Les discours critiques de l'intelligence artificielle dans Wikipédia

Critical discourse on artificial intelligence in Wikipedia

Aymeric Bouchereau[1], Antonin Segault[2]

[1] Univ Paris Est Créteil, LAB'URBA, F-94010 Créteil, France, aymeric.bouchereau@u-pec.fr
[2] Laboratoire Dicen, Université Paris Nanterre, asegault@parisnanterre.fr

RÉSUMÉ. Cet article examine la manière dont les discours critiques sur l'Intelligence artificielle (IA) sont représentés dans Wikipédia, en se concentrant sur les enjeux majeurs identifiés dans la littérature. À partir d'un corpus d'articles francophones et anglophones de l'encyclopédie, l'analyse porte sur la couverture des enjeux éthiques comme la protection de la vie privée et les biais algorithmiques. L'étude met en lumière les dynamiques discursives à l'œuvre dans Wikipédia et les biais éditoriaux influençant le traitement des controverses liées à l'IA. Les résultats offrent une meilleure compréhension des liens entre médias, science et encyclopédie participative sur des questions liées à l'IA.

ABSTRACT. This article studies how critical discourses on Artificial Intelligence (AI) are represented in Wikipedia, focusing on the major issues identified in the literature. Using a corpus of French- and English-language Wikipedia articles, the analysis focuses on the coverage of ethical issues such as privacy and algorithmic bias. The study sheds light on the discursive dynamics at work in Wikipedia and the editorial biases influencing the treatment of AI-related controversies. The results provide a better understanding of the links between the media, science and the participatory encyclopaedia on issues related to AI.

MOTS-CLÉS. Intelligence artificielle, enjeux éthiques, Wikipédia, discours critiques, controverses médiatiques
KEYWORDS. Artificial intelligence, ethical issues, Wikipedia, critical discourse, media controversies

1. Introduction

L'Intelligence artificielle (IA) connaît depuis la fin des années 2010 une évolution rapide suscitant un flot de discours allant d'annonces enthousiastes de nouvelles applications aux craintes d'un déclassement de l'humain face à l'IA [AND 23]. Une partie des discours sur les systèmes d'IA critique leurs développements actuels en pointant les risques et les enjeux éthiques qui se posent de manière concrète et immédiate aux citoyens, à l'exemple de la protection de la vie privée ou des biais algorithmiques [FLO 23]. Cependant, la prise de conscience nécessaire de ces risques par les citoyens est rendue complexe en raison d'une publicisation controversée de l'IA. En effet, les médias font une couverture intense, mais discutable de l'IA mettant davantage en avant les aspects sensationnels et prospectivistes que les discours critiques et les enjeux [BEND 25 ; CRE 22].

Wikipédia constitue un riche terrain pour analyser le déploiement des controverses scientifiques dans l'espace médiatique [BENJ 18 ; BENJ 23]. En raison de son adossement à des sources secondaires, l'encyclopédie se fait souvent le reflet de l'attention médiatique portée à un sujet donné [DOU 20], mais avec ses propres biais notamment liés à la démographie des contributeurs [MCD 21]. Dans nos travaux de recherche, nous nous intéressons à Wikipédia en tant que source d'informations publique sur l'IA, ses risques et ses enjeux. Dans cette optique, cet article porte sur la manière dont les discours critiques de l'IA sont intégrés dans l'encyclopédie Wikipédia, en se concentrant sur les enjeux majeurs identifiés dans la littérature scientifique. À partir d'un travail empirique sur un corpus d'articles de l'encyclopédie, il s'agit d'examiner les dynamiques discursives à l'œuvre au sein de Wikipédia sur la question des enjeux éthiques de l'IA.

L'article s'organise autour de trois axes principaux. Le premier revient sur l'évolution scientifique et médiatique de l'IA et de ses enjeux, ainsi que l'étude des discours en circulation dans l'espace public à travers Wikipédia. Le second axe présente la méthodologie pour constituer un corpus et analyser le

© 2026 ISTE OpenScience – Published by ISTE Ltd. London, UK – openscience.fr

traitement des enjeux éthiques de l'IA dans les articles de l'encyclopédie. Le troisième axe rapporte puis discute les résultats en termes de distribution de ces enjeux dans le corpus, leur fréquence et les références mobilisées par les contributeurs. Nous concluons par les limites et les apports de ce travail à la compréhension des dynamiques éditoriales et discursives à l'œuvre dans Wikipédia sur des questions liées à l'IA.

2. IA : du projet scientifique aux controverses

L'IA connaît une période faste depuis les années 2010, tant sur le plan scientifique que médiatique. Le déploiement massif de systèmes d'IA dans tous les domaines de la société suscite d'abondantes discussions dans l'espace public oscillant entre enthousiasmes, peurs et critiques.

2.1. L'intelligence artificielle et ses enjeux éthiques

L'IA est un projet technologique et scientifique ancré dans un contexte philosophique et historique riche [AND 23]. Initiée sous ce terme en 1955, l'IA désigne un champ de recherche visant à élaborer des « techniques permettant à des machines d'accomplir des tâches et de résoudre des problèmes normalement réservés aux humains » [LEC 15]. Avant les années 1990, ces techniques s'inscrivent principalement dans l'approche dite symbolique : les connaissances sont formalisées par des spécialistes avant d'être manipulées par des machines, à l'instar des systèmes experts [RUS 10]. Puis, les techniques relevant de l'approche connexionniste, fondées sur des réseaux de neurones artificiels, deviennent progressivement majoritaires jusqu'à l'emballement dû à l'apprentissage profond au début des années 2010 [CAR 18]. L'apprentissage profond consiste à inférer les connaissances sur la base de motifs statistiques identifiés dans de grands volumes de données [LEC 19]. Son efficacité rapproche les capacités des systèmes d'IA de celles des humains en vision artificielle, traitement du langage naturel et création de contenus.

À partir de 2015, le succès de l'apprentissage profond entraîne une multiplication des systèmes d'IA, surtout issus d'acteurs privés, ainsi qu'un engouement discursif où s'entremêlent des intérêts scientifiques, économiques, médiatiques et politiques [BEL 23]. L'IA est discutée par une variété d'acteurs portant des regards différents et parfois antagonistes [VAY 20 ; BEL 23] : promesses d'applications, catastrophisme, alertes sur les dangers et les risques, appels à régulation, etc. Les médias font une couverture intense et controversée de l'IA et de ses discours en mettant davantage l'accent sur des aspects sensationnels et spectaculaires [GOU 24 ; CRE 22 ; ROM 22], avec une fluctuation de l'attention portée aux risques [NGU 22]. Plusieurs études montrent notamment que la presse écrite tend à mettre en scène une confrontation de l'humain face à l'IA en puisant dans des imaginaires de science-fiction [BUN 22 ; CHU 19 ; WAN 23]. Les médias sont « prompts à activer un stock de représentations déjà disponibles et riche de nombreuses ressources culturelles et fictionnelles pour relayer la prophétie d'une guerre entre les humains et les robots » [CRE 22, p. 161]. Ces discours, qu'ils expriment une forme de « *dooming* » (catastrophisme) ou de « *boosting* » (engouement excessif), participent à l'inflation de la part de fantasme lié à l'IA et détournent l'attention des problématiques plus immédiates [BEND 25].

En parallèle de ce régime discursif sensationnel, un régime discursif plus critique de l'IA s'attache à exprimer les risques, expliquer les enjeux et appeler à des régulations, en accord avec l'état actuel du développement des systèmes d'IA [CRE 22 ; MAC 18]. Ces risques et enjeux découlent de limites intrinsèques au fonctionnement des systèmes d'IA fondés sur l'apprentissage profond [AND 23 ; LEC 19] : la nature statistique de cette méthode et sa dépendance aux données peuvent conduire à la reproduction de biais discriminatoires, à la collecte abusive de données, à l'atteinte à la vie privée ou encore à la production de fausses informations. Afin de réduire ces risques, l'éthique est mobilisée et des principes sont formulés dans le but d'encadrer le développement de l'IA [FLO 23 ; JOB 19]. Bien que critiquable en raison de la multiplicité des conceptions de l'éthique de l'IA [ROM 23], l'approche par l'éthique permet de mettre en lumière des enjeux de l'IA en accord avec la réalité sociotechnique de ses usages [FLO 23]. Cependant, les débats sur l'IA éthique tendent à rester confinés aux espaces limités que sont l'arène scientifique et les sphères

© 2026 ISTE OpenScience – Published by ISTE Ltd. London, UK – openscience.fr

politiques et administratives représentées notamment par les commissions, les missions de réflexion, les organes de régulation ou encore les *think tanks* [BEL 23].

Au milieu des controverses et des discours polarisés, la visibilité auprès du grand public des discours critiques sur l'IA pose donc question, d'autant plus dans un contexte où les citoyens sont amenés à interagir étroitement et quotidiennement avec des systèmes d'IA. Les représentations qu'ont les citoyens de l'IA, et donc les informations dont ils disposent sur elle, influencent leurs interactions avec les systèmes intelligents [INI 25 ; GOU 24]. Une faible compréhension de ces systèmes et de leurs potentiels risques tend à inhiber le recul critique et à augmenter la confiance qui leur ait accordée [TUL 25 ; LEE 25]. Il apparaît par conséquent crucial d'examiner les moyens d'information relatifs à l'IA et à ses enjeux, ainsi que leurs répercussions dans les arènes publiques. Cette nécessité est d'autant plus prégnante en 2024, année de constitution du corpus étudié dans cet article, qui voit une accélération sensible de l'essor des systèmes génératifs et des enjeux qui les accompagnent. D'une part, les performances des systèmes d'IA continuent de s'améliorer au-delà de la génération de textes (e.g. génération de vidéos réalistes avec Sora d'OpenAI) et ils sont progressivement intégrés à des outils déjà quotidiennement utilisés tels que les équipements Apple ou le moteur de recherche Google. D'autre part, le Conseil européen adopte en juin 2024 le règlement sur l'IA (*AI Act*) qui constitue un premier cadre important en matière de régulation dans ce domaine avec des implications à l'échelle européenne et internationale.

2.2. Étudier les controverses à travers Wikipédia

Wikipédia est plus qu'une encyclopédie, c'est avant tout un « lieu de production et de transmission de connaissances » [BEA 15]. Inlassablement, des milliers de contributeurs, réguliers ou occasionnels, identifiés ou anonymes, s'y affairent pour écrire, vérifier et actualiser des articles dans plus de trois cents langues. Leur travail est encadré par un large corpus de règles, de normes et de procédures administratives [LAN 15 ; MCD 21] réifiées dans les outils qu'ils utilisent pour produire et gérer les articles (principalement le logiciel MediaWiki). Ce dispositif, unique par son fonctionnement et par son échelle, fait l'objet d'une attention considérable des chercheurs, notamment en sciences humaines et sociales.

Wikipédia s'est notamment avérée être une source précieuse pour étudier les discours qui se déploient dans l'espace public [DOU 20 ; FOR 22]. Dans le cas des controverses socioscientifiques, l'encyclopédie permet de mettre en évidence l'évolution temporelle de certains sujets ou arguments, leur organisation à l'échelle d'un ou plusieurs articles, ainsi que les références mobilisées [BENJ 18 ; BENJ 23]. L'analyse croisée de plusieurs versions linguistiques des mêmes articles permet également de rendre visibles les différences de terminologie et de point de vue, notamment dans le cas d'articles relatifs à des conflits [ROG 12 ; OEB 20]. L'encyclopédie, imprégnée des biais de ses contributeurs et des sources que ces derniers mobilisent [MCD 21], n'offre qu'un reflet imparfait des événements qu'elle relate. Toutefois, la richesse des traces qu'elle enregistre (dans les historiques de version, les pages de discussions, etc.) et leur grande accessibilité (*via* l'interface wiki ou des interfaces de programmation, sans coûts ni entraves contractuelles) constituent une opportunité rare, dans l'écosystème numérique actuel, de retracer la construction et la circulation de discours médiatiques à travers le temps et l'espace.

Nos premiers travaux relatifs aux discours de l'IA dans l'article dédié de Wikipédia ont montré une surreprésentation de sujets sensationnels et controversés, au détriment des applications actuelles et des questionnements éthiques qu'elles suscitent [BOUa 21]. Ces résultats, cohérents avec ceux d'études sur la presse, pourraient s'expliquer par la forte dépendance de Wikipédia à l'égard des sources journalistiques, notamment autour des sujets d'actualité [DOU 20]. Nous avons montré que des articles voisins (Apprentissage machine, Apprentissage profond) n'adoptent pas le même point de vue sur le sujet, mobilisant un vocabulaire et des références (plus ou moins scientifiques ou journalistiques) différents, et ce à la fois dans les versions francophones et anglophones de Wikipédia [BOUb 23]. Nos observations rejoignent en cela celles sur la « bulle médiatique » de l'IA [BEL 23] caractérisée en partie par un mélange des registres scientifiques, médiatiques et politiques, ce qui pourrait brouiller la visibilité des enjeux de l'IA.

© 2026 ISTE OpenScience – Published by ISTE Ltd. London, UK – openscience.fr

3. Méthodologie

Dans la continuité des travaux présentés plus haut, cette étude vise à examiner la manière dont les enjeux éthiques de l'IA sont présentés dans Wikipédia. Nous faisons l'hypothèse que les controverses médiatiques autour de l'IA se manifestent par une présence hétérogène de discours critiques dans les articles de l'encyclopédie traitant de l'IA. En premier lieu, nous supposons que les articles relatifs à l'IA se distinguent par le type de discours adopté, construisant soit une représentation générale et peu critique de l'IA, soit une représentation centrée sur les enjeux éthiques. Cette distinction irait de pair avec un traitement inégal des différents enjeux éthiques ; certains enjeux seraient plus visibles que d'autres. Par ailleurs, nous présumons que ces discours peuvent être reliés à l'usage de différents types de sources parmi les références des articles. Afin d'éprouver ces hypothèses, il nous a fallu construire une typologie des enjeux éthiques de l'IA à examiner, délimiter un corpus multilingue d'articles Wikipédia, puis développer une méthode de traçage des occurrences des enjeux dans les articles.

3.1. Typologie des enjeux éthiques

Le développement de l'IA s'est accompagné d'une pléthore de discours relatifs aux enjeux et aux questions éthiques. Face à ce foisonnement, divers acteurs académiques ou institutionnels ont tenté de formaliser des typologies des risques ou des enjeux de l'IA. En raison de la diversité de ces acteurs, de leurs points de vue et de leur approche du sujet, les listes résultantes diffèrent fortement, tant par leur forme (nombre d'*items*, hiérarchisation) que par leur fond (éléments inclus ou non, organisation). Sans prétention à l'exhaustivité, nous allons présenter quelques-unes de ces typologies qui ont éclairé notre démarche :

– En 2017, un rapport de la Commission nationale de l'informatique et des libertés sur « les enjeux éthiques des algorithmes et de l'intelligence artificielle »[1] relevait six problématiques : l'autonomie humaine au défi de l'autonomie des machines ; biais, discrimination et exclusion ; fragmentation algorithmique : la personnalisation contre les logiques collectives ; entre limitation des mégafichiers et développement de l'IA : un équilibre à réinventer ; qualité, quantité, pertinence : l'enjeu des données fournies à l'IA ; l'identité humaine au défi de l'IA.

– En 2018, dans un article intitulé « Le nouvel âge de l'intelligence artificielle : une synthèse des enjeux éthiques » [MAC 18], Maclure et Saint-Pierre proposent, en réaction à des discours prospectivistes sur la menace que constituerait une super IA, une liste de problématiques éthiques plus immédiates : la responsabilité des algorithmes ; justice, équité et discrimination ; l'interprétabilité des décisions des systèmes d'IA ; autonomie et protection de la vie privée ; la reconfiguration du marché du travail.

– En 2018, le groupe AI4People, émanation de l'Atomium - European Institute for Science, Media and Democracy, a proposé dans un « Ethical Framework for a Good AI Society » [FLO 23] (repris plus tard par le GEHN IA de la Commission Européenne) une liste de cinq principes éthiques issus d'une revue de la littérature : Beneficence, Non maleficence, Autonomy, Justice, Explicability.

– En 2019, Benjamin Cheatham, Kia Javanmardian et Hamid Samandari proposent, dans l'article « Confronting the risks of artificial intelligence » [CHE 19], de classer les risques associés aux systèmes d'IA en fonction des types d'acteurs susceptibles d'être affectés : pour les individus, la sécurité physique, la vie privée et la réputation, la sécurité numérique, la santé économique, l'équité et un traitement juste ; pour les organisations, la performance économique, la performance non économique, la conformité juridique, l'intégrité réputationnelle ; pour la société, la sécurité de l'état, la stabilité économique, la stabilité politique, l'intégrité des infrastructures.

Il apparaît que certaines approches, telles que celle d'AI4People, privilégient des principes assez généraux et donc difficiles à observer dans les discours, ce qui s'avère rédhibitoire dans le cadre de cette étude. En revanche, le travail de Cheatham et ses collègues a permis de clarifier le positionnement de cette étude

[1] https://www.cnil.fr/fr/comment-permettre-lhomme-de-garder-la-main-rapport-sur-les-enjeux-ethiques-des-algorithmes-et-de

© 2026 ISTE OpenScience – Published by ISTE Ltd. London, UK – openscience.fr

autour des enjeux pour les individus. Enfin, l'approche de Maclure et Saint-Pierre, en accord avec les études citées dans la section précédente, invite à se concentrer sur des enjeux immédiatement concrets. Nous avons ainsi construit la typologie suivante, basée sur celle de Maclure et Saint-Pierre et enrichie de quelques points :

- Vie privée : questions relatives à l'exploitation et à la protection des données personnelles ;
- Interprétabilité (des décisions) : degré d'opacité ou d'explicabilité des décisions des IA ;
- Biais : risque de décisions injustes ou discriminatoires par des IA ;
- Autonomie (humaine) : libertés humaines face aux *nudges* et aux bulles de filtre ;
- Désinformation : problématiques liées à la diffusion et à la détection de contenus falsifiés ;
- Travail (et automatisation) : reconfigurations du monde du travail et risques de chômage ;
- Responsabilité (des décisions) : enjeux d'imputabilité (y compris légale) des décisions.

Cette typologie est assez simple, mais facilement opérationnalisable. Elle présente pour notre étude l'intérêt d'être représentative des enjeux fréquemment cités dans la littérature, de comporter des enjeux suffisamment précis pour être reliés à une problématique concrète rencontrée par les individus non-spécialistes, et qui sont ancrés dans la réalité sociotechnique plutôt que fondés sur des conjectures et de la prospective.

3.2. Corpus d'articles Wikipédia

En août 2024, Wikipédia comptait 63 millions d'articles en 331 langues, avec une distribution très inégale, l'encyclopédie anglophone étant composée de plus de 6 millions d'articles[2]. Afin de limiter les écueils liés à la traduction et à ses ambiguïtés, il a été décidé de ne procéder qu'à une comparaison entre les versions francophones et anglophones. Ce choix résulte de considérations pratiques (les compétences linguistiques des auteurs), mais aussi d'importantes différences observées lors d'études précédentes [BOUb 23] entre les versions françaises et anglaises de la paire d'articles Intelligence artificielle / Artificial intelligence de l'encyclopédie.

Nous avons constitué un corpus d'articles au sein desquels l'IA est citée de manière importante, en supposant que les citoyens souhaitant s'informer sur le sujet s'orienteront en priorité vers ces articles. Pour cela, nous avons repris la méthode et les outils présentés par Benjakob et ses collègues dans leur étude de la couverture des controverses autour de CRISPR [BENJ 23]. L'équipe a mis à disposition un *script* Python, WikiCorpusBuilder, qui identifie tous les articles contenant une certaine chaîne de caractères dans leur titre principal ou dans l'un de leurs intertitres. Le *script* collecte et calcule une série d'indicateurs pour chacun des articles trouvés (notamment sur les références mobilisées).

Le *script*, lancé avec la chaîne de caractères « artificial intelligence », remonte 361 articles sur l'encyclopédie anglophone (résultats actualisés en date du 26 juin 2024). Conçu spécifiquement pour cette version linguistique de l'encyclopédie, il doit être adapté afin de fonctionner pour l'encyclopédie francophone. Sur cette dernière, la chaîne « intelligence artificielle » permet d'identifier 194 articles (même date). Afin de permettre une comparaison fiable entre les deux environnements linguistiques, seuls les articles dont les sujets apparaissent dans les deux listes sont conservés. Un *script* Python *ad-hoc* est développé, utilisant les interfaces de programmation (API) de Wikipédia, afin de vérifier les liens inter-langues entre les articles. Sur cette base, une liste de quarante-trois paires d'articles francophones et anglophones (soit 86 articles au total) est finalement retenue, constituant le corpus de l'étude (Tableau 1).

[2] https://meta.wikimedia.org/wiki/List_of_Wikipedias

© 2026 ISTE OpenScience – Published by ISTE Ltd. London, UK – openscience.fr

Articles anglophones	Articles francophones
A.I. Artificial Intelligence	A.I. Intelligence artificielle
AlphaGo versus Lee Sedol	Match AlphaGo - Lee Sedol
Amy Karle	Amy Karle
Anticipation (artificial intelligence)	Anticipation (intelligence artificielle)
Applications of artificial intelligence	Applications de l'intelligence artificielle
Argumentation theory	Théorie de l'argumentation
Artificial Intelligence Act	Législation sur l'intelligence artificielle
Artificial intelligence and copyright	Intelligence artificielle et droit d'auteur dans l'art
Artificial intelligence arms race	Course aux armements intelligents
Artificial intelligence art	Art créé par intelligence artificielle
Artificial intelligence content detection	Détection de contenu généré par intelligence artificielle
Artificial intelligence in fiction	Intelligence artificielle dans la fiction
Artificial intelligence in healthcare	Intelligence artificielle dans la santé
Artificial intelligence in video games	Intelligence artificielle dans le jeu vidéo
Artificial Intelligence: A Modern Approach	Intelligence artificielle : une approche moderne
Artificial intelligence	Intelligence artificielle
Civic intelligence	Intelligence civique
Cognitive science	Sciences cognitives
Dalle Molle Institute for Artificial Intelligence Research	Institut Dalle Molle d'études sur l'intelligence artificielle
Diagnosis (artificial intelligence)	Diagnostic (intelligence artificielle)
Distributed artificial intelligence	Intelligence artificielle distribuée
Doina Precup	Doina Precup
Ethics of artificial intelligence	Éthique de l'intelligence artificielle
Frame (artificial intelligence)	Cadre (intelligence artificielle)
Friendly artificial intelligence	Intelligence artificielle amicale
Generative artificial intelligence	Intelligence artificielle générative
Geoffrey Hinton	Geoffrey Hinton
German Research Centre for Artificial Intelligence	Centre de recherche allemand pour l'intelligence artificielle
Global Partnership on Artificial Intelligence	Partenariat mondial sur l'intelligence artificielle
Hallucination (artificial intelligence)	Hallucination (intelligence artificielle)
Herbert A. Simon	Herbert Simon
History of artificial intelligence	Histoire de l'intelligence artificielle
Hubert Dreyfus's views on artificial intelligence	Point de vue d'Hubert Dreyfus sur l'intelligence artificielle
Ian Hogarth	Ian Hogarth
Open letter on artificial intelligence (2015)	Lettre ouverte sur l'intelligence artificielle
Paul Jorion	Paul Jorion
Philosophy of artificial intelligence	Philosophie de l'intelligence artificielle
Regulation of algorithms	Régulation des algorithmes
Regulation of artificial intelligence	Réglementation de l'intelligence artificielle
Sam Harris	Sam Harris
Symbolic artificial intelligence	Intelligence artificielle symbolique
Weak artificial intelligence	Intelligence artificielle faible
World Government Summit	Sommet mondial des gouvernements

Tableau 1. *Liste des articles du corpus*

La démarche d'identification des articles résulte en un corpus assez peu exhaustif. En premier lieu, l'usage du script WikiCorpusBuilder implique que seuls les articles dont les titres ou les intertitres contiennent l'expression exacte recherchée (« artificial intelligence » ou « intelligence artificielle ») sont sélectionnés. Cette approche laisse de côté des articles utilisant des termes proches (« IA », « machine intelligence ») ou dans lesquels des termes sont intercalés (« artificial general intelligence »). Elle permet toutefois d'éliminer tous les faux positifs, aucun article du corpus n'étant pas directement lié à la thématique de l'IA. Le

croisement des corpus anglophones et francophones se traduit également par une forte déperdition, de nombreux articles pertinents dans un corpus n'ayant pas de correspondance dans l'autre (parce que l'article équivalent n'existe pas, ou que ses titres ne contiennent pas l'expression recherchée). Là aussi, le rappel a été sacrifié au profit de la précision.

3.3. Détection des occurrences

Nous postulons que l'apparition d'un ou plusieurs des enjeux de notre typologie dans un article de l'encyclopédie signale plausiblement la présence d'un discours critique ou, à tout le moins, d'une mise en avant des risques immédiats posés par les systèmes d'IA. Bien que la mention d'un enjeu éthique puisse s'avérer uniquement cosmétique et ne pas correspondre à une réelle discussion critique, ce postulat nous permet d'avoir une lecture quantitative de la diffusion des discours critiques de l'IA dans les articles de l'encyclopédie. Nous avons sélectionné une série de termes ou d'expressions (dans les deux langues) pouvant être reliés à chacun de ces enjeux. Cette liste a été revue de manière itérative, en supprimant les termes trop généralistes (apparaissant sans lien avec l'enjeu en question) ou trop spécifiques (n'apparaissant pas dans le corpus), aboutissant à la version présentée dans le Tableau 2.

Enjeu	Termes anglais	Termes français
Vie privée	Privacy, personal data, GDPR, private data, data protect	Vie privée, données personnelles, RGPD, données privées, protection des données
Interprétabilité	Interpretability, explainable, black box, opaque	Interprétable, explicable, boite noire, opacité
Biais	Bias, discrimination, unequal, racism, inequality	Biais, discrimination, injustice, racisme, inégalité
Autonomie	Filter bubble, nudge, autonomy, freedom	Bulle de filtre, nudge, autonomie, liberté
Désinformation	Disinform, deep fake, fake news	Désinformation, deep fake, fake news
Travail	Jobs, job market, worker, unemployment	Emplois, marché du travail, travailleur, chômage
Responsabilité	Imputability, accountable, liability, responsibility	Imputable, responsable

Tableau 2. *Liste des termes anglais et français associés à chaque enjeu*

Le contenu des 86 articles du corpus a été téléchargé à travers les API de Wikipédia et manuellement nettoyé pour en retirer les sections de bibliographie ou de liens connexes. Un script Python *ad-hoc* a été employé pour rechercher dans chacun de ces textes les occurrences des termes associés aux enjeux éthiques, préalablement racinisés (la taille du corpus, dépassant les 1,6 millions de mots, faisant obstacle à l'usage de méthodes de détection manuelles, plus nuancées). La longueur des articles (en nombre de caractères) a également été calculée sur la base de ces mêmes fichiers texte.

3.4. Analyse des références

Depuis la fin des années 2000, les contributeurs de Wikipédia sont tenus d'ajouter des références bibliographiques à toute affirmation susceptible d'être contestée [LAN 15]. Des fonctionnalités dédiées ont été ajoutées au logiciel MédiaWiki, tandis que la communauté a défini une hiérarchie plus ou moins explicite des sources (favorisant notamment les publications scientifiques) et des listes de sources à éviter [SAH 16]. Différents travaux ont montré que l'analyse des références employées dans les articles participe à éclairer l'activité des contributeurs [BENJ 22 ; KOP 15 ; YAN 24].

Le *script* WikiCorpusBuilder utilisé pour sélectionner les articles relatifs à l'IA fournit également pour chaque article un « SciScore » [BENJ 23]. Cette métrique, calculée sur la base des références ajoutées par les contributeurs, correspond à la proportion de sources scientifiques (identifiées comme telles par la

© 2026 ISTE OpenScience – Published by ISTE Ltd. London, UK – openscience.fr

présence d'identifiants DOI, PMID, PMC ou NCBI) sur le nombre total de sources. Bien que sa fiabilité puisse varier selon les types de sources (de tels identifiants n'étant associés qu'aux articles de revues, et non aux ouvrages ou chapitres d'ouvrages) et la rigueur des pratiques des contributeurs en termes de formatage des références, ce SciScore constitue une ressource intéressante pour l'étude du processus d'écriture des articles. À noter que ce score ne prend pas en compte les références d'une éventuelle section Bibliographie, dont la fonction n'est pas d'étayer les assertions de l'article mais d'offrir des recommandations de lecture aux usagers[3].

Nos travaux précédents [BOUa 21 ; BOUb 23] ayant montré que la place occupée par la presse dans les références varie fortement selon les articles, nous avons développé un second indicateur. Ce « PressScore » correspond à la part des sources de presse parmi l'ensemble des références d'un article (également hors section Bibliographie). En l'absence d'indicateurs aussi clairs que les DOI, ce sont les URL intégrées dans les références qui ont été examinées : tous les noms de domaine employés (696) ont été extraits, visités et manuellement classés. Faute de listes exhaustives ou de critères pouvant aisément être appliqués à la grande diversité des cas étudiés (en termes de format, d'historicité, de zone géographique), ont été classés comme « presse » les domaines associés à des journaux (hors revues scientifiques), des magazines, des sites dédiés à l'actualité (hors sites ou blogs personnels). Nous avons ensuite dénombré les apparitions de ces domaines dans les listes de références fournies par WikiCorpusBuilder.

4. Résultats

4.1. Présence des enjeux éthiques dans les articles

Dans un premier temps, nous nous sommes concentrés sur la seule présence des enjeux dans les articles. Un enjeu est considéré comme présent dans un article donné si au moins l'un des termes qui lui sont associés apparaît au moins une fois dans cet article. Si chaque article est associé à deux enjeux en moyenne, cette métrique connaît une distribution plus complexe (Tableau 3).

Nbr d'enjeux	Articles fr	Articles en	Total articles
0	16	12	28
1	7	8	15
2	8	5	13
3	5	9	14
4	1	4	5
5	2	0	2
6	3	1	4
7	1	4	5

Tableau 3. *Nombre d'enjeux éthiques associés à chaque article (lecture : six des sept enjeux étudiés ont été détectés dans trois articles francophones du corpus)*

On constate tout d'abord que 28 articles (16 francophones et 12 anglophones) ne sont associés à aucun des enjeux. Il s'agit principalement d'articles relatifs à des personnalités (e.g. Amy Karle), des concepts techniques (e.g. anticipation, diagnostic) et des œuvres (e.g. Artificial Intelligence: A Modern Approach). Viennent ensuite des articles plus variés, portant sur des sujets de niche ou marginalement liés à l'IA, (e.g. Match AlphaGo - Lee Sedol, Intelligence artificielle distribuée). Ces articles sans enjeux éthiques associés sont en moyenne plus courts que les autres (8201 caractères, contre 19076.5 pour l'ensemble du corpus).

Les articles où six ou sept enjeux éthiques ont pu être repérés sont principalement des articles généralistes sur l'IA (Artificial intelligence et Applications of artificial intelligence avec sept enjeux, Intelligence

[3] https://fr.wikipedia.org/wiki/Wikip%C3%A9dia:Bibliographie

© 2026 ISTE OpenScience – Published by ISTE Ltd. London, UK – openscience.fr

artificielle avec six enjeux) et des articles dédiés aux questionnements éthiques qu'elle soulève (Ethics of artificial intelligence, Éthique de l'intelligence artificielle, Regulation of artificial intelligence avec sept enjeux, Intelligence civique avec six enjeux). La présence de six enjeux dans deux articles plus spécifiques, Artificial intelligence in healthcare et Intelligence artificielle générative, semble indiquer que certains sujets font l'objet d'un traitement plus riche en questionnements éthiques. Ce parti-pris des contributeurs transparaît directement dans la structure des deux articles : Artificial intelligence in healthcare s'achève sur une section Ethical concerns (composée des sous-sections Data collection, Automation et Bias), tandis qu'Intelligence artificielle générative comporte deux sections Transparence du système d'IA ? et Vers une régulation, une réglementation et une gestion des risques. Les articles associés à six ou sept enjeux sont en moyenne plus longs que les autres (49041.4 caractères).

4.2. Influence de la longueur des articles et de la langue

Afin de neutraliser l'effet de la longueur de l'article, nous avons calculé un ratio entre celle-ci (exprimée en milliers de caractères) et le nombre d'enjeux. Cette métrique met en évidence une série d'articles courts, mais abordant un certain nombre d'enjeux éthiques. Parmi les articles avec une grande densité d'enjeux, la plupart sont plus ou moins directement liés à la régulation de l'IA : Artificial intelligence content detection (0.67 enjeux pour mille caractères), Open letter on artificial intelligence (2015) (0.56), Intelligence artificielle symbolique (0.47), Détection de contenu généré par intelligence artificielle (0.39), Regulation of algorithms (0.38). Le calcul de la même métrique pour les articles dans lesquels les sept enjeux ont été identifiés montre en revanche des densités d'enjeux bien plus faibles : Applications of artificial intelligence (0.12), Artificial intelligence (0.10), Ethics of artificial intelligence (0.13), Éthique de l'intelligence artificielle (0.22), Regulation of artificial intelligence (0.17).

La distribution des différents enjeux repérés dans les articles (Tableau 4) doit être considérée avec précautions. L'importance relative des enjeux dépend fortement de la possibilité de trouver des mots clefs spécifiques à chaque enjeu, ce qui n'était pas toujours évident. Tout au plus peut-on constater une sous-représentation des enjeux de Travail et de Responsabilité dans l'encyclopédie francophone par rapport aux articles anglophones ce qui, dans le cas du second, pourrait s'expliquer par le plus faible nombre de mots clefs dans la typologie. La plus grande longueur des articles anglophones (21533.8 caractères en moyenne contre 16619.2 pour les articles francophones) contribue sans doute aussi à ce décalage.

Enjeu	Articles fr	Articles en	Total articles
Vie privée	7	9	16
Interprétabilité	7	9	16
Biais	16	18	34
Autonomie	14	14	28
Désinformation	11	10	21
Travail	8	14	22
Responsabilité	12	21	33

Tableau 4. *Nombre d'articles associés à chaque enjeu éthique (lecture : l'enjeu Désinformation a été détecté au sein de dix articles anglophones du corpus)*

Afin de mettre en lumière les écarts entre les deux langues, nous avons calculé, pour chaque couple d'articles français/anglais, le nombre de différences dans les enjeux identifiés (le nombre d'enjeux présents dans l'un des articles et absents dans l'autre). Cette métrique montre une certaine proximité des discours des deux langues (la différence moyenne étant de 1.16 enjeux), sans pour autant que les enjeux repérés soient similaires (seuls 15 couples d'articles sur 43 ne présentent aucune différence). Les différences les plus fortes (quatre enjeux) apparaissent dans des articles longs, témoignant de réelles divergences dans les approches des contributeurs les ayant écrits. Par exemple, Applications de l'intelligence artificielle (46573 caractères) n'évoque que les questions de Biais, d'Autonomie et de Travail, tandis que l'ensemble des sept enjeux apparaissent dans Applications of artificial intelligence (60237 caractères). La disparité est plus

© 2026 ISTE OpenScience – Published by ISTE Ltd. London, UK – openscience.fr

importante encore dans le cas de Generative artificial intelligence (27090 caractères), associé aux seuls enjeux de Vie privée, Biais, Désinformation et Travail, tandis que tous les enjeux à l'exception de Vie privée apparaissent dans Intelligence artificielle générative (pourtant plus court avec 22662 caractères).

4.3. Quantification de la représentation des enjeux éthiques

Nous avons cherché également à évaluer la représentation de chaque enjeu au sein du corpus, ce qui pourrait indiquer l'importance que les contributeurs leur accordent. Nous nous sommes appuyés sur le nombre d'occurrences des termes associés à chaque enjeu dont l'interprétation est soumise à caution (e.g. dépendance de la qualité des termes), mais demeure intéressante.

Enjeu	Corpus en (%)	Corpus fr (%)	Ensemble du corpus (%)
Vie privée	16	14	16
Interprétabilité	7	7	7
Biais	36	24	32
Autonomie	8	15	11
Désinformation	8	10	9
Travail	11	12	11
Responsabilité	12	18	14

Tableau 5. *Représentation des enjeux éthiques en fonction de leur nombre d'occurrences (lecture : dans le corpus francophone,36 % des occurrences d'enjeux éthiques détectées concernent les biais)*

Les articles en anglais comptabilisent davantage d'occurrences des enjeux qu'en français, mais les proportions sont relativement proches (voir Tableau 5). Les Biais sont l'enjeu le plus représenté dans le corpus, apparaissant dans 36 % des articles anglophones et 24 % des articles francophones. Viennent ensuite, dans le corpus anglophone, les enjeux Vie privée (16 %), Responsabilité (12 %) et Travail (11 %). On retrouve les mêmes enjeux dans les articles français, mais dans des proportions différentes : Responsabilité (18 %), Travail (15 %) et Vie privée (14 %). A contrario, les enjeux Désinformation (8 % en anglais et 10 % en français) et Interprétabilité (7 % dans les deux langues) sont les plus faiblement représentés dans le corpus.

© 2026 ISTE OpenScience – Published by ISTE Ltd. London, UK – openscience.fr



Figure 1. *Graphes de liens entre les enjeux et les articles du corpus anglais puis français*

Les disparités de représentation des enjeux sont également visibles sur les graphes de liens entre les enjeux et les articles Wikipédia (Figure 1), construits à l'aide du logiciel Gephi : un lien indique que l'enjeu est présent dans l'article, l'épaisseur du lien est fonction du nombre d'occurrences. On y retrouve la centralité de l'enjeu Biais, ainsi que ses liens forts avec les articles anglophones Ethics of AI, AI in healthcare, AI, Regulation of AI et les articles français Réglementation de l'IA, IA générative et Ethique de l'IA. Il en va de même pour les enjeux Responsabilité et Vie privée, avec une plus forte représentation dans les articles

© 2026 ISTE OpenScience – Published by ISTE Ltd. London, UK – openscience.fr

AI in healthcare et IA pour la santé. On remarque que si l'enjeu Désinformation est le moins représenté, il est fortement cité par l'article anglais Generative AI tandis qu'en français sa présence est due aux articles Course aux armements intelligents, Ethique de l'IA et IA.

4.4. Analyse des références bibliographiques

Les articles du corpus contiennent en moyenne 61 références (selon WikiCorpusBuilder), avec une forte différence selon les versions linguistiques : 46 en moyenne pour les articles francophones et 76 pour les articles anglophones. WikiCorpusBuilder n'a identifié aucune référence pour dix articles, assez courts, portant principalement sur des notions techniques, des institutions ou des personnes. Dans l'ensemble, le SciScore est corrélé avec le nombre d'enjeux éthiques détectés (les articles abordant un grand nombre d'enjeux citent davantage de sources académiques), mais pas le PressScore (Tableau 6).

Nbr d'enjeux	SciScore moyen	PressScore moyen
0	0.042	0.124
1	0.087	0.141
2	0.133	0.217
3	0.133	0.251
4	0.198	0.273
5	0.155	0.294
6	0.226	0.210
7	0.260	0.187

Tableau 6. *SciScores et PressScores des articles par nombre d'enjeux éthiques identifiés (lecture : dans les articles associés à trois enjeux éthiques, les sources scientifiques comptent pour 13,3 % des références)*

Les articles présentent en moyenne un SciScore de 0.114 (c'est-à-dire que 11.4 % des sources été classées comme scientifiques) et un PressScore de 0.182 (18.2 % de sources classées comme issues de la « presse »). Parmi les neuf articles ayant des SciScores élevés (supérieurs à 0.3), trois sont dédiés à des notions techniques (Frame (artificial intelligence), Symbolic artificial intelligence, Intelligence artificielle symbolique), deux à des domaines de recherche (Sciences cognitives, Argumentation theory), tandis que les quatre autres portent sur les usages et l'éthique de l'IA (e.g. Intelligence artificielle dans la santé, Ethics of artificial intelligence). Le SciScore exceptionnel (0.71) de Artificial intelligence in healthcare confirme le caractère spécifique de cet article.

Les PressScores élevés (supérieurs à 0.4) sont surtout associés à des usages controversés de l'IA (e.g. Course aux armements intelligents, Artificial intelligence content detection) et à des initiatives ou des acteurs liés à la régulation de l'IA (e.g. Lettre ouverte sur l'intelligence artificielle), mais aussi une œuvre de fiction (A.I. Intelligence artificielle) et un article plus généraliste (Generative artificial intelligence). On constate, plus généralement, une légère sur-représentation des sources scientifiques dans les articles anglophones (SciScore moyen de 0.144 contre 0.0819 pour les francophones), tandis que la proportion des sources de presse est assez similaire (PressScore moyen de 0.179 pour les articles francophones contre 0.185 pour les anglophones).

© 2026 ISTE OpenScience – Published by ISTE Ltd. London, UK – openscience.fr

Enjeu	SciScore moyen	SciScore fr	SciScore en	PressScore moyen	PressScore fr	PressScore en
Vie privée	0.211	0.155	0.255	0.238	0.198	0.269
Interpretabilité	0.265	0.223	0.297	0.189	0.208	0.175
Biais	0.175	0.143	0.203	0.238	0.228	0.248
Autonomie	0.166	0.107	0.225	0.198	0.182	0.214
Désinformation	0.124	0.062	0.191	0.290	0.261	0.321
Travail	0.137	0.060	0.187	0.235	0.252	0.224
Responsabilité	0.190	0.099	0.246	0.186	0.212	0.17

Tableau 7. *SciScore et PressScore des articles associés à chaque enjeu éthique (lecture : dans les articles anglophones associés à l'enjeu Travail, les sources de presse comptent pour 22,4 % des références)*

La distribution des SciScores en fonction des enjeux éthiques détectés montre des écarts non négligeables (voir Tableau 7). Les articles associés aux enjeux d'Interprétabilité et de Vie privée comportent en moyenne une plus forte proportion de sources scientifiques que ceux relatifs aux enjeux de Désinformation et de Travail, ce qui pourrait indiquer une approche différente des contributeurs traitant de ces sujets (en anglais et en français). Les articles associés à la Responsabilité présentent un SciScore plus élevé dans l'encyclopédie anglophone que dans l'encyclopédie francophone, montrant que les deux groupes de contributeurs ne mobilisent pas forcément les mêmes sources pour évoquer les mêmes thèmes.

On trouve d'importants écarts dans les PressScores, avec des valeurs plus élevées pour les articles liés à la Désinformation, puis à la Vie privée et aux Biais, et plus faibles pour les articles associés à la Responsabilité et à l'Interprétabilité. Les différences entre les langues montrent, là aussi, des approches différentes, avec une sur-représentation des sources de presse dans les articles anglophones relatifs à la Vie privée et à la Désinformation.

Enfin, il faut noter que la distribution des deux scores n'est pas corrélée. Si le SciScore élevé des articles liés à l'Interprétabilité est associé à un PressScore assez faible, ce n'est pas le cas des articles liés à la Vie privée, qui affichent des valeurs élevées pour les deux scores.

5. Discussions

Les résultats montrent que, selon les articles, les contributeurs abordent les enjeux de l'IA de manière différente. On distingue quatre principales catégories d'article : i) dans les longs articles généralistes sur l'IA (e.g. Intelligence artificielle), de nombreux enjeux sont abordés, mais de manière assez marginale ; ii) viennent ensuite des articles longs dédiés à l'éthique de l'IA, ou à son application en médecine, où les enjeux éthiques sont aussi nombreux, mais occupent une place plus importante ; iii) des articles courts portant sur des points spécifiques de la régulation de l'IA (e.g. Détection de contenu généré par intelligence artificielle) où les enjeux éthiques, bien que peu nombreux, conservent une place centrale ; et enfin iv) des articles courts portant sur des acteurs de l'IA ou des concepts techniques (e.g. Amy Karle), où les enjeux éthiques sont absents ou presque. Ces catégories appuient l'hypothèse d'une présence hétérogène des discours critiques de l'IA se traduisant par une distinction entre des articles à visée généraliste qui traite peu les enjeux éthiques et des articles spécialisés où se trouvent majoritairement les discours critiques. On observe effectivement une forme de spécialisation des articles, par laquelle les enjeux ne sont abordés de manière détaillée que dans des articles spécifiques. Cette tendance est cohérente avec les règles de l'encyclopédie, qui recommandent l'articulation d'« articles généraux » et d'« articles détaillés », plutôt qu'un seul article très long[4].

Ces différences de distribution des enjeux se combinent avec des approches différentes en matière d'usage des références. Les longs articles consacrés à l'éthique de l'IA et à la santé s'appuient sur une forte

[4] https://fr.wikipedia.org/wiki/Aide:Comment_r%C3%A9diger_un_bon_article#Taille

© 2026 ISTE OpenScience – Published by ISTE Ltd. London, UK – openscience.fr

proportion de sources académiques et assez peu d'articles de presse, tandis que la tendance s'inverse pour les longs articles généralistes sur l'IA. Ceci, ainsi que la corrélation entre le nombre d'enjeux et le SciScore, semble indiquer que les contributeurs préfèrent citer la littérature scientifique lorsqu'ils abordent les questions d'éthique. Mais cela pourrait aussi refléter un déséquilibre des sources disponibles, si les documents traitant de l'IA sous l'angle de l'éthique sont majoritairement issus du monde académique.

Le cas particulier de la sur-représentation des sources de presse dans les articles concernant la Désinformation pourrait appuyer l'hypothèse de la disponibilité des sources comme facteur expliquant le choix des contributeurs. La désinformation est un enjeu traité de manière récurrente par les médias — le cas de la diffusion de « fake news » sur les réseaux sociaux, par exemple — en raison de son lien étroit avec leur cœur de métier, la production d'informations fiables, qui s'est vu bousculée par l'essor de l'IA générative. Les formes de « panique médiatique » associées à ce sujet, que l'on peut qualifier de « moment de questionnement d'un média par un autre » [FRA 19, p.10], se traduisent par une forte couverture médiatique, qui se reflète alors dans les références des articles Wikipédia. De la même manière, la sur-représentation des enjeux de Biais et de Vie privée parmi les articles du corpus pourrait être expliquée par la popularité de ces thèmes dans la presse [NGU 22], dont l'encyclopédie se ferait alors le reflet. En revanche, les importantes différences constatées entre les deux langues, avec certains enjeux plus présents que d'autres, ou davantage associés à certains types de sources que d'autres, ne semblent pas trouver d'explication évidente et nécessiteraient un travail qualitatif plus poussé sur le processus d'écriture des articles et sur les discours dont il peut se nourrir.

Les particularités relevées sur plusieurs articles témoignent aussi du traitement spécifique dont bénéficient certains sujets. Certains événements semblent avoir un fort impact sur l'évocation et la représentation des enjeux de l'IA. C'est notamment le cas d'une lettre ouverte rédigée par le Future of Life Institute en 2015 et signée par plusieurs experts en IA et personnalités médiatiques, dont Stephen Hawking et Elon Musk. L'article anglophone dédié (Open letter on artificial intelligence (2015)) synthétise les risques de l'IA évoqués par la lettre, accordant une place centrale aux enjeux Travail et Vie privée. La forte part des références de presse (un PressScore de 0.5 et un SciScore nul) dans cet article confirme que des événements controversés et à forte portée médiatique ont une grande influence sur la diffusion des discours critiques dans l'espace public.

Parmi les articles contribuant le plus à la représentation des enjeux éthiques, ceux relatifs à la santé se démarquent également : par leur sujet, portant sur un domaine spécifique d'application des systèmes d'IA pour lequel il existe à la fois de potentiels usages et des risques importants, mais aussi par un recours inégalé à des sources académiques. Plusieurs facteurs peuvent expliquer ce cas [MIG 22] : l'ancienneté des liens entre l'IA et la santé (des systèmes d'IA dédiés à la santé sont créés dès 1960), sa relative maturité (avec une intégration dans les pratiques, notamment pour l'imagerie médicale), et les attentes fortes qui y sont associées. L'importante perspective éthique relevée dans ces articles de l'encyclopédie trouve sans doute ces racines dans le riche historique du domaine médical en matière d'analyse des risques et de questionnements éthiques. Les travaux sur la régulation de l'IA par l'éthique s'inspirent notamment des cadres développés par la bioéthique [FLO 23].

La longueur des articles dédiés à l'IA générative et le grand nombre d'enjeux éthiques qui y sont abordés peuvent être directement reliés à la « vague de l'IA générative ». L'important succès rencontré par ChatGPT après sa sortie en novembre 2022 a ouvert la voie à une course au développement de systèmes génératifs par les grandes entreprises du numérique donnant lieu à une multitude d'applications dans des domaines variés [AND 23]. Les controverses nombreuses et largement médiatisées entourant l'usage de ces outils expliquent sans doute la forte représentation des enjeux éthiques et des sources de presse dans les deux articles portant sur l'IA générative, mais également dans des articles connexes (e.g. Artificial intelligence content detection).

© 2026 ISTE OpenScience – Published by ISTE Ltd. London, UK – openscience.fr

6. Conclusion

Cette étude porte sur la manière dont les discours critiques sur l'IA sont intégrés dans l'encyclopédie Wikipédia. La représentation de sept enjeux majeurs est examinée dans un corpus de 43 paires d'articles francophones et anglophones portant sur l'IA, dont les références font également l'objet d'une analyse. Les résultats montrent des disparités dans le traitement des enjeux, en termes d'occurrences sur l'ensemble du corpus et de distribution dans différents articles. Marginalement présents dans les articles généralistes, les enjeux sont principalement développés dans des articles spécialisés dédiés à l'éthique et à la régulation de l'IA, mais aussi d'articles portant sur des sujets ayant fait l'objet de débats ou de controverses éthiques. On constate également d'importantes différences d'usage des références dans le corpus, certains sujets (l'IA générative) et certains enjeux éthiques (la désinformation) étant principalement documentés à partir de sources de presse tandis que d'autres sujets (la santé) et d'autres enjeux (notamment l'interprétabilité) sont plutôt associés à des sources académiques. Ces résultats confirment en outre le rôle central des médias de masse dans la structuration des débats sur les enjeux de l'IA, sur Wikipédia comme dans l'espace public.

Au-delà de cette conclusion, plusieurs points doivent encore être clarifiés. Le processus d'écriture, au cours duquel les contributeurs choisissent, parmi les angles possibles et les sources disponibles, ce qui sera présenté dans l'article, pourrait être éclairé par une étude qualitative à un niveau plus micro (e.g. pages de discussion, historiques des articles, contexte lexical d'apparition des enjeux), qui pourrait être complété par des entretiens avec les contributeurs. À un niveau plus macro, et diachronique, il serait fécond d'examiner la manière dont différents thèmes sont développés, transmis, réactivés, au gré des événements, par les journalistes, les chercheurs, les wikipédiens, etc. Ces deux pistes de recherche permettraient de mieux comprendre comment, sur le sujet de l'IA comme dans d'autres controverses socio-techniques, le discours singulier que l'on retrouve sur Wikipédia est relié à la multitude des discours produits et mis en circulation par les acteurs variés de l'espace public.

7. Références

[AND 23] ANDLER D., *Intelligence artificielle, intelligence humaine : la double énigme*, Éditions Gallimard, Paris, 2023.

[BEA 15] BEAUDE B., « De quoi Wikipédia est-elle le lieu ? », dans L. Barbe, L. Merzeau et V. Schafer (dir.), *Wikipédia, objet scientifique non identifié*, Presses universitaires de Paris Nanterre, Nanterre, pp. 41-54, 2015.

[BEL 23] BELLON A., VELKOVSKA J., « L'intelligence artificielle dans l'espace public : du domaine scientifique au problème public. Enquête sur un processus de publicisation controversé », *Réseaux*, n°°240, pp. 31-70, 2023.

[BENB 18] BENBOUZID B., CARDON D., « Machines à prédire », *Réseaux*, n°°211, pp. 9-33, 2018.

[BEND 25] BENDER E.M., HANNA A., *The AI Con: How to Fight Big Tech's Hype and Create the Future We Want*, Harper, New York, 2025.

[BENJ 18] BENJAKOB O., AVIRAM R., « A Clockwork Wikipedia: From a Broad Perspective to a Case Study », *Journal of Biological Rhythms*, vol. 33, n°°3, pp. 233-244, 2018.

[BENJ 22] BENJAKOB O., AVIRAM R., SOBEL J.A., « Citation needed? Wikipedia bibliometrics during the first wave of the COVID-19 pandemic », *GigaScience*, vol. 11, 2022.

[BENJ 23] BENJAKOB O., GULEY O., SEVIN J.-M., BLONDEL L., AUGUSTONI A., COLLET M., JOUVESHOMME L., AMIT R., LINDER A., AVIRAM R., « Wikipedia as a tool for contemporary history of science: A case study on CRISPR », *PLoS One*, vol. 18, n°°9, 2023.

[BOUa 21] BOUCHEREAU A., SEGAULT A. « Évolution des représentations de l'intelligence artificielle : De quelle(s) IA Wikipédia est-elle l'encyclopédie ? », dans I. SALEH, N. BOUHAI, S. LELEU-MERVIEL, I. ROXIN, M. ZACKLAD, ET L. MASSOU (dir.), *H2PTM'21 Information : Enjeux et nouveaux défis*, ISTE Editions, Londres, 2021.

[BOUb 23] BOUCHEREAU A., SEGAULT A. « Une synthèse raisonnée de l'IA ? Usage des références bibliographiques des articles de Wikipédia sur l'Intelligence artificielle », dans I. SALEH, N. BOUHAÏ, S. LELEU-MERVIEL, I. ROXIN, M. ZACKLAD, L. MASSOU, H. BOULEKBACHE-MAZOUZ, P. USEILLE (dir.), *H2PTM'23 - La fabrique du sens à l'ère de l'information numérique : enjeux et défis*, ISTE Editions, Londres, 2023.

© 2026 ISTE OpenScience – Published by ISTE Ltd. London, UK – openscience.fr

[BUN 22] BUNZ M., BRAGHIERI M., « The AI doctor will see you now: assessing the framing of AI in news coverage », *AI & Society*, vol. 37, n°°1, pp. 9-22, 2022.

[CAR 18] CARDON D., COINTET J.-P., MAZIERES A., « La revanche des neurones : l'invention des machines inductives et la controverse de l'intelligence artificielle », *Réseaux*, n°°211, pp. 173-220, 2018.

[CHE 19] CHEATHAM, Benjamin, JAVANMARDIAN, Kia, SAMANDARI, Hamid. « Confronting the risks of artificial intelligence », *McKinsey Quarterly*, vol. 2, n°°38, pp. 1-9, 2019.

[CHU 19] CHUAN C.-H., TSAI W.-H.S., CHO S.Y., « Framing artificial intelligence in American newspapers », dans *Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society*, ACM, pp. 339-344, 2019.

[CRE 22] CREPEL M., CARDON D., « Robots vs algorithmes : prophétie et critique dans la représentation médiatique des controverses de l'IA », *Réseaux*, n°°232-233, pp. 129-167, 2022.

[DOU 20] DOUTREIX M.-N., *Wikipédia et l'actualité : qualité de l'information et normes collaboratives d'un média en ligne*, Presses Sorbonne Nouvelle, Paris, 2020.

[FOR 22] FORD H., *Writing the revolution: Wikipedia and the survival of facts in the digital age*, MIT Press, Cambridge, 2022.

[FLO 23] FLORIDI L., *L'éthique de l'intelligence artificielle : principes, défis et opportunités*, trad. E. Panai et E. Goffi, Éditions Mimésis, Sesto San Giovanni, 2023.

[GOU 24] GOURLET P., RICCI D., CRÉPEL M., « Reclaiming artificial intelligence accounts: A plea for a participatory turn in artificial intelligence inquiries », *Big Data & Society*, vol. 11, n°°2, 2024.

[INI 24] INIE N., DRUGA S., ZUKERMAN P., BENDER E.M., « From "AI" to Probabilistic Automation: How Does Anthropomorphization of Technical Systems Descriptions Influence Trust? », dans *Proceedings of the ACM Conference on Fairness, Accountability, and Transparency*, Association for Computing Machinery, New York, pp. 2322-2347, 2024.

[JOB 19] JOBIN A., IENCA M., VAYENA E., « The global landscape of AI ethics guidelines », *Nature Machine Intelligence*, vol. 1, pp. 389-399, 2019.

[KOP 15] KOPPEN L., PHILLIPS J., PAPAGEORGIOU R., « Analysis of reference sources used in drug-related Wikipedia articles », *Journal of the Medical Library Association*, vol. 103, n°°3, pp. 140-144, 2015.

[LAN 15] LANGLAIS P.-C., « L'émergence d'une norme wikipédienne », dans L. Barbe, L. Merzeau et V. Schafer (dir.), *Wikipédia, objet scientifique non identifié*, Presses Universitaires de Paris Ouest, Nanterre, 2015.

[LEC 15] LECUN Y., « L'apprentissage profond : une révolution en intelligence artificielle », Collège de France, 2015, disponible à l'adresse : http://www.college-de-france.fr/site/yann-lecun/course-2016-02-12-14h30.htm.

[LEC 19] LECUN Y., *Quand la machine apprend : la révolution des neurones artificiels et de l'apprentissage profond*, Odile Jacob, Paris, 2019.

[LEE 25] LEE H.-P., SARKAR A., TANKELEVITCH L., DROSOS I., RINTEL S., BANKS R., WILSON N., « The Impact of Generative AI on Critical Thinking: Self-Reported Reductions in Cognitive Effort and Confidence Effects From a Survey of Knowledge Workers », dans *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems, Association for Computing Machinery*, New York, article 1121, pp. 1-22, 2025.

[MAC 18] MACLURE J., SAINT-PIERRE M.-N., « Le nouvel âge de l'intelligence artificielle : une synthèse des enjeux éthiques », *Les Cahiers de propriété intellectuelle*, n°°30, p. 741-766, 2018.

[MCD 21] MCDOWELL Z.J., VETTER M.A., *Wikipedia and the Representation of Reality*, Routledge, New York, 2021.

[MIG 22] MIGNOT L., SCHULTZ É., « Les innovations d'intelligence artificielle en radiologie à l'épreuve des régulations du système de santé », *Réseaux*, n°°232-233, pp. 65-97, 2022.

[NGU 24] NGUYEN D., HEKMAN E., « The news framing of artificial intelligence: a critical exploration of how media discourses make sense of automation », *AI & Society*, vol. 39, pp. 437-451, 2024.

[OEB 20] OEBERST A., VON DER BECK I., MATSCHKE C., IHME T.A., CRESS U., « Collectively biased representations of the past: Ingroup Bias in Wikipedia articles about intergroup conflicts », *British Journal of Social Psychology*, n°°59, pp. 791-818, 2020.

[ROG 12] ROGERS R., SENDIJAREVIC E., « Neutral or National Point of View? A Comparison of Srebrenica Articles across Wikipedia's Language Versions », 2012, disponible à l'adresse : https://dare.uva.nl/search?identifier=608ec3d8-fde6-4a0d-ba1c-948e1af97b22, consulté le 22 avril 2023.

© 2026 ISTE OpenScience – Published by ISTE Ltd. London, UK – openscience.fr

[ROM 22] ROMELE A., « Images of Artificial Intelligence: a Blind Spot in AI Ethics », *Philosophy & Technology*, vol. 35, 2022.

[ROM 24] ROMELE A., « Éthique de l'intelligence artificielle comme signifiant flottant : considérations théoriques et analyse critique des discours de presse », *Interfaces numériques*, vol. 13, n°°1, 2024.

[RUS 10] RUSSELL S., NORVIG P., *Intelligence artificielle*, 3e édition, Pearson France, Paris, 2010.

[SAH 16] SAHUT G., « L'autorité importée dans Wikipédia : la question de la qualité des sources citées », *Quaderni. Communication, technologies, pouvoir*, n°°91, pp. 93-105, 2016.

[TUL 25] TULLY S.M., LONGONI C., APPEL G., « Lower Artificial Intelligence Literacy Predicts Greater AI Receptivity », *Journal of Marketing*, vol. 89, n°°5, 2025.

[VAY 20] VAYRE J.-S., GAGLIO G., « L'intelligence artificielle n'existe-t-elle vraiment pas ? Quelques éléments de clarification autour d'une science controversée », *Diogène*, n°°269-270, pp. 107-120, 2020.

[WAN 23] WANG W., DOWNEY J., YANG F., « AI anxiety? Comparing the sociotechnical imaginaries of artificial intelligence in UK, Chinese and Indian newspapers », *Global Media and China*, 2020.

[YAN 24] YANG P., COLAVIZZA G., « Polarization and reliability of news sources in Wikipedia », *Online Information Review*, vol. 48, n°°5, pp. 908-925, 2024.

© 2026 ISTE OpenScience – Published by ISTE Ltd. London, UK – openscience.fr