA qui pourrait être extrait à travers l'identification des points communs entre les codes éthiques étudiés.

Exploration des principaux codes éthiques concernant l'utilisation de l'IA

Cette partie du travail expose les instructions et les codes éthiques de l'usage de l'IA d'ASILOMAR conférence, de l'OCDE, de l'UNESCO, de la Chine, des Emirats Arabes Unis, de l'Union Européenne, des Etats Unis d'Amérique, du Canada, du Japon, du Royaume Uni et des GAFAM (Google, Apple, Facebook, Amazon et Microsoft).

Les principes éthiques de l'intelligence artificielle qui ont été convenus lors de la conférence d'Asilomar

En 2017, un groupe de chercheurs en intelligence artificielle (IA) venant à la fois du monde académique économique, droit, éthique et philosophique a créé des principes éthiques pour l'IA lors de la conférence de l'Asilomar dans le but d'encadrer le développement de l'intelligence artificielle.

Il s'agit de 23 lignes directrices en relation avec la recherche et le développement exponentiel de l'intelligence artificielle (IA) appelées « Asilomar AI Principles ». Le choix de la conférence d'Asilomar comme référence émane de la volonté d'élargir notre champ d'étude tout en confirmant

l'importance du concept d'IA bénéfique, qui est un élément clé de cette conférence (Principes de l'IA d'Asilomar, 2017)¹.

En effet, les 23 lignes directrices pour guider l'usage responsable et bénéfique de l'IA convenu lors de la conférence d'Asilomar ont été réduites à 10 principes éthiques clés après des discussions et des débats : « *La sécurité, la transparence et l'explicabilité, la responsabilité, les valeurs humaines, le respect de la vie privée, les bénéfices et la prospérité partagés, le contrôle humain et l'interdiction de la course aux armements de l'IA* ». (Asilomar IA principes, 2017).

Les Principes éthiques de l'IA selon l'OCDE²

En 2019, L'Organisation de Coopération et de Développement Économiques (OCDE) présente des recommandations et des conseils sur l'intelligence artificielle en s'appliquant d'ores et déjà à l'IA à l'instar des principes relatifs à la protection de la vie privée et des données, à la gestion du risque de sécurité numérique et à la conduite responsable des entreprises. L'OCDE a mis en avant cinq nouveaux principes éthiques recommandés pour promouvoir une utilisation responsable et fiable de l'IA, en complément des principes précédents de l'organisation. Ces cinq principes sont considérés comme un ensemble cohérent. Ces principes regroupent : « *la croissance inclusive, le développement durable et le bien-être, les valeurs centrées sur l'humain, l'équité, la transparence et l'explicabilité, la robustesse, la sûreté et sécurité, et la responsabilité* » (Collection : Instruments juridiques de l'OCDE).

Les Principes éthiques de l'IA de l'UNESCO

En 2020, lors de sa Conférence générale, l'UNESCO a examiné les questions éthiques liées à l'intelligence artificielle. Elle a présenté l'éthique

de l'IA comme une réflexion normative systématique qui repose sur un cadre global et en constante évolution de valeurs et de principes. L'objectif de des recommandations dégagées est l'analyse des caractéristiques des systèmes d'IA qui ont une importance cruciale d'un point de vue éthique. (UNESCO, 2020).

Les questions éthiques soulevées par les systèmes de l'IA selon l'UNESCO se focalisent sur la prise de décision éthique de l'IA, notamment leur impact sur l'emploi, sur les interactions sociales, les soins de santé, l'éducation, les médias, la liberté d'expression, l'accès à l'information, la protection de la vie privée, la démocratie, la discrimination et la militarisation. Ces enjeux peuvent s'appliquer à la fois à des personnes physiques et à des personnes morales. Ils sont liés aux principes : « *de proportionnalité et innocuité, de sûreté et sécurité, d'équité et non-discrimination, de durabilité, de surveillance et décision humaines, de transparence et l'explicabilité, de responsabilité et redevabilité, de sensibilisation et éducation ainsi qu'à la gouvernance et la collaboration multipartites et adaptatives* » (Noiseau et al., 2021).

L'émergence des principes éthiques de l'IA en Chine

En 2019, le comité national de gouvernance en Chine a publié les principes de la gouvernance de la prochaine génération de système d'intelligence artificielle (IA) pour développer une intelligence artificielle responsable (Huang et Gong, 2020). Le but de ces principes est d'émerger la gouvernance éthique de l'IA dont le développement a atteint de nouvelles caractéristiques telles que l'intégration transversale, la coordination homme-machine et l'intelligence ouverte et collective (Zeng et al., 2019).

¹ La conférence dénommée Bénéficiaire AI 2017 organisée par le Futur of Life Institute (FLI, institut pour l'avenir de la vie). La conférence s'est tenue à Asilomar, en Californie au mois de janvier dernier.

Pour rappel, le Future of Life Institute (FLI) est une association de volontaires basée dans la région de

Boston, cherchant à diminuer les risques existentiels menaçant l'humanité, en particulier ceux provenant de l'intelligence artificielle (IA).

² OCDE, Recommandation du Conseil sur l'intelligence artificielle, OECD/LEGAL/0449 consulté le 17/05/2022 au Conditions d'utilisation - OCDE (oecd.org)

Dans ce sens, le ministère des sciences et de la technologie de la République Populaire de Chine (MOST) a proposé huit principes éthiques reflétant non seulement la bonne volonté et les croyances morales de la société, mais exigeant également des solutions techniques réalisables et complètes pour mettre en œuvre des contraintes éthiques dans les modèles, les algorithmes et les produits de l'IA (Laney, 2019).

Ces huit Principes éthiques de l'IA sont liés à « *la sûreté, la confidentialité, la sécurité, la fiabilité, la transparence, la responsabilité, l'impartialité, et la gouvernance agile* ».

Les Principes éthiques de l'IA selon Les Emirats Arabes Unis

La ville écosystème « Smart Dubaï » (Bint Butti Bin Bisher, 2018) a été établie avec pour objectif de promouvoir l'IA éthique dans sa stratégie de développement. Cette initiative vise également à sensibiliser les entreprises, les universités et les particuliers à la manière dont les systèmes d'IA peuvent être utilisés de manière responsable.

Selon les Emirats Arabes (EAU), les directives éthiques de l'intelligence artificielle peuvent être résumées en quatre principes éthiques fondamentaux notamment « *l'équité, la transparence, la responsabilité et la compréhension* » (Bint Butti Bin Bisher, 2018).

Les Principes éthiques de l'IA de l'Union Européenne

Les droits fondamentaux constituent une source majeure d'orientation pour l'élaboration de cadres éthiques pour les systèmes d'IA. Ces droits sont considérés comme des normes universelles qui garantissent les libertés et les droits essentiels de chaque individu, et sont généralement inclus dans les constitutions et les lois nationales (GEHN, 2019).

En 2019, un groupe d'experts de haut niveau en matière de l'IA a soumis un document révisé à la commission européenne (GEHN, 2019). Ce groupe présente un ensemble de pays qui ont signé une déclaration commune dans laquelle ils s'engagent à accentuer leur coopération dans le

domaine de l'intelligence artificielle. Selon l'Union européenne, la responsabilité et la dignité de confiance de l'IA sont deux fondements éthiques fondamentaux de cette nouvelle technologie (Livre blanc de la Commission Européenne, 2020).

Pour parvenir à ces deux fondements majeurs, les organisations de coopération de l'UE s'engagent à tenir en compte un ensemble de réglementations et d'exigences applicables visant à respecter les principes du « *facteur humain et du contrôle humain, de la robustesse et de la sécurité, du respect de la vie privée et de la gouvernance des données, de la transparence et la traçabilité, de la diversité, la non-discrimination et de l'équité, du bien-être sociétal et environnemental et de la responsabilisation* » (Lausson, 2019).

Les Principes éthiques de l'IA aux États-Unis d'Amérique

Dans le but de promouvoir un usage digne de confiance de l'intelligence artificielle, la constitution et les lois des États-Unis d'Amérique publient en Décembre 2020, un décret exécutif pour la promotion de cet usage éthique (Bietti et al., 2019).

Le questionnement sur la manière d'assurer une intelligence artificielle fiable et digne de confiance, permet au gouvernement fédéral de développer une infrastructure et une technologie réglementée par des principes éthiques lors de la conception, du développement, et de l'acquisition de l'IA pour les organismes adhérant à ce gouvernement (Trump, 2019).

Ces principes éthiques sont : « *la légalité et le respect des valeurs de la nation, la détermination et l'axe sur la performance, la précision, la fiabilité et l'efficacité, la sûreté et la résilience, la compréhension, la responsabilité et la traçabilité, la surveillance régulière, la transparence et la responsabilité* » (Bietti et al., 2019).

Les Principes éthiques de l'IA selon le Canada

Les trois objectifs principaux pour une évolution responsable de l'intelligence artificielle sont publiés en 2018 par Montréal dans le but

d'élaborer un ensemble de principes éthique pour le développement et le déploiement de la technologie de l'IA (Dihac et al.2018). Ce cadre permet d'orienter la transition numérique vers un développement inclusif, équitable et écologiquement durable de l'IA pour bénéficier de la révolution de cette technologie et accéder à l'ouverture d'un espace de dialogue éthique à l'échelle nationale et internationale.

La déclaration de Montréal prend en considération dix codes éthiques incluant « *le bien-être, le respect de l'autonomie, la protection de l'intimité et de la vie privée, la solidarité, la participation démocratique, l'équité, l'inclusion de la diversité, la prudence, la responsabilité et le développement durable* » (Dihac et al., 2018).

Les Principes éthiques de l'IA selon le Japon

En Mai 2017, les membres de la Société Japonaise d'intelligence artificielle (JSAI) formalisent neuf lignes directrices éthiques pour servir de base morale et prendre conscience des responsabilités sociales et morales à l'égard de la société. Les principes éthiques de l'IA dérivés de ces lignes directrices sont : « *la contribution à l'humanité, le respect des lois et des règlements, le respect de la vie privée d'autrui, l'équité, la sécurité, l'intégrité, l'obligation de rendre compte et la responsabilité sociale, la communication avec la société et le développement personnel ainsi que le respect des directives éthiques par l'IA* » (JSAI, 2017).

Les Principes éthiques de l'IA dans le Royaume Uni

Au Royaume-Uni, le comité spécial de l'IA de la Chambre des Lords a publié en 2017, un rapport pouvant être adopté à l'échelle nationale et internationale ((Hall et al., 2017)). Ce rapport souligne la nécessité d'encadrer la technologie de l'IA d'une manière éthique, légale et sociale afin

qu'elle soit toujours bénéfique et au profit de la société. Le gouvernement britannique en 2017 a publié un rapport pour recommander plusieurs mesures concernant l'impact de l'IA sur l'économie (Wendy, 2017)³.

Lloyd en 2018, président de comité de la chambre de Lords, affirme que ces principes éthiques sont liés au « *développement pour le bien commun, aux principes d'intelligibilité, l'équité, au droit de conserver ses données, l'interdiction de donner à l'IA le pouvoir autonome de blesser les être humain, l'interdiction de donner à l'IA le pouvoir autonome de détruire ou de tromper les êtres humains (L'interdiction de l'innocuité)* ». (Lloyd Liguéri, 2018)

Les GAFAM et Les principes éthiques de l'IA

Les GAFAM (Google, Apple, Facebook, Amazon et Microsoft) sont considérés comme les leaders de la révolution numérique (Sushcheva et Fontanel, 2018), notamment en raison de leur capacité à innover et à développer des technologies émergentes telles que l'IA (Fontanel et Sushcheva, 2019). Cependant, leur succès est également accompagné de préoccupations concernant les implications éthiques de leurs activités.

En termes de réglementation, les organismes qui constituent les GAFAM évoquent souvent des principes éthiques de l'IA d'une manière individuelle afin d'encadrer leurs activités. Ces principes comprennent notamment « *le respect de la vie privée, le respect des valeurs humaines, la surveillance et le contrôle humain, la sécurité, la fiabilité et la responsabilité, la justice, la solidarité, l'innovation et l'innocuité* ».

En effet, le respect de la vie privée est un principe éthique important pour les GAFAM. Les entreprises doivent respecter les normes de confidentialité et être transparentes quant à la

³ Hall, W., & Pentsenti, J. (2017). Growing the Artificial Intelligence Industry in the UK. London, UK: UK Government.

Cette référence correspond au rapport intitulé "Growing the Artificial Intelligence Industry in the UK" rédigé par Dame Wendy Hall et Jérôme Pesenti

en 2017. Le rapport examine les mesures recommandées pour stimuler l'industrie de l'intelligence artificielle au Royaume-Uni. Il a été publié par le gouvernement britannique.

manière dont elles utilisent les données personnelles des utilisateurs. Par exemple, Google a été critiqué pour ses pratiques de collecte de données et pour la manière dont il utilise ces données pour la publicité ciblée (Electronic Frontier Foundation, 2019).

Le respect des valeurs humaines est également crucial pour les GAFAM. Les entreprises doivent garantir que leurs technologies ne violent pas les droits de l'homme ou ne contribuent pas à des pratiques discriminatoires. Par exemple, Microsoft a développé une politique de responsabilité sociétale de l'entreprise pour guider ses activités en matière de droits de l'homme, de durabilité, de diversité et d'inclusion (Microsoft, 2021).

La surveillance et le contrôle humain sont des préoccupations importantes en matière d'IA. Les GAFAM doivent garantir que leurs systèmes d'IA ne prennent pas de décisions qui pourraient nuire aux individus ou à la société dans son ensemble. Par exemple, Amazon a été critiqué pour son système de reconnaissance faciale, « Rekognition », qui a été accusé d'être biaisé et pour sa possible violation de la vie privée (ACLU, 2018).

La sécurité et la fiabilité sont des principes éthiques clés pour les GAFAM. Les entreprises doivent garantir que leurs systèmes d'IA ne soient pas vulnérables aux attaques ou aux piratages, et qu'ils fonctionnent de manière fiable et prévisible. Par exemple, Apple a mis en place des mesures de sécurité robustes pour protéger les données de ses utilisateurs, y compris le chiffrement des données de bout en bout (Apple, 2021).

La responsabilité est un autre principe éthique important pour les GAFAM. Les entreprises doivent assumer la responsabilité de leurs actions et de l'impact de leurs technologies. Par exemple, Facebook a été critiqué pour sa gestion de l'affaire Cambridge Analytica, où les données personnelles de millions d'utilisateurs ont été utilisées sans leur consentement (The Guardian, 2018).

La justice et la solidarité sont des principes éthiques importants pour garantir que les avantages de l'IA sont équitablement répartis. Les GAFAM doivent veiller à ce que leurs technologies ne créent pas de déséquilibres économiques ou sociaux. Par exemple, Google a lancé un programme pour former des personnes à faible revenu à des compétences en technologie de pointe (Google, 2021).

L'innovation et l'innocuité sont des principes éthiques importants pour garantir que les avantages de l'IA sont maximisés tout en minimisant les risques. Les GAFAM doivent continuer à innover tout en veillant à ce que leurs technologies soient sûres et ne causent pas de préjudice. Par exemple, Microsoft a lancé un programme pour développer des technologies d'IA éthiques et responsables (Microsoft, 2021).

En définitive, bien que les GAFAM soient souvent considérés comme les leaders de la révolution numérique, leur succès est étroitement lié à leur capacité à respecter les principes éthiques de l'IA. Les préoccupations concernant la protection de la vie privée, la surveillance et le contrôle humain, la sécurité et la responsabilité doivent être prises en compte pour garantir que ces entreprises continuent de bénéficier de la confiance des consommateurs et des régulateurs.

Synthèse et comparaison des principes éthiques de l'intelligence artificielle

A travers l'étude exploratoire des codes éthiques phares de l'IA réalisée dans la première section de ce travail, nous avons pu extraire 31 principes éthiques relatifs à l'IA et déclarés par les entités objet de notre étude, qui sont : la sécurité, la sûreté, la transparence, la robustesse, la traçabilité, l'explicabilité, l'efficacité, la fiabilité, la responsabilité, la confiance, la collaboration, l'équité et la non-discrimination, l'égalité, la justice, la surveillance et le contrôle humain, les valeurs et principes humains, la diversité, le développement durable, le respect de la dignité et la liberté, la confidentialité et le respect de la vie privée, la prospérité, le bien être, la souveraineté, le respect des valeurs de la nation, la gouvernance,

D'après cette (Figure 2), nous constatons que :

- Les codes éthiques de l'IA liés à la sécurité, la responsabilité, la confidentialité et le respect de la vie privée détiennent la 1^{ère} position avec 91% d'affirmations dans toutes les entités objet de notre étude.
- Les codes éthiques réservés à la transparence, la surveillance et le contrôle humain ainsi que les résultats bénéfiques occupent un pourcentage de 82% d'affirmations.
- La 3^{ème} position, est occupée par les principes liés à l'équité et la non-discrimination, la fiabilité, le développement durable, et les valeurs et principes humains avec 64% d'affirmations.
- Les principes d'égalité, d'innocuité, de l'IA non meurtrière, de sûreté et d'explicabilité sont classés en 4^{ème} position avec un pourcentage de 55% d'affirmations et sont suivis par la traçabilité, la confiance et la diversité avec 45% d'affirmations.

- La collaboration et le respect de la dignité et la liberté prennent la 6^{ème} position avec 37% d'affirmations.
- En dernier lieu, on trouve la légalité, la souveraineté, la gouvernance, le bien être, la justice, la robustesse ainsi que l'apprentissage et l'éducation et l'intérêt pour la société avec un pourcentage de 27% d'affirmations.

Afin de grouper ces principes éthiques explorés dans des catégories éthiques bien définies, nous avons choisi l'Analyse en Correspondance Multiple (ACM) (Desbois, 2008) et (Meyer, 2004), afin de proposer un référentiel catégorisant les principes éthiques de l'usage de l'IA.

Résultats de la synthèse des principes éthiques de l'IA: vers un référentiel commun par catégorie des principes éthiques de l'IA

L'analyse des Correspondances Multiples (ACM) des 31 principes éthiques explorés et mentionnés ci-dessus est faite par l'usage du logiciel SPSS 25.

Le (Tableau 1) ci-dessous, récapitule les résultats obtenus et engendrés par SPSS 25 :

Tableau 1: Récapitulatif des modèles

Catégories	Alpha de Cronbach	Variance expliquée		
		Total (valeur propre)	Inertie	Pourcentage de variance expliquée
1	,867	6,208	,200	20,026
2	,859	5,937	,192	19,153
3	,779	4,061	,131	13,099
4	,739	3,516	,113	11,343
5	,686	2,972	,096	9,586
6	,564	2,202	,071	7,105
Total		24,896	,803	
Moyenne	,784 ^a	4,149	,134	13,385

a. La valeur Alpha de Cronbach moyenne est basée sur la valeur propre moyenne.

En réduisant les 31 principes éthiques de l'IA issus de notre étude exploratoire à travers l'ACM, nous obtenons six dimensions principales qui représentent une forte inertie totale de 80,3 %.

Les six catégories engendrées ont toutes des valeurs propres supérieures à 1 d'où la pertinence des résultats produits (Inertie totale >60 %).

Les résultats engendrés démontrent que les six catégories engendrées après la réduction des données détiennent un Alpha de Cronbach moyen de 0,784 (>0,7) (Tableau 1), ce qui témoigne d'une très bonne fiabilité de ces dimensions selon les instructions de (Nunnally et al, 1994).

Nous avons choisi l'ACM comme méthode d'analyse car c'est une méthode qui fournit une représentation visuelle des relations entre les catégories éthiques ce qui permet d'identifier les regroupements ou les tendances émergentes parmi les codes éthiques.

Nous signalons que c'est une méthode adaptée pour analyser des données catégorielles et identifier des relations entre des variables qualitatives. Car elle permet de synthétiser les informations contenues dans un tableau croisé, où les lignes représentent les individus et les colonnes correspondent aux variables catégorisant les informations sur ces individus. Les 31 principes éthiques de l'IA ont été regroupés en six catégories en se basant sur les contributions factorielles les plus significatives pour chacune de ces catégories, qui sont respectivement :

- Catégorie 1 : regroupe les principes de prospérité, de bien être, des résultats bénéfiques, d'apprentissage, éducation et innovation ainsi que les principes d'innocuité et de l'IA non meurtrière. Cette catégorie détient 20 % de la variance explicative des six catégories extraites que nous avons nommé « Principes pour une IA bénéfique » ;
- Catégorie 2 : Nous l'avons nommé « Principes appliqués à la technologie pour une IA sûre ». Cette catégorie regroupe les codes éthiques liés à la sûreté, la transparence, la robustesse, la traçabilité et l'explicabilité. Cette catégorie détient 19 % de la variance explicative des six catégories extraites ;
- Catégorie 3 : composée de la gouvernance de l'IA, la fiabilité, la responsabilité, la confiance, l'efficacité et la surveillance et le contrôle humain. Cette catégorie nous l'avons nommé « Principes pour une IA fiable et responsable », et détient un pourcentage de 13% d'inertie ;
- Catégorie 4 : Nous l'avons nommé « Principes pour une l'IA sociétale ». Elle est composée de la collaboration, la diversité, le développement durable et l'intérêt pour la société. Cette catégorie détient 11% de la

variance explicative des six catégories extraites ;

- Catégorie 5 : Nous l'avons nommé « Principes pour une IA citoyenne » car elle se compose de concepts comme la souveraineté, le respect des valeurs de la nation, la légalité et la justice. Cette catégorie complète l'information concernant les quatre premières catégories par un taux de 10% d'inertie ;
- La catégorie 6 : regroupe les « Principes pour une IA sécurisée et respectant les droits humains », dont la sécurité, l'égalité, l'équité et la non-discrimination, la confidentialité et le respect de la vie privée, les valeurs et les principes humains ainsi que le respect de la dignité et la liberté. Cette dimension détient 7% d'inertie.

D'après les résultats obtenus, nous avons élaboré un cadre de référence pour les principes éthiques universels de l'IA, qui est composé de six catégories présentées dans le (tableau 2) ci-dessous.

Tableau 2: Référentiel proposé pour des principes éthiques universels de l'IA

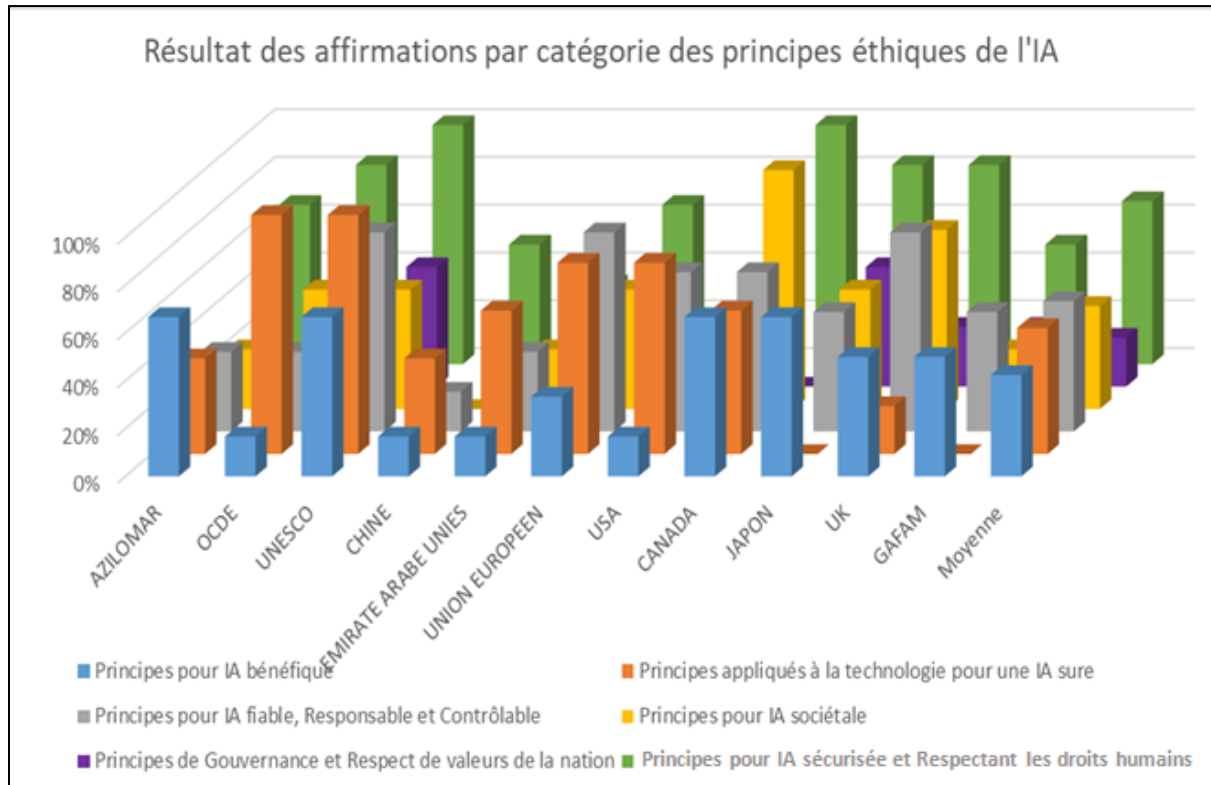
Principes pour une IA bénéfique	Principes appliqués à la technologie pour une IA sûre	Principes pour une IA Fiable, Responsable	Principes pour une IA Sociétale	Principe pour une IA Citoyenne	Principes pour une IA sécurisée et Respectant les droits humains
Résultats bénéfiques	Sûreté	Efficacité	Collaboration	Souveraineté	Sécurité
Prosperité	Transparence	Fiabilité	Diversité	Respect des valeurs de la nation	Egalité
Innocuité	Robustesse	Confiance	Développement durable	Légalité	Respect de la dignité et liberté
IA non meurtrière	Traçabilité	Responsabilité	Intérêt pour la société (Solidarité)	Justice	Confidentialité et Respect de la vie privée
Bien être	Explicabilité	Surveillance et Contrôle humain			Respect des valeurs liées à l'humanité de l'autre.
Apprentissage, Education, Innovation		Gouvernance			Equité, et non-discrimination

Analyse des Résultats

Après avoir réalisé une étude comparative des codes éthiques de l'IA de onze entités internationales et arriver à les regrouper en six catégories essentielles, nous avons pu imaginer comme présenté dans le (Tableau 2) ce que pourrait être un référentiel

commun de principes éthiques de l'IA décliné en catégories bien définies. La (Figure 3) quant à elle, propose une synthèse de l'ensemble des principes éthiques de l'IA objet de notre étude par le biais des pourcentages d'affirmations de chaque catégorie de principes éthiques.

Tableau 3: Synthèse des principes éthiques de l'IA des entités étudiées



D'après la (Figure 3), relative à la moyenne des affirmations par catégorie des principes éthiques de l'IA, nous constatons que :

- Les « Principes pour une IA sécurisée et respectant les droits humains » présentent la catégorie la plus utilisée dans toutes les entités étudiées avec une moyenne de 68% des affirmations. En effet une IA sécurisée et respectant la dignité et la vie privée représente la première préoccupation éthique des utilisateurs de cette technologie (Mittelstadt et Floridi, 2016) (Brundage et al., 2018) (Azencott, 2018). Ceci exige le développement des plateformes sécurisées pour
- les traitements des données personnelles respectant le règlement général sur la protection des données (Tambou, 2016) (Gruson, 2018).
- Les « Principes appliqués à la technologie pour une IA sûre » détiennent la deuxième position en termes d'utilisation dans toutes les entités étudiées avec un pourcentage de 55%. Dans ce contexte, plusieurs travaux de recherche soulignent la nécessité de la sûreté et la robustesse de cette technologie dans le but d'améliorer son usage et éviter toute utilisation malveillante et non éthique (Brundage, et al., 2018) (Voarino, 2020) (Meiller, 2017).

- La classe des « Principes pour une IA fiable, responsable et contrôlable » présente un pourcentage de 53% d'utilisation. Cette classe occupe la troisième position des principes éthiques universels de l'IA. Dans ce sens, La réflexion sur l'intelligence artificielle responsable doit être planifiée afin de prendre la meilleure décision (Zouinar, 2020). Selon (Klein et al., 2004) pour qu'un système « intelligent » soit véritablement collaboratif, il doit se faire guider dans les décisions et actions selon des concepts de régulation et de contrôle.
- En 4ème position, on trouve les « Principes une pour IA sociétale » avec 43% d'affirmations des entités explorées. Dans cette perspective, (Gruson, 2018) précise que l'usage de l'IA doit appliquer la notion de responsabilité sociétale et s'engager à des mesures d'accompagnement des effets sociaux (Gruson, 2018). De même, (Voarino, 2020) souligne l'importance du respect des intérêts collectifs et individuels par l'IA dans le but d'encourager la mutualisation et la collaboration pour son usage pour arriver à des objectifs communs.
- Les « Principes pour une IA bénéfique » présentent 42% d'affirmations. Cette catégorie mène une réflexion sur l'IA non nuisible et bénéfique avec un usage ayant une approche humaniste assurant ainsi le bien-être des humains et leur accès égal à l'apprentissage et au savoir (UNESCO, 2019). **Plusieurs recherches soulignent que l'IA peut changer le monde vers le meilleur ou le pire selon la façon dont elle est utilisée (Ayoh,2021).**

Pour cette raison, le développement des algorithmes de cette technologie doit assurer le contrôle des données et algorithmes pour que ces derniers ne soient pas nuisibles ou meurtriers afin de pouvoir auditer les

programmes de l'IA pour éviter son innocuité (Ayoh,2021).

- *Viennent à la fin, les « Principes de gouvernance et respect des valeurs de la nation » avec seulement 20% d'affirmations.*

Les innovations, liées au développement et l'utilisation de l'intelligence artificielle, ont soulevé des nouvelles inquiétudes dans le monde entier. Le principe de la gouvernance, demeure primordial pour garantir la responsabilité éthique des décisions prises par les systèmes d'IA (Marques, 2022). La gouvernance de l'IA s'inscrit donc dans une continuité des questionnements éthiques posés par la régulation d'Internet (Berleur et al., 2002).

La gouvernance englobe deux aspects importants : d'une part, assurer un fonctionnement efficace des institutions, et d'autre part, veiller à la qualité des actions de ces institutions dans l'espace social, comme souligné par (Pitseys, 2010). Elle implique une réflexion approfondie sur la manière d'impliquer davantage les acteurs dans le processus décisionnel, tout en établissant des critères clairs pour garantir le bon fonctionnement de la société⁴.

La gouvernance éthique de l'IA est un élément crucial pour s'assurer que les systèmes d'IA sont développés et utilisés de manière responsable et éthique, en accordant une attention particulière aux impacts sur les personnes et les sociétés. Pour cela, nous évoquons le concept de la bonne gouvernance qui se réfère au processus de la prise de décision et de réglementation concernant les questions d'intérêt commun⁵.

Le Conseil des Droits de l'Homme⁶, affirme qu'il existe cinq piliers fondamentaux sur lesquels repose une gouvernance efficace : la transparence, la responsabilité, l'obligation de rendre compte, la participation et la capacité de répondre aux besoins de la population. Autrement dit, pour garantir une

⁴ World Bank, Development in practice: Governance – the World's Bank experience, Report n° 13134, 1994, p. vii.

⁵ ibid

⁶ Bureau du Haut-Commissariat aux droits de l'homme

bonne gouvernance, il est essentiel de favoriser la transparence dans les actions entreprises, d'assumer la responsabilité de ses actes, de rendre compte de ses décisions, de promouvoir la participation des acteurs concernés et d'être en mesure de répondre aux besoins de la population.

Dans cette optique une gouvernance éthique de l'IA peut aider à :

- Prévenir les biais et la discrimination dans les décisions prises par les systèmes d'IA ;
- Garantir la transparence et la responsabilité de leur utilisation ;
- Protéger la vie privée des individus ;
- Etablir la confiance du public dans l'IA ;
- Assurer une adoption responsable et durable de cette technologie émergente⁷.

La gouvernance de l'IA peut se faire à plusieurs niveaux, notamment au niveau de l'entreprise, de l'industrie et du gouvernement.

- Au niveau de l'entreprise, la gouvernance de l'IA peut inclure la mise en place de politiques internes pour assurer la responsabilité et la transparence de l'utilisation de l'IA.
- Au niveau de l'industrie, des normes et des réglementations peuvent être mises en place pour garantir que les entreprises travaillant dans le domaine de l'IA respectent des normes éthiques minimales.
- Au niveau du gouvernement, les politiques publiques peuvent être élaborées pour réglementer l'utilisation de l'IA dans des secteurs sensibles, tels que la santé, la sécurité et le respect des valeurs de la nation.

CONCLUSION

Bien que l'IA joue un rôle croissant dans le futur de l'humanité, il est incontestable que son utilisation

abusive peut être préjudiciable à la société et entraîner des conflits avec l'intérêt public.

Dans ce sens, nous avons réalisé cette étude exploratoire pour extraire les codes éthiques de l'usage de l'IA de 11 entités internationales et les synthétiser à travers l'analyse des Correspondances Multiples correspondante à l'analyse factorielle.

Les résultats ont engendré six catégories principales des codes éthiques de l'IA, catégories qui nous ont permis par la suite d'imaginer ce que pourrait être un référentiel commun par catégorie des principes éthiques de l'IA pour un usage bénéfique, sûr et confiant de cette technologie

L'un des enjeux majeurs de l'intelligence artificielle (IA) est la nécessité de développer une approche éthique prenant en compte les valeurs et les normes sociétales. Cela implique de s'assurer que les décisions prises par les systèmes d'IA sont justes, transparentes et responsables. Pour atteindre cet objectif, les concepteurs et les utilisateurs de l'IA doivent parler un langage commun pour discuter des implications éthiques de la technologie.

En effet, l'utilisation de l'IA soulève de nombreuses questions éthiques complexes. Par exemple, un système d'IA peut-il être tenu responsable de ses actions ? Comment s'assurer que les décisions prises par les systèmes d'IA sont justes et non discriminatoires ? Comment assurer la confidentialité et la sécurité des données collectées par les systèmes d'IA ?

Pour répondre à ces questions, les acteurs impliqués dans le développement et l'utilisation de l'IA doivent pouvoir discuter de manière productive et se comprendre. Il s'agit de développer un langage commun permettant de décrire différents concepts et problématiques liés à l'IA.

Cependant, développer un tel langage universel n'est pas facile. Les concepts et les problèmes de l'intelligence artificielle sont complexes et

⁷ Bureau du Haut-commissariat aux droits de l'homme

multidimensionnels, et il est difficile de les exprimer de manière simple et claire. De plus, l'éthique et les valeurs varient d'une culture à l'autre, ce qui rend encore plus difficile le développement d'un langage commun.

Toutefois, il est essentiel de travailler à cet objectif, car cela établira un terrain d'entente pour les discussions sur l'éthique de l'IA. Il favorisera également la collaboration entre les différentes parties prenantes, y compris les développeurs d'IA, les chercheurs en éthique, les décideurs politiques et les utilisateurs.

En conclusion, la normalisation de l'IA en tant qu'approche éthique nécessite le développement d'un langage commun permettant aux parties prenantes de discuter des implications éthiques de la technologie. Bien que cela puisse être un défi, il est impératif de veiller à ce que les décisions prises par les systèmes d'IA soient justes, transparentes et responsables.

BIBLIOGRAPHIE

- ACLU. (2018). Amazon's Face Recognition Falsely Matched 28 Members of Congress with Mugshots. Consulté depuis <https://www.aclu.org/blog/privacy-technology/surveillance-technologies/amazons-face-recognition-falsely-matched-28>.
- Apple. (2021). Privacy - Encryption and Data Protection. <https://www.apple.com/privacy/encryption/>
- Ayoh, C. (2021). Les Legs de partialité: les enjeux éthiques dans l'intelligence artificielle.
- Azencott, C. A. (2018). Machine learning and genomics: precision medicine versus patient privacy. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 376(2128), 20170350.
- Belouali, S., Belouali, A., Saber, M., Jaafar, K., & Belkasmi, M. G. (2019). Ethics of AI or Ethical AI, Topical Point of View. In International conference on smart Information & communication Technologies (pp. 535-540). Springer, Cham.
- Belouali, S., Demazeau, Y., Bouchentouf, T., & Brahmi, H. (2022). Interrogation de l'universalité d'une éthique pour l'usage de l'IA au travers d'une lecture du code publié par Dubaï. *Ethique et Numérique*, 1(1), 1-17.
- Berleur, J. & Pouillet, Y. (2002). Réguler Internet. *Études*, 397, 463-475. <https://doi.org/10.3917/etu.975.0463>.
- Besse, P., Castets-Renard, C., Garivier, A., & Loubes, J. M. (2018). L'IA du Quotidien peut-elle être Éthique ? Loyauté des Algorithmes d'Apprentissage Automatique [Can Everyday AI be Ethical? Machine Learning Algorithm Fairness]. doi:10.13140/RG.2.2.31345.94566
- Bietti, E. (2020, January). From ethics washing to ethics bashing: A view on tech ethics from within moral philosophy. In *Proceedings of the 2020 conference on fairness, accountability, and transparency* (pp. 210-219). <https://doi.org/10.1145/3351095.3372860>
- Brundage, M., Avin, S., Clark, J., Toner, H., Eckersley, P., Garfinkel, B., ... & Amodei, D. (2018). The malicious use of artificial intelligence: Forecasting, prevention, and mitigation. *arXiv preprint arXiv:1802.07228*.
- Commission européenne, Direction générale des réseaux de communication, du contenu et des technologies, Lignes directrices en matière d'éthique pour une IA digne de confiance –, Publications Office, 2019, consulté depuis : <https://data.europa.eu/doi/10.2759/74304>
- Conference, Beneficial AI. (2017). Asilomar AI Principles. Récupéré de <https://futureoflife.org/bai-2017/>
- Desbois, D. (2008). L'analyse des correspondances multiples « à la hollandaise »: introduction à

- l'analyse d'homogénéité. *Revue Modulad*, 38(194-244).
- Digital Dubai. (n.d.). AI principles & ethics. Récupéré depuis <https://www.digitaldubai.ae/initiatives/ai-principles-ethics>
- Dilhac, M. A., Abrassart, C., & Voarino, N. (2018). Rapport de la déclaration de Montréal pour un développement responsable de l'intelligence artificielle.
- Fontanel, J., & Sushcheva, N. (2019). La puissance des GAFAM : réalités, apports et dangers. *Annuaire français de relations internationales*, XX. Récupéré depuis <https://hal.archives-ouvertes.fr/hal-02196915>
- Fontanel, J., & Sushcheva, N. (2019). La puissance des GAFAM. *Annuaire français de relations internationales*, 20.
- Groupe d'Experts Ad Hoc (GEAH) pour l'élaboration d'un avant-projet et de recommandation sur l'éthique de l'intelligence artificielle. (2020). Document final avant-projet de recommandation sur l'éthique de l'intelligence artificielle (SHS/BIO/AHEG-AI/2020/4 REV.2). Paris, septembre 2020. Récupéré depuis https://unesdoc.unesco.org/ark:/48223/pf0000373434_fr
- Gruson, D. (2018). L'intelligence artificielle en santé. *Inf Psychiatr*, 94(1), pp. 47-52.
- Hall, W., & Pesenti, J. (2017). Growing the artificial intelligence industry in the UK. Department for Digital, Culture, Media & Sport and Department for Business, Energy & Industrial Strategy. Part of the Industrial Strategy UK and the Commonwealth.
- IA : vers une meilleure régulation aux Etats-Unis - ZDNet. Consulté en juin 2021 depuis (<https://www.zdnet.fr/actualites/ia-vers-une-meilleure-regulation-aux-etats-unis-39897031.htm>).
- Klien, G., Woods, D. D., Bradshaw, J. M., Hoffman, R. R., & Feltovich, P. J. (2004). Ten challenges for making automation a "team player" in joint human-agent activity. *IEEE Intelligent Systems*, 19(6), 91-95.
- Laney, Zhang. 2019. « China: AI Governance Principles Released ». Consulté en janvier 2020 depuis : <https://www.loc.gov/law/foreign-news/article/china-ai-governance-principles-released/>.
- Lausson, Julien. 2019. « 7 principes pour une IA éthique vont être mis à l'épreuve par l'Europe ». *Numerama*. Consulté en novembre 2019 depuis : <https://www.numerama.com/politique/479517-7-principes-pour-une-ia-ethique-vont-etre-mis-a-lepreuve-par-leurope.html>.
- Ligueri-Richard, L. (2018). Le développement de l'intelligence artificielle au Royaume-Uni "IA éthique". Consulté depuis : <https://www.irsem.fr/document/file/n176b>
- Marques, J. (2022). Le principe de justice dans la gouvernance de l'Intelligence artificielle au prisme du genre, de classe et de race. *Instruments, définitions et limites*. Terminal. Technologie de l'information, culture & société, pp.132-133.
- Meiller, Y. (2017). Intelligence artificielle, sécurité et sûreté. *Sécurité et stratégie*, 28(4), pp.75-84.
- Meyer, N., Ferlicot, S., Vieillefond, A., Peyromaure, M., & Vielh, P. (2004). Intérêt de l'analyse factorielle des correspondances multiples en histopathologie. In *Annales de Pathologie*, Vol. 24, N° 2, pp. 149-160. Elsevier Masson.
- Microsoft. (2021). AI for Good. [online]. Available : <https://www.microsoft.com/en-us/>.
- Mittelstadt, B. D., Allo, P., Taddeo, M., Wachter, S., & Floridi, L. (2016). The ethics of algorithms: mapping the debate. *Big Data & Society*, 3(2), 2053951716679679.

- Noiseau, P., Lanteigne, C., Echaiz, L. F., Salazar, F. G. G., Mai, V., Dilhac, M. A., & Mörch, C. M. (2021). Le dialogue inclusif sur l'éthique de l'IA: délibération en ligne citoyenne et internationale pour l'UNESCO. *Communication, technologies et développement*, (10).
- Nunnally, J., & Bernstein, I. (1994). *Psychometric Theory* 3rd edition, MacGraw-Hill, New York. ohchr-archives@un.org
- Organisation pour la Coopération et le développement économique. (2019). Quarante-deux pays adoptent les nouveaux Principes de l'OCDE sur l'intelligence artificielle. Consulté en janvier 2021 depuis <http://www.oecd.org/fr/numerique/quarante-deux-pays-adoptent-les-nouveaux-principes-de-l-ocde-sur-lintelligence-artificielle.htm>
- Pitseys, J. (2010). Le concept de gouvernance. *Revue interdisciplinaire d'études juridiques*, 65, 207-228. <https://doi.org/10.3917/riej.065.0207>
- Tambou, O. (2016). L'Introduction de la certification dans le règlement général de la protection des données personnelles: quelle valeur ajoutée? (The introduction of the Certification in the EU General Data Protection Regulation: What Added Value?). *The introduction of the Certification in the EU General Data Protection Regulation: What Added Value*, pp.51-54.
- The Guardian. (2018, March 17). Revealed: 50 million Facebook profiles harvested for Cambridge Analytica in major data breach. <https://www.theguardian.com/news/2018/mar/17/cambridge-analytica-facebook-influence-us-election>
- The Japanese Society for Artificial Intelligence: Rapport. "Les lignes directrices éthiques de la Société japonaise d'intelligence artificielle". Consulté en août 2020 depuis <http://www.ai-elsi.org/wp-content/uploads/2017/05/JSAI-Ethical-Guidelines-1.pdf> (2017).
- Thomasset, A. (2020). Éditorial. « Quelle éthique pour l'intelligence ? artificielle ? ». *Revue d'éthique et de théologie morale*, (3), pp. 7-9.
- Trump, D. J. (2019). Executive Order, E. (2019). 13859 of February 11, 2019 on Maintaining American Leadership in Artificial Intelligence. *Federal Register*, 84(31), pp.3967-3972.
- UNESCO. (2019). Actes de la Conférence générale, 40e session, Paris, 12 novembre-27 novembre 2019, volume 1: Résolutions. Récupéré depuis https://unesdoc.unesco.org/ark:/pf0000372579_fre
- Voarino, N. (2020). Systèmes d'intelligence artificielle et santé: les enjeux d'une innovation responsable.
- Wu, W., Huang, T., & Gong, K. (2020). Ethical principles and governance technology development of AI in China. *Engineering*, 6(3), 302-309, Consultée en octobre 2020. depuis <https://doi.org/10.1016/j.eng.2019.12.015>.
- Zeng, Y., Lu, E., & Huangfu, C. (2018). Linking artificial intelligence principles. arXiv preprint arXiv : 1812.04814 Consultée en octobre 2020. Depuis <https://arxiv.org/abs/1812.04814> (2018).
- Zhang, L. (2019). China: AI governance principles released. In *Library of Congress*. Consulté en Janvier 2020 depuis <https://www.loc.gov/item/global-legal-monitor/2019-09-09/china-ai-governance-principles-released/>
- Zouinar, M. (2020). Évolutions de l'intelligence artificielle : quels enjeux pour l'activité humaine et la relation Humain-Machine au travail? *Activités*, (17-1), Consulté en octobre 2020. Depuis <https://doi:10.4000/activites.4941>.