



L'IA générative et l'édition savante

by Britt Amell | 10 December 2024 | French, Insights and Signals Reports



Lisez-le en anglais

Ce rapport sur les tendances et perspectives (“Insights and Signals Report”) a été rédigé par Brittany Amell, avec des remerciements à John Willinsky, John Maxwell et William Bowen pour leurs commentaires et leurs contributions.

En bref

Perspectives
et signaux
Domaine

L'IA générative et l'édition savante

thématique	
Principaux participants	Ithaka S+R, cOAlition S, DOAJ, Direction générale de la recherche et de l'innovation de la Commission européenne, Public Knowledge Project
Cadre temporel	2022 – aujourd'hui
Mots clés ou thèmes principaux	IA générative, bourses d'études ouvertes, confiance, crédibilité, libre accès

Résumé

Les rapports “*Policy Insights and Signals*” scrutent l’horizon afin d’identifier et d’analyser les tendances émergentes et les signaux précurseurs susceptibles d’influer sur les orientations politiques futures en matière de libre accès et d’érudition ouverte et sociale. Ils ont tendance à mettre en évidence les changements dans la technologie, l’opinion et les sentiments du public, et/ou les changements réglementaires à l’intérieur et à l’extérieur du Canada. Tout comme les observations politiques de l’OSPO, les rapports sur les perspectives et les signaux visent à aider les partenaires à élaborer des stratégies proactives, réactives et tournées vers l’avenir.

Ce rapport sur les perspectives et les signaux poursuit l’examen par l’OSPO de l’évolution du dialogue sur les implications de l’IA générative pour les bourses d’études ouvertes et l’édition en libre accès. “L’IA générative fait référence à une classe d’algorithmes qui guident la création de divers types de contenu (Dobrin, 2023). Parfois évoquée dans le même temps que ChatGPT, DALL-E 2 ou Gemini, nous commençons déjà à voir l’impact de l’IA générative sur l’édition savante, malgré son introduction relativement récente.

Compte tenu de l’intérêt croissant pour les implications de l’IA générative pour le libre accès et la communication savante en général, ce rapport s’articule autour des discussions récentes sur les risques et les opportunités potentiels pour le secteur. Vous êtes intéressé par d’autres rapports “Insights and Signals” consacrés à l’IA ?

Vous les trouverez [ici](#) ou [ici](#).

Les points abordés dans ce rapport sont les suivants :

- Une **annonce récente** d'Ithaka S+R concernant un nouveau projet de recherche axé sur l'IA générative et l'édition savante.
- Un **article invité de** Shen et Ball (DOAJ) sur le Scholarly Kitchen concernant l'augmentation du nombre d'articles rétractés et la crise de confiance envers les revues en libre accès dans un contexte d'innovations en matière d'IA générative.
- Une **proposition de** la cOAlition S concernant l'édition responsable
- Les applications potentielles des principes FAIR à l'élaboration de politiques en réponse à l'IA générative, ainsi que le chevauchement entre ces principes et ceux mentionnés dans un **récent guide sur l'utilisation responsable de l'IA** dans la recherche publié par la direction générale de la recherche et de l'innovation de la Commission européenne.
- Le **Publication Facts Label** (PFL – également : [ici](#)), ainsi qu'une **annonce récente du Public Knowledge Project** concernant l'essai du PFL pour les revues fonctionnant avec OJS (v 3.3 ou plus).

Remarque : Nous vous présentons nos excuses en avance, car la plupart des sources mentionnées dans ce rapport sont disponibles en anglais uniquement.

Le groupe de recherche sur l'édition savante annonce un nouveau projet explorant les implications de l'IA générative

Le groupe de recherche **Ithaka S+R** a récemment annoncé son **intention d'entreprendre un nouveau projet de recherche axé sur l'IA générative et l'édition savante** : “L'évolution rapide des besoins et des attentes des utilisateurs, le potentiel de l'IA générative pour atténuer les défis systémiques tenaces dans le secteur de l'édition savante, et la prise de conscience des risques que l'IA générative fait peser sur les connaissances des experts exigent que nous trouvions le temps de réfléchir en profondeur à ce que l'IA générative *signifie* pour l'édition savante comme secteur et à sa valeur en tant que composante de l'infrastructure partagée qui soutient la recherche et les communications savantes et scientifiques”

(Ruediger et Bergstrom, 2024). Ce projet examinera les opportunités, les risques et les implications stratégiques que l'IA générative pourrait avoir pour l'édition savante. Il s'appuie sur l'étude d'Ithaka S+R de 2023, "[The Second Digital Transformation of Scholarly Publishing](#)", qui examinait les besoins partagés en matière d'infrastructures de communication savante à la lumière des transformations numériques (Bergstrom, Rieger et Schonfeld, 2024).

Comme l'écrivent Shen et Ball (2024) [dans leur article invité pour The Scholarly Kitchen](#), bien que les menaces à la confiance et à la crédibilité aient toujours été au centre des préoccupations d'organisations telles que le Directory of Open Access Journals (DOAJ), ces menaces sont devenues encore plus urgentes. Constatant une augmentation du nombre d'articles rétractés, qui a atteint le chiffre record de 10 000 en 2023, Shen et Ball (2024) expliquent qu'il est devenu nécessaire de prendre des mesures supplémentaires pour préserver la confiance et la crédibilité à l'ère de l'IA générative.

Pour le DOAJ, l'une de ces mesures a consisté à former une équipe chargée d'enquêter sur les "cas présumés de pratiques douteuses". Ces pratiques douteuses sont signalées soit par des membres de la communauté du DOAJ au sens large, soit par ceux qui participent au processus d'évaluation de la demande d'inclusion d'une revue dans le DOAJ. Une fois le signalement effectué, les membres de l'équipe examinent de près les articles publiés par la revue, ainsi que la composition et la compétence de son comité de rédaction, ses pratiques d'évaluation par les pairs et d'autres facteurs.

"Comme les pratiques prédatrices continuent d'évoluer, nos enquêtes deviennent de plus en plus complexes", écrivent Shen et Ball (2024). "Nous consultons parfois des experts externes en la matière pour obtenir leurs conseils. Rien qu'en 2023, nous avons mené un total de 409 enquêtes sur des revues et des éditeurs, dont beaucoup ont abouti à des exclusions du DOAJ d'au moins un an."

Aider les éditeurs et le public à évaluer la crédibilité d'une publication est une chose à laquelle les membres du Public Knowledge Project (PKP) ont beaucoup réfléchi. Le Publication Facts Label, conceptualisé à l'origine par John Willinsky (fondateur du Public Knowledge Project et présenté plus en détail [ici](#) et [ici](#)), en est un exemple. Basé sur le Nutrition Facts Label – ce tableau bien connu que l'on trouve sur les emballages alimentaires au Canada et aux États-Unis – le Publication Facts Label (ou

PFL en abrégé) regroupe huit normes en un guide facile à consulter que les plateformes d'édition peuvent utiliser pour présenter l'intégrité d'une publication à un large public (Willinsky et Pimental). PKP **a récemment annoncé** la mise à l'essai du Publication Facts Label pour les revues utilisant l'Open Journal Systems (v 3.3 ou plus). En installant un plug-in, comme expliqué **ici**, le PFL peut être affiché automatiquement sur la page d'accueil d'un article.

La cOAlition S élabore une proposition de publication responsable

Alors qu'Ithaka S+R se prépare à avancer dans son projet, la **cOAlition S** (associée au **Plan S**) a terminé la phase de consultation de sa proposition intitulée “**Vers une édition responsable**”.

Poussés par le même “devoir de diligence pour le bon fonctionnement du système scientifique” qui a inspiré le Plan S, les bailleurs de fonds qui forment la cOAlition S explorent maintenant une nouvelle vision de la communication savante ; une vision qui promet d'être plus efficace, plus abordable et plus équitable, pour finalement bénéficier à la société dans son ensemble”, écrivent Bodo Stern (chef des initiatives stratégiques à l'Howard Hughes Medical Institute) et Johan Rooryck (directeur exécutif de la cOAlition S) dans un **billet de blog annonçant la proposition**. La cOAlition S a révisé la proposition sur la base des commentaires reçus entre novembre 2023 et avril 2024, et a partagé la proposition révisée avec une réunion des bailleurs de fonds du cOAlition S en juin.

Les réactions à la proposition ont été rendues publiques par cOAlitionS **ici**.

La proposition “Vers une publication responsable” met en avant un ensemble de principes qui peuvent être utilisés pour guider les décisions concernant la manière de soutenir “la diffusion de la recherche de manière responsable, équitable et durable” (Stern et al. 2023, 2) – une phrase qui évoque les priorités clés nommées dans l'**Agenda de la science ouverte 2025 de l'Association européenne des universités**. L'Open Science Agenda de l'Association, qui a été publié en février 2022 en prévision de la conférence et de l'assemblée générale qui se tiendront en 2025-2026, énonce plusieurs priorités et objectifs, dont celui de faire en sorte que

toutes les universités européennes fassent partie d'un "écosystème d'édition savante juste" d'ici à 2025 (Gaillard, 2022).

Vinciane Gaillard (directrice adjointe pour la recherche et l'innovation, Association européenne des universités) décrit un écosystème d'édition savante juste comme étant "transparent, diversifié, économiquement abordable et durable, techniquement interopérable et piloté par la communauté des chercheurs et ses institutions par le biais de politiques coordonnées" (diapositive 4).

Elsevier lance ScopusAI

Au début de l'année 2024, Elsevier a annoncé le lancement de ScopusAI, un outil d'intelligence artificielle proposé aux institutions sur une base d'abonnement. Selon le site web d'Elsevier, ScopusAI sert de "guide expert" que les chercheurs peuvent utiliser pour naviguer dans "la vaste étendue de la connaissance humaine dans Scopus". En plus de résumer la littérature disponible dans Scopus, l'outil est apparemment aussi capable de "repérer" ce qu'Elsevier appelle "l'espace blanc" dans la littérature afin que les chercheurs puissent ostensiblement mieux identifier les types de contributions qu'ils peuvent apporter. Il est inquiétant de constater que les résumés générés par l'outil renvoient à ce qu'Elsevier et l'outil ScopusAI considèrent comme des documents "fondamentaux" pour un sujet donné – il s'agit "d'articles à fort impact qui sont le plus souvent cités par les articles utilisés dans les résumés". Outre les décisions algorithmiques concernant les articles considérés comme fondamentaux, l'outil offre également la possibilité de "découvrir des experts" dans un domaine. Cependant, l'outil ne prend en compte que les articles et les profils qui figurent déjà dans Scopus qui est déjà connu comme une base de données qui surreprésente les universitaires et les chercheurs d'Europe, d'Océanie et d'Amérique du Nord par rapport aux universitaires et aux chercheurs d'autres régions du monde (Asubiaro et al. 2024).

Questions et considérations clés

Les idées et les signaux décrits ci-dessus et ailleurs indiquent que l'intégration de l'IA générative dans l'édition savante présente à la fois des opportunités et des

défis, ainsi que des implications significatives pour la qualité, l'intégrité et l'accessibilité des résultats de l'édition savante. Plusieurs questions se posent pour l'élaboration des politiques, notamment

- Quelle est la place éventuelle de l'IA générative dans l'édition en libre accès ?
- Comment les éditeurs en libre accès peuvent-ils guider l'utilisation responsable, éthique et crédible de l'IA générative ?

Par exemple, John Willinsky (fondateur du Public Knowledge Project) nous a fait savoir que le PKP étudie activement la manière dont l'IA basée sur de grands modèles linguistiques pourrait contribuer à la durabilité de l'édition en libre accès :

L'objectif principal de ce travail est d'établir si les LLM peuvent être suffisamment adaptés pour automatiser de manière fiable le balisage HTML et JATS XML des manuscrits des auteurs (étant donné que ce balisage nécessite actuellement des compétences techniques ou des paiements qui dépassent la capacité de la plupart des revues Diamond OA). Ce travail a atteint le stade initial de la preuve de concept, et les travaux se poursuivent sur sa valeur comparative (par rapport à d'autres outils) et sur les moyens d'intégrer et de maintenir un tel service de balisage dans le flux de travail éditorial. (Lisez l'intégralité du commentaire de John Willinsky ci-dessous).

En plus de ces questions, les lecteurs pourraient également se demander quelles leçons et perspectives clés du mouvement de l'accès libre, du discours, de la recherche et de la littérature pourraient être appliquées au paysage en évolution de l'IA générative, le cas échéant. Par exemple, les principes FAIR pour la gestion et l'intendance des données ([Wilkinson et al. 2016](#)) – envisagés à l'origine comme un moyen de soutenir la réutilisation des données savantes en veillant à ce qu'elles soient trouvables, accessibles, interopérables et réutilisables – pourraient servir d'exemple. Les principes FAIR ont été réinterprétés pour s'appliquer aux logiciels, aux flux de travail, aux outils, aux algorithmes et, de plus en plus, aux modèles d'IA ([Huerta et al. 2023](#)). Dans le même ordre d'idées, la Direction générale de la recherche et de l'innovation de la Commission européenne a récemment **publié un ensemble de principes visant à orienter l'utilisation responsable de l'IA générative dans la recherche**. On peut dire que ces principes recoupent les principes FAIR, ce qui offre d'autres possibilités de réflexion. (Pour mémoire, les quatre principes clés sont la fiabilité, l'honnêteté, le respect et la responsabilité).

Commentaires du partenariat INKE

Réponse de John Willinsky (fondateur, Public Knowledge Project) :

Bien qu'il y ait des raisons de s'inquiéter des récentes avancées de l'IA, les universitaires ont également la responsabilité d'explorer les contributions et les avancées potentielles de l'IA pour la recherche et l'érudition. Depuis un certain temps, le **Public Knowledge Project** se tourne vers l'IA pour résoudre les problèmes urgents liés à l'équité et à la qualité des ressources dans les communications savantes, avec un succès limité. Il mène actuellement des recherches sur la capacité des grands modèles de langage à relever le défi de longue date que représente le développement d'un moyen durable pour les revues Diamond OA de publier dans les formats standard HTML et PDF, ainsi que d'exporter des fichiers en JATS XML. L'objectif principal de ce travail est d'établir si les LLM peuvent être suffisamment adaptés pour automatiser de manière fiable le balisage HTML et JATS XML des manuscrits des auteurs (étant donné que ce balisage nécessite actuellement des compétences techniques ou des paiements qui dépassent la capacité de la plupart des revues Diamond OA). Ce travail a atteint le stade initial de la preuve de concept, et les travaux se poursuivent sur sa valeur comparative (par rapport à d'autres outils) et sur les moyens d'intégrer et de maintenir un tel service de balisage dans le flux de travail éditorial.

Références

Ahari, Juni. 2024. "Generative AI and Scholarly Publishing". *Ithaka S+R* (blog). 23 avril 2024. <https://sr.ithaka.org/blog/generative-ai-and-scholarly-publishing/>.

Asubiaro, Toluwase, Sodiq Onaolapo et David Mills. 2024. "Regional disparities in Web of Science and Scopus journal coverage." *Scientometrics* 129 (3) : 1469–91. <https://doi.org/10.1007/s11192-024-04948-x>.

Bergstrom, Tracy, Oya Y. Rieger et Roger C. Schonfeld. 2024. "The Second Digital Transformation of Scholarly Publishing: Strategic Context and Shared Infrastructure." *Ithaka S+R*. <https://doi.org/10.18665/sr.320210>.

Chiarelli, Andrea, Ellie Cox, Rob Johnson, Ludo Waltman, Wolfgang Kaltenbrunner,

André Brasil, Andrea Reyes Elizondo et Stephen Pinfield. 2024. “Towards Responsible Publishing”: Findings from a global stakeholder consultation. cOAlition S. Zenodo. <https://doi.org/10.5281/zenodo.11243942>.

Directorate-General for Research and Innovation. 2024. “Living Guidelines on the Responsible Use of Generative AI in Research (Version 1).” Brussels: European Commission. https://research-and-innovation.ec.europa.eu/document/download/2b6cf7e5-36ac-41cb-aab5-0d32050143dc_en?filename=ec_rtd_ai-guidelines.pdf.

Dobrin, Sidney I. 2023. “Talking about Generative AI: A Guide for Educators.” Version 1. Broadview Press. <https://sites.broadviewpress.com/ai/talking/>.

Gaillard, Vinciane. 2022. “Encouraging/Supporting Sustainability in the Diamond Action Plan Community.” Presented at the 2022 Diamond Open Access Conference, September. <https://www.scienceeurope.org/media/yg3ho4tp/doa-conf-vinciane-gaillard.pdf>

Huerta, E. A., Ben Blaiszik, L. Catherine Brinson, Kristofer E. Bouchard, Daniel Diaz, Caterina Doglioni, Javier M. Duarte, et al. 2023. “FAIR for AI: An Interdisciplinary and International Community Building Perspective.” *Scientific Data* 10 (1): 487. <https://doi.org/10.1038/s41597-023-02298-6>.

Shen, Cenyu, and Joanna Ball. 2024. “DOAJ’s Role in Supporting Trust in Scholarly Journals: Current Challenges and Future Solutions.” *The Scholarly Kitchen* (blog). June 6, 2024. <https://scholarlykitchen.sspnet.org/2024/06/06/guest-post-doajs-role-in-supporting-trust-in-scholarly-journals-current-challenges-and-future-solutions/>.

Stern, Bodo, et Johan Rooryck. 2023. “Introducing the ‘Towards Responsible Publishing’ Proposal from cOAlition S | Plan S.” sOApbox : A Plan S Blog (blog). 31 octobre 2023. <https://www.coalition-s.org/blog/introducing-the-towards-responsible-publishing-proposal-from-coalition-s/>.

Stern, Bodo, Zoé Ancion, Andreas Björke, Ashley Farley, Marte Qvenild, Katharina Rieck, Jeroen Sondervan, et al. 2023. “Towards Responsible Publishing : Seeking Input from the Research Community to a Draft Proposal from cOAlition S”, octobre. <https://doi.org/10.5281/ZENODO.8398480>.

Wilkinson, Mark D., Michel Dumontier, IJsbrand Jan Aalbersberg, Gabrielle Appleton, Myles Axton, Arie Baak, Niklas Blomberg, et al. 2016. “The FAIR Guiding Principles for Scientific Data Management and Stewardship.” *Scientific Data* 3 (1): 160018. <https://doi.org/10.1038/sdata.2016.18>.

Search



Archives



Categories

Community News

English

French

Insights and Signals Reports

Observations

Observations and Responses

Policies

Responses

Uncategorized

Tags

[Berlin Declaration / Déclaration de Berlin](#)

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

[Budapest Statement / Déclaration de Budapest](#) [Canada](#) [Canadiana.org](#)

[Canadian government/le gouvernement du Canada](#) [CAPOS](#) [CARL / ABRC](#)

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

[CRKN / RCDR](#) [data management / gestion des données](#)

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

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

[funding agencies / organismes de financement](#) [genAI](#)

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

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

[international policy / politique internationale](#) [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 scien](#)

[open science / science ouverte](#) [open source software / les logiciels libres](#) [ORCID](#)

[peer review / critique des pairs](#) [PKP](#) [Plan S](#) [Plan S update / mise à jour du Plan S](#)

[policy / politique](#) [promotion et titularisation](#) [publishing / édition](#) [RDC / DRC](#)

[RDM](#) [RECODE](#) [recommendations / recommandations](#) [reports / les rapports](#)

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

[research evaluation / 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 / 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](https://creativecommons.org/licenses/by-sa/4.0/).

