



# Une synthèse raisonnée de l'IA ? Usage des références bibliographiques des articles de Wikipédia sur l'Intelligence artificielle

Aymeric Bouchereau, Antonin Segault

## ► To cite this version:

Aymeric Bouchereau, Antonin Segault. Une synthèse raisonnée de l'IA ? Usage des références bibliographiques des articles de Wikipédia sur l'Intelligence artificielle. H2PTM'23 : La fabrique du sens à l'ère de l'information numérique : enjeux et défis, Oct 2023, Valenciennes, France. pp.117-131. hal-04296175

**HAL Id: hal-04296175**

**<https://hal.science/hal-04296175v1>**

Submitted on 1 Aug 2025

**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.

---

# Une synthèse raisonnée de l'IA ?

## Usage des références bibliographiques des articles de Wikipédia sur l'Intelligence artificielle

Aymeric Bouchereau\*, Antonin Segault\*\*

*\*Laboratoire Lab'Urba, Université Paris-Est Créteil  
61 avenue du Général de Gaulle, 94000 Créteil  
aymeric.bouchereau@u-pec.fr*

*\*\*Laboratoire Dicen IdF, UFR PHILLIA, Université Paris Nanterre  
200 avenue de la République, 92000 Nanterre  
antonin.segault@parisnanterre.fr*

---

**RÉSUMÉ.** Cet article examine la manière dont les contributeurs de l'encyclopédie Wikipédia s'efforcent de construire une synthèse raisonnée des connaissances relatives à l'intelligence artificielle. Il s'appuie pour cela sur l'étude des références bibliographiques mobilisées dans six articles, couvrant trois sujets dans deux langues. Les analyses qualitatives et quantitatives montrent des approches différentes selon les articles, certains sujets étant plus fréquemment associés à des discours prospectifs ou relatifs à des controverses. On observe également que, si ces discours sont souvent appuyés par des sources issues de la presse, ils ne sont pas pour autant absents des références scientifiques.

**ABSTRACT.** This article focuses on how the editors of Wikipedia try to write a compendium of all knowledge on artificial intelligence. It therefore studies the bibliographic references used in six articles related to three topics in two languages. The quantitative and qualitative analysis show differences between the articles, with some topics more frequently associated to prospective or controversial discourses. It also appears that, while such discourses are often backed by press sources, they can also be found in scientific references.

**MOTS-CLÉS :** wikipédia, encyclopédisme, intelligence artificielle, controverse, point de vue, bibliographie, médiation des savoirs

**KEYWORDS:** wikipedia, encyclopedism, artificial intelligence, controversy, point of view, bibliography, knowledge mediation

---

## 1. Introduction

Pas une semaine ne passe sans que la presse n'évoque les derniers exploits de systèmes d'Intelligence artificielle (IA) et les nouvelles perspectives qu'ils pourraient ouvrir. Cette omniprésence du sujet n'est pas seulement le résultat d'une abondance de travaux de recherche dans le domaine, elle est aussi révélatrice de l'intensité et de la variété des controverses qui les entourent. Témoins de ce phénomène, les systèmes d'IA génératifs conçus depuis 2020 (e.g. ChatGPT, Midjourney) représentent un succès scientifique du fait de la vraisemblance des contenus textuels et visuels qu'ils produisent (Cao *et al.*, 2023). Dans le même temps, ces systèmes font l'objet d'une forte attention médiatique ; la presse annonçant des prouesses, comme la réussite à un examen<sup>1</sup>, ou des dangers liés au remplacement ou au surpassement des humains<sup>2</sup> (Bunz & Braghieri, 2022).

Dans un tel contexte, cette recherche s'intéresse aux dispositifs socio-numériques qui, en opérant une médiation entre les connaissances scientifiques et les individus, permettent d'appréhender l'IA et ses enjeux. En particulier, nous étudions le processus par lequel les individus peuvent donner du sens aux concepts relatifs à l'IA à partir des informations présentes dans Wikipédia. Pierre angulaire du web et de nombreux dispositifs de recherche d'information, cette encyclopédie libre joue un rôle central dans l'agrégation et la diffusion de connaissances relatives à une large variété de sujets, incluant les événements d'actualités et les controverses les plus en vue. Dans un précédent travail (Bouchereau & Segault, 2021), il a été montré que le contenu des articles de l'encyclopédie portant sur l'IA comportait des discours prospectifs, voire sensationnalistes, et des controverses. Poursuivant ce travail, cette contribution examine comment et dans quelle mesure les contributeurs de l'encyclopédie ont pu développer une « synthèse raisonnée » des multiples points de vue concernant l'IA.

Nous reviendrons dans un premier temps sur les évolutions récentes du domaine de l'IA et les controverses qu'elles ne manquent pas d'occasionner. Nous décrirons ensuite les processus par lesquels les contributeurs de Wikipédia cherchent à assurer la neutralité de point de vue des articles qu'ils produisent. Puis, nous présenterons une étude de cas s'appuyant sur l'analyse des références bibliographiques mobilisées dans six articles de Wikipédia, et dont les résultats mettent en évidence les différentes approches adoptées par les contributeurs afin de décrire les concepts relatifs à IA.

## 2. Comprendre l'IA à partir de Wikipédia

L'histoire de l'IA est marquée par des périodes intenses, caractérisées par une accumulation de résultats scientifiques et un fort intérêt de la part d'acteurs variés (e.g.

---

<sup>1</sup> [https://www.bfmtv.com/tech/intelligence-artificielle/royaume-uni-grace-a-chat-gpt-des-etudiants-obtiennent-la-meilleure-note-possible-a-un-examen\\_AV-202304110184.html](https://www.bfmtv.com/tech/intelligence-artificielle/royaume-uni-grace-a-chat-gpt-des-etudiants-obtiennent-la-meilleure-note-possible-a-un-examen_AV-202304110184.html)

<sup>2</sup> <https://start.lesechos.fr/innovations-startups/tech-futur/est-ce-que-chatgpt-va-mettre-la-france-au-chomage-1936321>

investisseurs, industriels, politiques, journalistes) (Cardon *et al.*, 2018; Vayre & Gaglio, 2020). Nous connaissons l'une de ces périodes depuis le début des années 2010 en raison des performances obtenues par les méthodes d'apprentissage automatique (*machine learning*) et, surtout, d'apprentissage profond (*deep learning*). Ces méthodes sont particulièrement efficaces pour analyser de grands volumes de données (e.g. images, textes), en extraire des informations utiles (e.g. détection d'objets, classification de textes) pour, *in fine*, effectuer des tâches auparavant réalisées par des humains (LeCun, 2019). Ces succès scientifiques se sont accompagnés d'investissements conséquents et d'applications concrètes, rapidement intégrées dans les dispositifs socio-numériques du quotidien, soulevant ainsi des questions multiples et complexes (Benbouzid & Cardon, 2018).

### 2.1. Représentations de l'IA et de ses enjeux

Les enjeux afférents à l'IA sont abondamment discutés par la communauté scientifique, notamment *via* les travaux sur l'éthique (Floridi *et al.*, 2021), ainsi que par les entreprises et les institutions qui créent des instances de régulation et des cadres éthiques et juridiques. Ces enjeux concernent également les citoyens qui interagissent désormais étroitement et quotidiennement avec des systèmes d'IA, et sont confrontés à des problématiques concrètes et immédiates telles que le respect de la vie privée, les discriminations liées aux biais algorithmiques ou encore l'explicabilité des décisions algorithmiques prises par les institutions (Benbouzid & Cardon, 2018; O'Neil, 2018; Zacklad *et al.*, 2022). Ainsi, le développement d'une littératie de l'IA devient nécessaire afin que les citoyens acquièrent un regard critique, une compréhension des traitements appliqués aux données et des enjeux éthiques et sociaux (Le Deuff & Roumanos, 2022).

Dans une précédente étude sur l'évolution du contenu de l'article « Intelligence artificielle » de Wikipédia (Bouchereau & Segault, 2021), nous avons montré que les enjeux sus-cités étaient sous-représentés face aux discours plus prospectifs et controversés. Ce constat est partagé par plusieurs études sur la presse écrite, rapportant qu'elle tend à se faire le relais de discours controversés, notamment centrés sur le remplacement ou le déclassement des humains face à l'IA (Bunz & Braghieri, 2022; Chuan *et al.*, 2019; Crépel & Cardon, 2022; Vergeer, 2020). On trouvera par exemple des articles sur les dangers que fait peser l'automatisation sur le marché de l'emploi, le dépassement de l'intelligence humaine ou encore la destruction de l'humanité par une IA devenue surpuissante.

Ces observations suggèrent que la presse traite davantage l'IA en mobilisant des imaginaires élaborés à partir de récits fictionnels mettant en scène des IA et des robots bienveillants ou hostiles (Romele, 2022). Plusieurs raisons peuvent expliquer ce biais, à commencer par la mauvaise qualité sémantique de l'expression « IA », qui obscurcit la réalité qu'elle recouvre plutôt qu'elle ne l'éclaire (Kyrou, 2020). Pour d'autres, à l'instar de Vayre et Gaglio (2020), l'ambiguïté viendrait plutôt de son

utilisation faite par des acteurs divers (e.g. chercheurs, industriels, journalistes, décideurs) aux intérêts différents (e.g. scientifique, économique, politique). Il semble que la diversité des acteurs alimentant les discussions, leurs différences de points de vue et les imaginaires qu'ils mobilisent génèrent de la confusion sur ce qu'est l'IA et ses enjeux.

## 2.2. Points de vue et références dans les articles Wikipédia

Créée en 2001, l'encyclopédie en ligne Wikipédia est depuis quinze ans l'un des dix sites web les plus consultés au monde (Alexa, s. d.). Ses articles apparaissent fréquemment en première position dans les résultats des moteurs de recherche tels que Google Search ou Bing et leur contenu nourrit largement les réponses des assistants vocaux comme Siri et Alexa (Minhaj & Urs, 2021). Cette encyclopédie s'inscrit dans la lignée du projet encyclopédique en s'efforçant de « présenter une synthèse raisonnée de l'ensemble du savoir humain établi » (Wikipédia, s. d.-c). Elle s'en distingue en revanche par les moyens employés : dans cette encyclopédie libre, les articles peuvent être modifiés, complétés, corrigés par toute personne disposant d'un accès au web. Cette approche, souvent décriée, a pourtant permis à Wikipédia d'atteindre des taux d'erreur du même ordre de grandeur que dans les encyclopédies traditionnelles (Casebourne *et al.*, 2012; Giles, 2005).

L'activité des contributeurs à l'encyclopédie est encadrée par un large corpus de règles, elles-mêmes en grande partie produites et amendées par la communauté. Parmi les « cinq piliers », le principe de Neutralité de point de vue consiste à « présenter tous les points de vue pertinents, en les attribuant à leurs auteurs, mais sans en adopter aucun » (Wikipédia, s. d.-b). La mention des sources, rare dans les premières années de Wikipédia, s'est progressivement développée, d'abord avec l'inclusion de courtes bibliographies à la fin des articles puis, à partir des années 2005, avec le développement d'un robuste système de référence, régit par des règles strictes (Langlais, 2015). Cette norme, adossée à une hiérarchisation plus ou moins explicite de l'autorité de différentes sources (les publications académiques sont à privilégier, certains sites sont considérés comme problématiques), contribue à la légitimation du contenu des articles (Cardon & Levrel, 2009; Sahut, 2016).

Dès qu'un sujet est sujet à controverse, le principe de Neutralité de point de vue implique de présenter les différents points de vue en leur accordant « une place proportionnelle à leur importance dans les études sur le sujet » (Wikipédia, s. d.-b). Concrètement, cet équilibre se base sur « la proportion octroyée par les sources fiables », aboutissant généralement à une prédominance de l'opinion majoritaire dans les médias, la littérature scientifique ou le sens commun (Doutreix, 2020, p. 118). Ces normes ne produisent toutefois pas un lissage total des points de vue, laissant la une certaine marge d'expression d'opinions. Par exemple, sur des sujets conflictuels tels que les Guerres de Yougoslavie, la comparaison de plusieurs versions linguistiques d'un même article révèle d'importantes différences dans la couverture du sujet (Ro-

gers & Sendijarevic, 2012). Influencés par les points de vue de leurs auteurs, mais aussi ceux des sources sur lesquelles ces derniers appuient leur travail, les articles de Wikipédia peuvent ainsi être considérés comme simultanément victimes et sources de biais informationnels (McDowell & Vetter, 2021, p. 59).

Au vu de ces éléments, il apparaît que le point de vue adopté dans les articles de Wikipédia et les références bibliographiques qui l'appuient pourrait jouer un rôle important dans la diffusion des discours prospectifs et controverses et, conséquemment, sur le sens que donnent les individus à l'IA.

### 3. Cas d'étude et méthodologie

Dans la continuité du premier article (Bouchereau & Segault, 2021), cette étude cherche à expliciter l'origine de la mise en avant de discours prospectifs sur l'IA au détriment de ceux portant sur des réalités socio-techniques. L'étude examine les références bibliographiques mobilisées pour appuyer ces discours et cherche à les mettre en perspectives en comparant plusieurs articles de l'encyclopédie dans des langues différentes. Il s'agit ainsi de montrer comment les contributeurs wikipédiens construisent une « synthèse raisonnée » des savoirs sur l'IA, à partir de laquelle les lecteurs construisent leurs représentations.

Pour cette étude, six articles ont été sélectionnés, correspondant à trois sujets liés à l'IA (intelligence artificielle, apprentissage automatique et apprentissage profond) en français et en anglais<sup>3</sup>. L'analyse portera sur les références bibliographiques de ces articles tels qu'ils étaient visibles début décembre 2022. L'ensemble des références (dans la section « Références », donc, mais aussi dans la section « Bibliographie » qui coexiste dans plusieurs articles) a été collecté et les doublons (la même référence mobilisée plusieurs fois dans un même article) supprimés. Dans le cas des références incomplètes (assez fréquentes), le titre du document et sa date de publication ont été manuellement recherchés (dans les quelques cas où la date était introuvable, notamment pour des sites web, elle a été notée « none »).

Ces références bibliographies ont été classées par type, selon leur origine :

- science : articles de revues à comité de lecture, actes de conférences, chapitres d'ouvrages collectifs publiés auprès d'éditeurs scientifiques ;
- presse : articles de journaux ou de médias en ligne reconnus ;
- livre : monographies, manuels scolaires ;
- institution publique : documents produits par des agences gouvernementales, des institutions publiques nationales ou supranationales (ONU, UE) ;

---

<sup>3</sup> Pour des raisons pratiques, nous désignerons les six articles du corpus comme suit : articles sur l'IA en anglais (article IA/en) et en français (article IA/fr), articles sur l'apprentissage automatique en anglais (article ML/en) et en français (article ML/fr), articles sur l'apprentissage profond en anglais (article DL/en) et en français (article DL/fr).

– autres : toutes les autres références (notamment un grand nombre de sites web).

Les titres des références, enfin, ont été analysés. Une première étape quantitative s'est appuyée sur des outils de textométrie : le vocabulaire employé fait l'objet d'une racinisation (avec les Stemmers du projet Snowball, utilisés à travers la librairie Python NLTK), permettant d'identifier les termes les plus fréquents (notamment visualisés sous forme de nuages de mots clefs). Une analyse qualitative des titres et résumé des références a également été réalisée afin d'étudier les choix effectués par les contributeurs, à la fois en termes de temporalité, d'origine et de sujet traité.

#### 4. Résultats

Les sections suivantes présentent les résultats obtenus par l'analyse et le croisement des données ainsi obtenues.

##### 4.1. Type de références mobilisées

La distribution des références dans les six articles du corpus (Figure 1) fournit quelques premiers indicateurs intéressants. On constate que, si les articles anglophones incluent presque exclusivement des références en anglais, il n'en est pas de même pour les articles francophones, au sein desquels les références anglaises gardent une place importante, voire prédominante (entre 33,6% et 85,5%). Ces résultats sont cohérents avec les observations d'autres auteurs, selon lesquelles les références anglophones sont largement utilisées dans les versions non-anglophones de Wikipédia, tandis que, à l'inverse, l'encyclopédie anglophone n'accueille que très peu de références non-anglophones (Doutreix, 2020, p. 118).

| Sujet de l'article              |              | Intelligence artificielle |        |          |        | Apprentissage machine |         |          |        | Apprentissage profond |        |          |        |
|---------------------------------|--------------|---------------------------|--------|----------|--------|-----------------------|---------|----------|--------|-----------------------|--------|----------|--------|
| Langue de l'article             |              | anglais                   |        | français |        | anglais               |         | français |        | anglais               |        | français |        |
| Nombre de références            |              | 237                       |        | 214      |        | 134                   |         | 76       |        | 225                   |        | 34       |        |
| Nombre de références par langue | anglais      | 235                       | 99,2 % | 72       | 33,6 % | 134                   | 100,0 % | 65       | 85,5 % | 224                   | 99,6 % | 24       | 70,6 % |
|                                 | français     | 0                         | 0,0 %  | 141      | 65,9 % | 0                     | 0,0 %   | 11       | 14,5 % | 0                     | 0,0 %  | 10       | 29,4 % |
|                                 | allemand     | 2                         | 0,8 %  | 1        | 0,5 %  | 0                     | 0,0 %   | 0        | 0,0 %  | 1                     | 0,4 %  | 0        | 0,0 %  |
| Nombre de références par type   | science      | 98                        | 41,4 % | 20       | 9,3 %  | 64                    | 47,8 %  | 29       | 38,2 % | 167                   | 74,2 % | 21       | 61,8 % |
|                                 | presse       | 58                        | 24,5 % | 84       | 39,3 % | 29                    | 21,6 %  | 10       | 13,2 % | 24                    | 10,7 % | 6        | 17,6 % |
|                                 | livre        | 31                        | 13,1 % | 23       | 10,7 % | 24                    | 17,9 %  | 6        | 7,9 %  | 9                     | 4,0 %  | 1        | 2,9 %  |
|                                 | institutions | 14                        | 5,9 %  | 12       | 5,6 %  | 0                     | 0,0 %   | 1        | 1,3 %  | 0                     | 0,0 %  | 1        | 2,9 %  |
|                                 | autres       | 36                        | 15,2 % | 75       | 35,0 % | 17                    | 12,7 %  | 30       | 39,5 % | 25                    | 11,1 % | 5        | 14,7 % |

**Figure 1.** *Distribution (nombre et proportion) des références bibliographiques, par langue et type de la référence, dans les six articles analysés*

On observe par ailleurs la place très importante des références scientifiques dans les articles DL (74,2% en anglais et 61,8% en français) et ML (47,8% en anglais,

38,2% en français). Ce taux est plus faible pour les articles IA, en particulier pour la version francophone (9,3%), où la part des sources de presse est à l'inverse beaucoup plus élevée que dans les autres cas (39,9%, alors qu'elle n'excède pas 24,5% dans les autres articles). Il apparaît ainsi que l'article IA/fr s'appuie sur un corpus bibliographique assez différent de ceux des autres articles du corpus.

On peut noter la part importante de sources non classées (« autre ») dans les articles IA/fr et ML/fr par rapport aux versions anglophones (le taux est également légèrement supérieur pour DL/fr, mais la différence est peu significative au regard du nombre de références). Bien que ces sources non classées n'aient pas fait l'objet d'un dénombrement précis, on y rencontre un grand nombre de sites web d'acteurs du domaine (entreprises, chercheurs) et d'articles de blogs (s'apparentant parfois à des plateformes de veille ou des médias d'entreprise). Le recours à des sources de ce type, considérées comme « peu fiables » dans les règles de l'encyclopédie (Wikipédia, s. d.-a), montre des différences de pratiques entre les contributeurs francophones et anglophones.

#### 4.3. Terminologie employée dans les sources

L'analyse textométrique des titres des références est également instructive. On constate en premier lieu une forte différence dans le vocabulaire associé aux trois sujets des articles (Figure 2). Les articles ML et DL mobilisent un langage plutôt technique : concepts informatiques (réseaux de neurones, apprentissage, entraînement), méthode scientifique (recherche, modèle, données) et domaines d'application précis (reconnaissance vocale, images, classification, détection). À l'inverse, les titres des références utilisés dans les articles IA comportent davantage de termes liés à des controverses (humain, futur, éthique, danger), à des sujets prospectifs ou larges (robots, emploi) et à des acteurs (Google, Elon Musk, etc.).

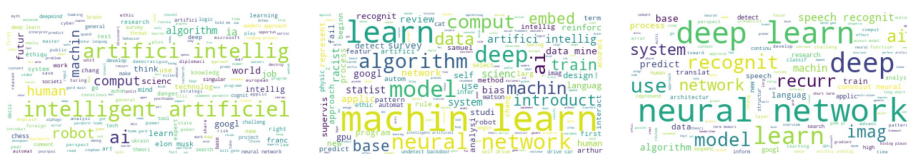
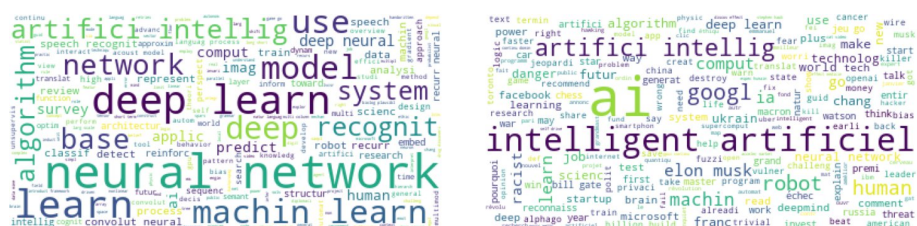


Figure 2. Nuages des mots (racinisés) utilisés dans les titres des références des articles IA (à gauche), ML (au centre) et DL (à droite).

Cette différence s'explique partiellement par celle, plus nette encore, qui apparaît entre les terminologies de la presse et du monde universitaire. Les références journalistiques emploient massivement le terme d'« intelligence artificielle » (ou « IA/AI »), tandis que les références scientifiques emploient plutôt des termes techniques (« modèle », « algorithme », « système »). La presse utilise davantage des verbes d'action, souvent au conditionnel (pourrait, ouvre, étendre, expliquer) et



En revanche, la comparaison entre le vocabulaire des sources de presse mobilisées dans les articles IA/en et IA/fr ne montre guère de différences. On retrouve, dans les deux langues, un discours prospectif, des comparaisons avec l'être humain, l'apparition récurrente de quelques acteurs industriels, autant de thèmes considérés comme des écueils sur lesquels le journalisme bute régulièrement lorsqu'il s'agit de traiter de l'intelligence artificielle (Kapoor & Narayanan, 2022).

L'analyse qualitative des références bibliographiques datant de 2000 à 2022 en fonction de leurs types permet d'affiner les observations précédentes.

S’agissant des articles ML et DL, on constate une certaine homogénéité dans les choix de références effectués par les contributeurs, quelle que soit la langue. Comme cela a été évoqué, les références bibliographiques de type « science » sont majoritaires et portent essentiellement sur des aspects techniques liés à la création, au perfectionnement et à la comparaison de méthodes d’apprentissage. Par exemple, les références scientifiques récentes dans l’article anglophone sur l’apprentissage profond évoquent l’entraînement de réseaux de neurones pour la modélisation acoustique (Tüske *et al.*, 2014) et la reconnaissance de lésions tissulaires (Lyakhov *et al.*, 2022). Du côté de la presse, les références choisies par les contributeurs rapportent des succès scientifiques (e.g. victoires au jeu de Go4) ou présentent de nouvelles applications (e.g. génération d’un manuel scolaires5).

<sup>4</sup> <https://www.scientificamerican.com/article/how-the-computer-beat-the-go-master/>

<sup>5</sup> <https://www.theverge.com/2019/4/10/18304558/ai-writing-academic-research-book-springer-nature-artificial-intelligence>

| Presse                                                                                                                                                                                                                                                                                                                | Science                                                                                                                                                                                                                                                                                                                                                                                                                              |
|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 57 références.<br><i>Sujets</i> : l'IA et la nature humaine, robots, Jeopardy!, dépassement de l'intelligence humaine par l'IA, dangers (humanité, emploi, démocratie), promesses, marché de l'IA, jeu de Go, Deepfake, biais et discriminations, fonctionnement d'IA, enjeux militaires (drones, guerre en Ukraine). | 72 références.<br><i>Sujets</i> : robots, conscience artificielle, théorie de l'esprit, découverte de connaissances, informatique affective, réseau de neurones, système expert, présent et futur de l'IA, robots (Asimov, ingénierie, effets et régulation), apprentissage automatique, apprentissage profond, traitement du langage naturel, dangers (humanité, démocratie, politique), régulation, emploi, étude des conséquences |

**Figure 4.** *Sujets traités par type pour les références de l'article IA/en.*

Concernant les articles IA, il a déjà été observé que la version anglaise comporte une majorité de références scientifiques par rapport à la presse et inversement pour la version française. L'examen des sujets traités par les références montre, pour les deux langues, la présence non négligeable de discours prospectifs évoquant, par exemple, la conscience artificielle, le dépassement de l'intelligence humaine, ou des dangers pour la démocratie (voir Figure 4). Cette observation vaut pour les références de type « presse », dont le biais pour les sujets prospectifs et controversés a déjà été montré (Crépel & Cardon, 2022), mais aussi, de façon plus inattendue, pour les références scientifiques. Certains articles scientifiques choisis par les contributeurs traitent de sujets similaires à ceux de la presse. Par exemple, dans la version anglaise, on trouve les articles de presse « *Is artificial intelligence really an existential threat to humanity?* »<sup>6</sup> et « *Will robots create more jobs than they destroy?* »<sup>7</sup> ainsi que les articles scientifiques « *Malevolent soft power, AI, and the threat to democracy* »<sup>8</sup> et « *AI-deploying organizations are key to addressing 'perfect storm' of AI risks* »<sup>9</sup>. Dans le cas des articles IA, en particulier la version française, les contributeurs ont sélectionné davantage de références liées à des discours prospectifs et controversés, y compris lorsqu'il s'agissait d'articles scientifiques.

## 5. Discussions et conclusion

Cet article s'intéressait à la multiplicité des sens qui sont associés aux concepts de l'IA, tiraillés entre des usages concrets, des discours prospectifs et des controverses variées. Il s'agissait, plus précisément, d'étudier la manière dont les contributeurs de

<sup>6</sup> <https://mambapost.com/2023/04/tech-news/ai-are-an-existential-threat-to-humanity/>

<sup>7</sup> <https://www.theguardian.com/technology/2015/sep/06/will-robots-create-destroy-jobs>

<sup>8</sup> <https://www.brookings.edu/research/malevolent-soft-power-ai-and-the-threat-to-democracy/>

<sup>9</sup> <https://link.springer.com/article/10.1007/s43681-022-00163-7>

l'encyclopédie Wikipédia construisent une « synthèse raisonnée » de ce domaine, à travers l'analyse des références employées dans six articles (trois en anglais, trois en français) relatifs à l'IA.

Il apparaît que les types de références utilisés diffèrent selon les sujets et les langues, avec une prédominance des sources scientifiques pour les articles ML et DL (en anglais comme en français), tandis que les articles IA se distinguent (particulièrement en français) par une forte proportion de sources de presse. Par ailleurs, les différents types de sources ne mobilisent pas le même vocabulaire, avec une place importante des termes relatifs aux controverses et aux discours prospectifs dans les titres des références issues de la presse. L'analyse qualitative du contenu des références montre toutefois que ces thématiques ne sont pas absentes des références scientifiques mobilisées dans les articles IA, en dépit d'un recours à un vocabulaire plus neutre.

Dans ces articles transparaît la grande variété des discours qui circulent dans l'espace public, à travers une large gamme de sources. Selon les sujets, les contributeurs de chaque version linguistique de l'encyclopédie font des choix, parfois différents, qui résultent en des articles aux tonalités distinctes. Coexistent alors plusieurs points de vue, tous solidement sourcés, constituant chacun une « synthèse raisonnée » d'une partie du sujet. Ainsi, face à la polysémie du sujet et de son traitement par diverses sources, la norme de Neutralité de point de vue ne semble pas tant s'appliquer à l'échelle d'un seul article, qui réunirait l'ensemble des points de vue dans des proportions parfaitement représentatives, qu'à l'échelle d'un ensemble d'articles, dont les points de vue se complètent et s'équilibrent entre eux.

Plusieurs éléments doivent toutefois nuancer la lecture de ces résultats. Tout d'abord, l'étude ne porte que sur un nombre très restreint d'articles, ce qui limite la généralisation des observations à d'autres langues ou à davantage de concepts du champ de l'IA. La classification des références constitue également une faiblesse de ces travaux, en raison de la difficulté rencontrée pour tracer et opérationnaliser une stricte délimitation du champ de la presse ou de la science, mais aussi du grand nombre et de la variété des références non classées (catégorie « autres »). Enfin, la méthode employée se concentre sur les références, mais n'examine pas la manière dont elles sont mobilisées dans le corps de l'article ni sur le profil des contributeurs dont les intentions peuvent varier (e.g. vulgarisation, promotion).

Si la prise en compte de ces limites méthodologiques constitue une piste évidente pour approfondir ces résultats, plusieurs études connexes pourraient également s'y adosser. Un travail centré sur l'activité des contributeurs, à travers l'étude des pages de discussions ou des entretiens, permettrait de mieux comprendre la manière dont les différents points de vue peuvent s'exprimer et se concilier dans l'écosystème réglementaire de Wikipédia. L'étude de la présentation d'autres sujets faisant l'objet de controverses, tels que l'énergie nucléaire, la vaccination, les organismes génétiquement modifiés ou les nanotechnologies seraient également intéressantes.

## 6. Bibliographie

- Alexa, « The top 500 sites on the web », s. d., <https://www.alexa.com/topsites>
- Benbouzid, B., & Cardon, D., « Machines à prédire », *Réseaux*, vol 5, n° 211, 2018, p. 9-33.
- Bouchereau, A. & Segault, A. « Évolution des représentations de l'intelligence artificielle : De quelle(s) IA Wikipédia est-elle l'encyclopédie ? », *H2PTM 2021, Information : Enjeux et nouveaux défis*. ISTE Éditions, 2021, p. 216-232.
- Bunz, M., & Braghieri, M. « The AI doctor will see you now : Assessing the framing of AI in news coverage », *AI & SOCIETY*, Springer, vol 37, n°1, 2022, p. 9-22.
- Cao, Y., Li, S., Liu, Y., Yan, Z., Dai, Y., Yu, P. S., & Sun, L., « A Comprehensive Survey of AI-Generated Content (AIGC) : A History of Generative AI from GAN to ChatGPT », 2023, *arXiv*. <https://doi.org/10.48550/arXiv.2303.04226>
- Cardon, D., Cointet, J.-P., & Mazières, A., « La revanche des neurones », *Réseaux*, vol 5, n° 211, 2018, p. 173-220.
- Cardon, D., & Levrel, J., « La vigilance participative. Une interprétation de la gouvernance de Wikipédia », *Réseaux*, n°154, 2009, p.51-89.
- Casebourne, I., Davies, D. C., Fernandes, D. M., & Norman, D. N., *Assessing the accuracy and quality of Wikipedia entries compared to popular online encyclopaedias*. Epic, 2012. [https://upload.wikimedia.org/wikipedia/commons/2/29/EPIC\\_Oxford\\_report.pdf](https://upload.wikimedia.org/wikipedia/commons/2/29/EPIC_Oxford_report.pdf)
- Chuan, C.-H., Tsai, W.-H. S., & Cho, S. Y., « Framing artificial intelligence in American newspapers », *Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society*, 2019, p. 339-344.
- Crépel, M., & Cardon, D., « Robots vs algorithmes. Prophétie et critique dans la représentation médiatique des controverses de l'IA », *Réseaux*, vol 2, n°232-233, 2022, p. 129-167.
- Doutreix, M.-N., *Wikipédia et l'actualité : Qualité de l'information et normes collaboratives d'un média en ligne*. Presses Sorbonne Nouvelle, 2020.
- Floridi, L., Cows, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., Luetge, C., Madelin, R., Pagallo, U., Rossi, F., Schafer, B., Valcke, P., & Vayena, E., « An Ethical Framework for a Good AI Society : Opportunities, Risks, Principles, and Recommendations », Dans L. Floridi (Éd.), *Ethics, Governance, and Policies in Artificial Intelligence*. Springer International Publishing, 2021, p. 19-39.
- Giles, J., « Internet encyclopaedias go head to head », *Nature*, vol 7070, n°438, 2005.
- Kapoor, S., & Narayanan, A., « A checklist of eighteen pitfalls in AI journalism », 2022, <https://www.cs.princeton.edu/~sayashk/ai-hype/ai-reporting-pitfalls.pdf>
- Kyrou, A., « Dans les imaginaires de l'IA », *Multitudes*, n°1, 2020, p. 75-83.
- Langlais, P.-C., « L'émergence d'une norme wikipédienne », Dans L. Barbe, L. Merzeau, & V. Schafer (Éds.), *Wikipédia, objet scientifique non identifié*. Presses Universitaires de Paris Ouest (PUPO), 2015, p. 77-90
- LeCun, Y., *Quand la machine apprend : La révolution des neurones artificiels et de l'apprentissage profond*. Odile Jacob, 2019.

12 Nom de l'ouvrage

- Le Deuff, O., & Roumanos, R., « Enjeux définitionnels et scientifiques de la littérature algorithmique : Entre mécanologie et rétro-ingénierie documentaire », *Tic & société*, vol. 15, n°1-2, 2022, p. 325-360.
- Lyakhov, P. A., Lyakhova, U. A., & Nagornov, N. N., « System for the Recognizing of Pigmented Skin Lesions with Fusion and Analysis of Heterogeneous Data Based on a Multimodal Neural Network », *Cancers*, vol 7, n°14, 2022, p. 1819.
- McDowell, Z. J., & Vetter, M. A., *Wikipedia and the Representation of Reality*. Routledge, 2021.
- Minhaj, M., & Urs, S., « Wikipedia Infoboxes : The Big Data Source for Knowledge Bases behind Alexa and Siri Virtual Assistants », *Information Matters*, vol 1, n°11, 2021.
- O'Neil, C., *Algorithmes : La bombe à retardement*. Les Arènes, 2018.
- Rogers, R., & Sendjarevic, E., « Neutral or National Point of View? A Comparison of Srebrenica Articles across Wikipedia's Language Versions », *Wikipedia Academy: Research and Free Knowledge*, 2012
- Romele, A., « Images of Artificial Intelligence : A Blind Spot in AI Ethics », *Philosophy & Technology*, vol 35, n°1, 2022, p. 4.
- Sahut, G., « L'autorité importée dans Wikipédia : La question de la qualité des sources citées », *Quaderni. Communication, technologies, pouvoir*, n°91, 2016.
- Tüske, Z., Golik, P., Schlüter, R., & Ney, H., « Acoustic modeling with deep neural networks using raw time signal for LVCSR », *Interspeech 2014*, 2014, p. 890-894.
- Vayre, J.-S., & Gaglio, G., « L'intelligence artificielle n'existe-t-elle vraiment pas ? Quelques éléments de clarification autour d'une science controversée », *Diogène*, vol 1-2, n°269-270, 2020, p. 107-120.
- Vergeer, M., « Artificial intelligence in the dutch press : An analysis of topics and trends », *Communication Studies*, vol 3, n°71, 2020, p. 373-392.
- Wikipédia, « Wikipédia:Citez vos sources », s. d.-a, consulté à l'adresse [https://fr.wikipedia.org/wiki/Wikip%C3%A9dia:Citez\\_vos\\_sources](https://fr.wikipedia.org/wiki/Wikip%C3%A9dia:Citez_vos_sources)
- Wikipédia, « Wikipédia:Neutralité de point de vue », s. d.-b, consulté à l'adresse [https://fr.wikipedia.org/wiki/Wikip%C3%A9dia:Neutralit%C3%A9\\_de\\_point\\_de\\_vue](https://fr.wikipedia.org/wiki/Wikip%C3%A9dia:Neutralit%C3%A9_de_point_de_vue)
- Wikipédia, « Wikipédia:Wikipédia est une encyclopédie », s. d.-c, consulté à l'adresse [https://fr.wikipedia.org/wiki/Wikip%C3%A9dia:Wikip%C3%A9dia\\_est\\_une\\_encyclop%C3%A9die](https://fr.wikipedia.org/wiki/Wikip%C3%A9dia:Wikip%C3%A9dia_est_une_encyclop%C3%A9die)
- Zacklad, M., & Rouvroy, A., « L'éthique située de l'IA et ses controverses », *Revue française des sciences de l'information et de la communication*, vol 25, 2022.