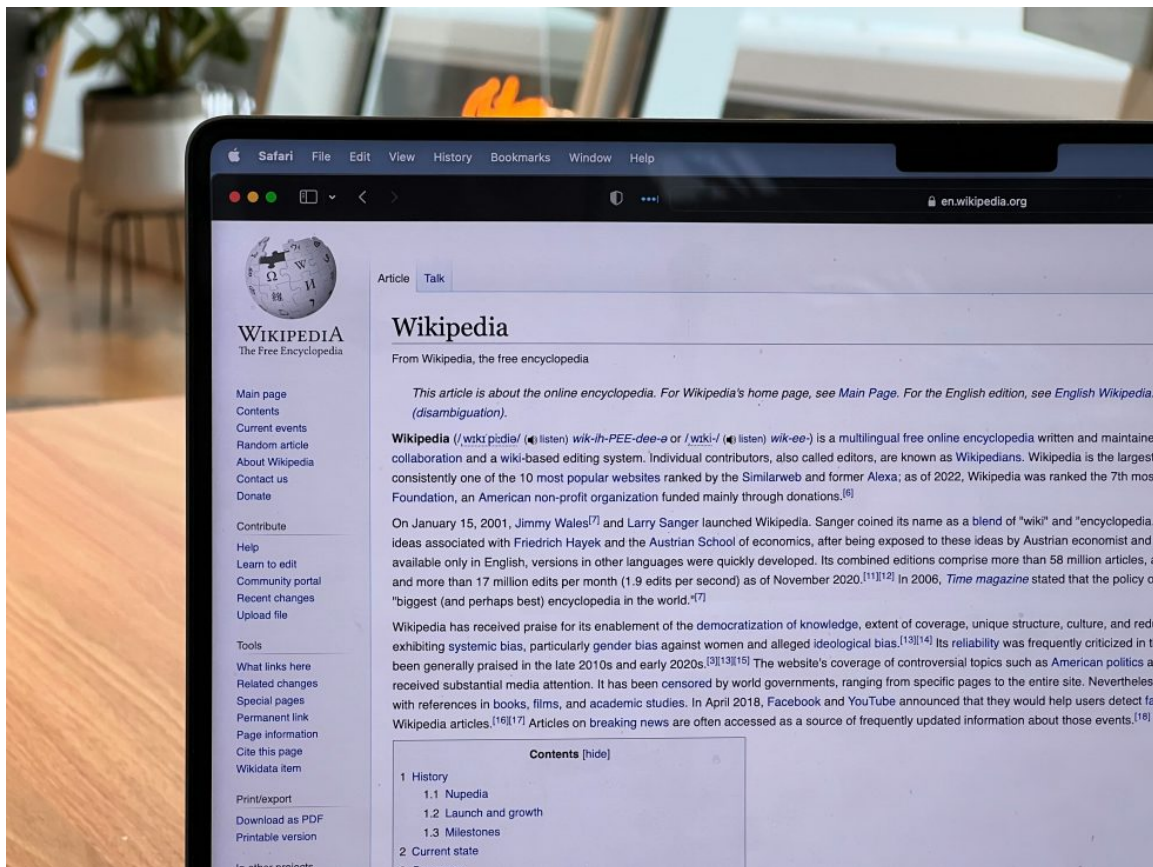


Rapport sur la nouvelle stratégie d'IA de Wikipédia, juillet 2025 à juin 2028

by Alan Colin Arce | 13 October 2026 | French, Observations and Responses



Read in English

Cette observation a été rédigée par Kyle Dase et Alan Colin-Arce.

En un coup d’œil

Topic / Titre	Rapport sur la nouvelle stratégie d’IA de Wikipédia, juillet 2025 à juin 2028
Key Participants / Participants clés	Wikipédia, Fondation Wikimedia
Date / Période	juillet 2025 – juin 2028
Keywords / Mots-clés	open science / science ouverte, open social scholarship / approches sociales des savoirs ouverts, open access / libre accès, policy / politique

Résumé

Cet article rend compte de la nouvelle stratégie de Wikipédia en matière d’intelligence artificielle, en portant une attention particulière à une **déclaration** publiée par Chris Albon, directeur de l’apprentissage automatique, et Leila Zia, directrice de la recherche à la Fondation Wikimedia. Alors que **la stratégie plus large d’IA** présentée par la Fondation Wikimedia offre un aperçu du processus et des alternatives envisagées par la Fondation avant d’arriver à leur trajectoire actuelle, la brève déclaration d’Albon et Zia cristallise les principes et les valeurs en jeu dans leur décision et la nature de l’engagement de Wikipédia avec la technologie.

La stratégie d’IA de Wikipédia

En 2025, la Fondation Wikimedia (WMF), l’organisme sans but lucratif qui héberge Wikipédia, a publié une stratégie à long terme sur la manière dont l’organisation

s'engagerait dans les développements de l'intelligence artificielle générative, décrivant d'abord plusieurs voies possibles et comment chacune s'alignerait ou entrerait en conflit avec les valeurs fondamentales de la Fondation.

En tant que l'un des dix sites les plus populaires sur Internet et peut-être l'exemple d'érudition ouverte le plus connu au monde, la compréhension et l'application de l'IA générative par Wikipédia sont appelées à avoir un impact en tant qu'exemple le plus visible de l'IA et de la politique d'approches sociales des savoirs ouverts.

Bien sûr, la WMF a plus d'une décennie d'expérience dans l'utilisation de l'IA de manière limitée. Wikipédia utilise des robots pour détecter les hyperliens morts, vérifier la qualité des citations et détecter le vandalisme et le pourriel, et a développé des modèles d'apprentissage automatique associés ainsi que des cartes modèles qui stockent les métadonnées essentielles sur un modèle donné. Cependant, ce travail est antérieur à l'avènement des LLM facilement disponibles (tels que ChatGPT), de sorte que la nouvelle stratégie se concentre sur la manière dont la WMF s'engagera avec les LLM.

L'IA a rapidement modifié le paysage Internet et Wikipédia a déjà dû s'adapter : l'an dernier, Wikipédia a noté une croissance de 50 % de son utilisation de la bande passante entre janvier 2024 et avril 2025 (Edwards, 2025) : les entreprises formaient les technologies de l'IA en grattant le site (comme nous l'avons déjà évoqué sur l'OSPO dans nos articles sur le [commonswashing](#) et les [robots IA](#)). La réponse de Wikipédia a été un [partenariat avec Kaggle](#) (une plateforme de science des données et une communauté en ligne appartenant à Google), qui héberge désormais une version allégée de Wikipédia basée sur JSON en anglais et en français, conçue pour le scraping de l'IA, disponible sous licence Creative Commons. Ce partenariat permet aux développeurs d'IA de télécharger un ensemble de données structurées et préanalysées contenant plus de 40 Go de données provenant d'articles Wikipédia en anglais et en français. Selon les statistiques de Kaggle, [l'ensemble de données](#) a été téléchargé plus de 16 000 fois en juin 2026.

La position la plus forte de la nouvelle stratégie est peut-être que **Wikipédia n'utilisera pas l'IA générative pour créer du contenu**. La WMF a décidé de ne pas investir dans la génération de connaissances sur l'IA ; leur document stratégique admet qu'un tel pivot aurait pu être attrayant, mais la liste des inconvénients potentiels est trop longue pour être ignorée :

[C]ontrôle humain limité, vulnérabilité aux préjugés, aux hallucinations, potentiel de propagation de la désinformation et de la désinformation, contexte local limité, travail humain invisible important et faible capacité à traiter des sujets nuancés. (Méta-Wiki)

Chacun de ces problèmes sape l'engagement humain au cœur de Wikipédia, et WMF craint que l'intégration de l'IA pour produire du contenu ne dissuade également les bénévoles au cœur de l'organisation.

Dans le même temps, **les dirigeants de la WMF ne pensent pas non plus que le simple fait d'ignorer les développements récents de la technologie de l'IA générative servira au mieux la communauté.** Au-delà du simple fait de manquer des occasions de réduire la charge de travail des éditeurs, cette stratégie ne refléterait pas la manière dont Internet continue d'évoluer et de se développer.

Cela désavantagerait aussi les humains au centre de Wikipédia. Comme l'expliquent les dirigeants de la WMF, « les progrès de l'IA ont facilité la génération de contenu » et « le coût de la vérifiabilité reste élevé pour les éditeurs », les exposant « à un risque important de surcharge et d'épuisement professionnel » (WMF). Reconnaisant que Wikipédia évolue normalement plus lentement que le reste d'Internet, les dirigeants du WMF craignent que ne pas s'engager de manière plus significative dans l'IA mettra Wikipédia à la traîne et risquera de perdre de sa pertinence à long terme.

Au lieu de cela, la Fondation Wikimedia a choisi « d'investir dans l'IA pour soutenir les éditeurs dans les domaines où [elle] peut avoir un avantage unique sur les autres technologies pour résoudre les problèmes d'impact et donner la priorité à l'action des éditeurs dans l'interaction avec l'IA » (WMF). Cette décision est motivée en partie par l'anticipation selon laquelle Internet regorgera bientôt de contenus trompeurs et de mauvaise qualité et il existe un besoin important de « connaissances vérifiables et de haute qualité » (WMF). Mais qu'implique exactement cette stratégie ?

Une approche « l'humain d'abord » : des éléments exploitables

La WMF décrit certaines tâches spécifiques pour lesquelles l'IA peut aider sur Wikipédia. Plusieurs impliquent l'optimisation des flux de travail grâce à l'aide de l'IA pour réduire le travail requis des modérateurs, des éditeurs et des patrouilleurs. L'idée centrale derrière ce qu'Albon et Zia appellent une approche « l'humain d'abord » est que l'IA peut être utilisée pour limiter certaines tâches monotones et chronophages afin que les êtres humains puissent consacrer plus de temps à discuter, à prendre des décisions et à produire le contenu dont dépend Wikipédia. Les tâches principales de l'IA dans cette approche centrée sur l'humain consistent à intégrer de nouveaux éditeurs, rationaliser les flux de travail et aider à la modération et à la patrouille. Malgré les promesses de prioriser la transparence, les technologies libres et le multilinguisme, le WMF n'a pas précisé comment il atteindra ces priorités ni comment il mesurera l'efficacité de l'optimisation des flux de travail.

IA (vraiment) ouverte ?

Un principe fondamental important derrière la nouvelle stratégie de la Fondation Wikimedia est la promesse de « donner la priorité à l'utilisation de technologies d'IA open source ou de pondérations ouvertes, et [...] de développer uniquement une IA open source » (Meta-Wiki Contributors, 2026). Cependant, les dirigeants de la WMF admettent qu'ils ont dû faire des compromis dans le cadre de cette nouvelle stratégie. Le compromis le plus difficile est peut-être leur choix « d'utiliser des modèles de pondération ouverts lorsque cela est nécessaire pour créer des fonctionnalités destinées à prendre en charge les éditeurs » (WMF).

Contrairement aux modèles open source qui assurent la transparence à tous les niveaux, les LLM à poids ouverts cachent toujours les données de formation et certains processus/méthodologies. Alors que les dirigeants de la WMF « espèrent que des modèles fondamentaux open source capables de rivaliser en matière d'évaluation des meilleures pratiques seront publiés à l'avenir », ils déplorent que Wikipédia n'ait tout simplement pas les ressources nécessaires pour développer un modèle fondamental open source à partir de zéro. Cette situation révèle les obstacles et les limites du développement de modèles d'IA open source, même lorsque les organisations disposent de données de haute qualité, un paradoxe qui pourrait être analysé plus en détail au sein de la communauté des chercheurs ouverts.

Politique contre pratique

Malgré les principes clairs énoncés dans la stratégie d'avril, la Fondation Wikimedia testait les « résumés d'articles simples », des résumés d'articles générés par l'IA, dès juin 2025. Ces résumés ont immédiatement suscité de vives critiques de la part de la communauté, et les résumés ont été interrompus au cours du même mois (Wikipédia). Au lieu de cela, en août 2025, Wikipédia a développé une politique pour l'identification, la nomination et la suppression ultérieure du contenu généré par l'IA sur Wikipédia. L'organisation a également rédigé un guide connexe sur les « Signes de l'écriture de l'IA » destiné à aider les réviseurs, les éditeurs et les lecteurs à identifier le contenu de l'IA. La nature proactive de la communauté Wikipédia et son ouverture ont aidé ses membres dans leur ensemble à façonner les politiques et les prises de décision liées à l'IA.

Questions et considérations clés

En résumé, la nouvelle stratégie de Wikipédia exploite et donne la priorité à l'IA générative open source (ou du moins au modèle à poids ouverts) pour soutenir ses éditeurs, réduisant ainsi la charge de travail des éditeurs afin qu'ils puissent consacrer plus de temps à s'engager dans les tâches humaines qui défendent les valeurs de l'organisation à but non lucratif. La WMF donnera la priorité à la qualité et

à l'intégrité du contenu plutôt qu'à la quantité, et l'utilisation de l'IA sera limitée à des domaines d'impact distincts, principalement la charge de travail éditoriale. Ils ont l'intention d'utiliser l'IA open source là où ils le peuvent, mais ne peuvent pas se permettre de développer leur propre modèle et admettent qu'ils utiliseront également des modèles open source à l'avenir. Ce compromis s'accompagne d'un manque de clarté concernant les projets de traduction automatique de Wikipédia, malgré les promesses d'une « approche nuancée ».

L'introduction des politiques d'IA de Wikipédia survient à un moment où les chercheurs mettent en garde contre une baisse du nombre de lecteurs et de contributeurs humains de Wikipédia. Dans un article intitulé *Death by AI: Will large Language Models diminish Wikipedia?* (La mort par l'IA: Est-ce que les LLM vont diminuer Wikipédia ?), Wagner et Jiang (2026) soutiennent que « les progrès des grands modèles de langage (LLM) et de l'intelligence artificielle (IA) générative diminueront la valeur de Wikipédia, en raison d'un retrait des producteurs de contenu humain, qui retiendront leurs efforts, percevant moins de besoin de leurs efforts et une « concurrence accrue en matière d'IA » » (p. 743). Si Wikipédia reçoit moins de visiteurs humains, le bassin de contributeurs potentiels diminuera probablement aussi, dans la mesure où les Wikipédiens commenceront comme lecteurs avant de modifier le site. Wagner et Jiang prédisent également que d'ici 5 à 8 ans, plus de la moitié de l'activité en ligne de Wikipédia sera réalisée par l'IA, ce qui pourrait amener les utilisateurs humains à quitter le site en raison d'une diminution de la reconnaissance.

Cependant, une autre étude examinant les changements dans l'utilisation et la contribution de Wikipédia après le lancement de ChatGPT n'a révélé aucune baisse globale du nombre de pages vues, du nombre de visiteurs, du nombre de modifications et de l'activité des éditeurs dans 12 langues (Reeves et al., 2025). Bien qu'il y ait eu une augmentation de l'utilisation et des contributions de Wikipédia dans presque toutes les langues analysées, les augmentations les plus importantes se sont produites dans les langues où ChatGPT n'était pas disponible. Par conséquent, l'impact des LLM sur les utilisateurs et les éditeurs de Wikipédia reste une question ouverte.

Bien que la WMF semble avoir changé de cap en ce qui concerne le contenu de l'IA assez soudainement avec les prototypes de résumés d'articles simples, les commentaires de la communauté ont amené l'organisation à corriger rapidement son

cap et à s'en tenir à ses principes de garantie de l'intégrité du contenu. En bref, l'approche centrée sur l'humain s'est avérée aussi efficace que les personnes qui la composent.

Avis : Alan Colin-Arce travaille présentement pour Wikimedia Canada. Ce message a été rédigé avant qu'il n'occupe son poste chez Wikimedia Canada et il ne représente pas les opinions de l'organisation.

Références

Albon, C., & Zia L. (2025). "Our new AI strategy puts Wikipedia's humans first." *Wikimedia Foundation*. <https://wikimediafoundation.org/news/2025/04/30/our-new-ai-strategy-puts-wikipedias-humans-first/>

"Artificial intelligence in Wikimedia Projects." https://en.wikipedia.org/wiki/Artificial_intelligence_in_Wikimedia_projects#:~:text=Artificial%20intelligence%20and%20machine%20learning,others%20for%20the%20Wikipedia%20community.

Edwards, B. (2025). *AI bots strain Wikimedia as bandwidth surges 50%*. Ars Technica. <https://arstechnica.com/information-technology/2025/04/ai-bots-strain-wikimedia-as-bandwidth-surges-50/>

Maxwell, T. (2025). "Wikipedia Is Making a Dataset for Training AI Because It's Overwhelmed by Bots." *Gizmodo*. <https://gizmodo.com/wikipedia-is-making-a-dataset-for-training-ai-because-its-overwhelmed-by-bots-2000590704>

Meta-Wiki Contributors. (2026). "Strategy/Multigenerational/Artificial intelligence for editors." *Meta-Wiki*. https://meta.wikimedia.org/w/index.php?title=Strategy/Multigenerational/Artificial_intelligence_for_editors&oldid=29927690

Reeves, N., Yin, W., & Simperl, E. (2025). Exploring the impact of ChatGPT on Wikipedia engagement. *Collective Intelligence*, 4(3), 26339137251372599. <https://doi.org/10.1177/26339137251372599>

Wagner, C., & Jiang, L. (2025). Death by AI: Will large language models diminish Wikipedia? *Journal of the Association for Information Science and Technology*, 76(5), 743–751. <https://doi.org/10.1002/asi.24975>

“Wikipedia Kaggle Dataset using Structured Contents Snapshot.” *Wikimedia Enterprise*. <https://enterprise.wikimedia.com/blog/kaggle-dataset/>

Woodcock, C. (2023). “AI Is Tearing Wikipedia Apart.” *VICE*. <https://www.vice.com/en/article/ai-is-tearing-wikipedia-apart/>

Search



Archives



Categories

Community News

English

French

Insights and Signals Reports

Observations

Observations and Responses

Policies

Portuguese

Responses

Spanish

Tags

[AI bots / Robots d'indexation IA](#) [AI governance / Gouvernance de l'IA](#)

[AI safety / Sécurité de l'IA](#) [Berlin Declaration / Déclaration de Berlin](#)

[Bethesda Statement / Déclaration de Bethesda](#) [bibliodiversity / bibliodiversité](#)

[Bilingualism / bilinguisme](#) [Budapest Statement / Déclaration de Budapest](#) [Canada](#)

[Canadiana.org](#) [Canadian government/le gouvernement du Canada](#) [CAPOS](#)

[CARL / ABRC](#) [collaboration](#) [commonswashing](#)

[community engagement / engagement communautaire](#)

[Compute Canada / calcul Canada](#) [copyright / droits d'auteurs](#)

[credibility / crédibilité](#) [CRKN / RCDR](#) [Cybersecurity / Cybersécurité](#)

[data management / gestion des données](#)

[diamond open access / le libre accès diamant](#)

[diamond open access / libre accès diamant](#) [Digital Commons / Commun numérique](#)

[digital scholarship / version numérique](#) [en français / French](#) [English / en anglais](#)

[events and gatherings / événements et rassemblements](#)

[Federation for the HSS / Fédération des sciences humaines](#)

[funding agencies / organismes de financement](#)

[generative artificial intelligence / l'intelligence artificielle générative](#)

[identity management / gestion de l'identité](#) [implementation / mise en oeuvre](#)

[Indigenous Knowledge / Le savoir autochtone](#) [INKE](#)

[International OA Week / Semaine internationale du libre accès](#)

[international policy / politique internationale](#)

[Knowledge diversity / Diversité des connaissances](#)

[Knowledge diversity / Diversité des savoirs](#)

[Knowledge mobilization / Mobilisation des connaissances](#)

[licensing agreements / accords de licence](#) [Multilingualism / Multilinguisme](#)

[Naylor Report / le rapport Naylor](#) [open access / libre accès](#)

[open data / données ouvertes](#) [open education / éducation ouverte](#)

[open government / gouvernement ouvert](#) [open infrastructure / infrastructure ouverte](#)

[Open Scholarship Press](#) [open science / science ouverte](#)

[open social scholarship / approches sociales des savoirs ouverts](#)

[open source software / les logiciels libres](#) [ORCID](#) [peer review / critique des pairs](#)

[Perpetual Access / Accès perpétuel](#) [PKP](#) [Plan S](#)

[Plan S update / mise à jour du Plan S](#) [policy / politique](#)

[policy guide / guide des politiques](#) [promotion et titularisation](#) [publishing / édition](#)

[RDC / DRC](#) [RDM](#) [RECODE](#) [recommendations / recommandations](#)

[reports / les rapports](#) [repositories / les dépôts](#)

[research creation / recherche-crédation](#)

[research evaluation / l'évaluation de la recherche](#)

[research libraries / les bibliothèques de recherche](#)

[research output / les résultats de la recherche](#)

[research security / sécurité de la recherche](#) [RPT / révision](#)

[scholarly communication / communication savante](#)

[scholarly communication / la communication savante](#) [SFU Library / Bibliothèque](#)

[social media / les médias sociaux](#) [Tri-Agency / des trois organismes](#) [UK](#)

[UK / Royaume-Uni](#) [UNESCO](#) [UVic Libraries](#) [Érudit](#)



This work is licensed under a [Creative Commons Attribution-ShareAlike 4.0 International License](#).

