

Diplôme de conservateur de bibliothèque

**Documentation électronique et IA
généralive : recompositions d'un
écosystème et adaptations
professionnelles**

Solène Méhat

Sous la direction de Florence Barré
Responsable des Ressources électroniques – SCD Nîmes Université

Remerciements

Je remercie avant toute chose ma tutrice Florence Barré pour sa disponibilité, ses conseils avisés et son accompagnement tout au long de ce travail.

Je remercie, par ailleurs, tous les collègues et professionnels ayant accepté d'échanger avec moi lors de mon travail d'enquête. Leurs perspectives et conseils ont été précieux.

Je remercie chaleureusement mes camarades de la promotion DCB 34 Madeleine Riffaud pour leur présence et leur joie malgré les moments d'adversité.

Je remercie mes proches. Jaime, Pascale, Alba et Iris pour leur accueil et leur générosité ; ma mère pour ses efforts pour naviguer les moments difficiles ; mes amis pour leur soutien indéfectible, proche ou lointain. Une pensée particulière pour Victor et Wazou pour les cuillères.

Résumé :

Depuis 2022, les technologies d'intelligence artificielle générative recomposent en profondeur l'écosystème de la publication scientifique. Entraînées sur des corpus massifs dont les publications en accès ouvert constituent une ressource privilégiée, ces technologies peuvent être utilisées tant pour de la rédaction que pour l'évaluation par les pairs, la découverte bibliographique ou la lecture. Les possibilités d'usage dans l'ensemble de la chaîne de production et de diffusion de l'information scientifique et technique bouleversent l'écosystème de la publication scientifique et les missions des professionnels des bibliothèques. Ce mémoire analyse les tensions entre les potentiels technologiques de l'IA générative et les fragilités qu'elle engendre, en particulier la remise en cause des équilibres sur lesquels repose la science ouverte. Il examine les solutions technologiques développées par des acteurs privés et publics, le cadre normatif en cours de constitution, et les adaptations de la profession.

Descripteurs :

intelligence artificielle générative

documentation électronique

recherche documentaire

enseignement supérieur et recherche

publication scientifique

Abstract :

Since 2022, generative artificial intelligence technologies have been profoundly reshaping the scientific publishing ecosystem. Trained on massive data, with open access publications being a privileged resource, these technologies can be used for writing, peer review, bibliographic discovery, and reading. The possibilities for use throughout the entire chain of production and dissemination of scientific and technical information are disrupting the scientific publishing ecosystem and the missions of library professionals. This thesis analyses the tensions between the technological potential of generative AI and the vulnerabilities it creates, in particular the challenge to the balance on which open science is based. It examines the technological solutions

developed by private and public actors, the regulatory framework currently being established, and the adaptations being made by the profession.

Keywords :

generative artificial intelligence

electronic documentation

document research

academia

scientific publication

Droits d’auteurs

CC BY NC ND

Sommaire

SIGLES ET ABRÉVIATIONS.....	9
INTRODUCTION.....	11
L'ÉCOSYSTÈME DE LA PUBLICATION SCIENTIFIQUE BOULEVERSE PAR LES IA GÉNÉRATIVES.....	15
Contexte général.....	15
<i>Vision d'ensemble.....</i>	<i>15</i>
<i>Le mouvement pour la science ouverte et l'accès ouvert.....</i>	<i>17</i>
<i>Inflation des publications scientifiques.....</i>	<i>19</i>
<i>Remise en question du modèle de la revue et de l'article.....</i>	<i>21</i>
Un écosystème changé par l'IA générative.....	22
<i>Un potentiel d'utilisation dans la totalité des étapes de la publication.....</i>	<i>22</i>
<i>La pertinence de ces usages en question.....</i>	<i>23</i>
<i>Des questionnements de natures multiples.....</i>	<i>25</i>
<i>Du refus à l'encadrement.....</i>	<i>25</i>
<i>De nouvelles opportunités commerciales.....</i>	<i>26</i>
Un écosystème permettant aux IA génératives d'exister, mais fragilisé par les IA génératives.....	28
<i>Les données en libre accès comme source d'entraînement pour les IA génératives.....</i>	<i>28</i>
<i>Une remise en cause du contrat social autour de l'open access.....</i>	<i>30</i>
<i>Tentation de la fermeture et ses limites.....</i>	<i>32</i>
<i>Adaptations.....</i>	<i>33</i>
LES ÉVOLUTIONS TECHNOLOGIQUES DANS LA DOCUMENTATION ÉLECTRONIQUE.....	37
Aux fondements technologiques de l'IA générative.....	37
<i>Des outils statistiques ayant des limites structurelles.....</i>	<i>37</i>
<i>La technologie RAG : réduire les biais par la sélection documentaire.....</i>	<i>39</i>
Un continuum de solutions privées.....	40
<i>Outils d'aides à la rédaction et à la soumission.....</i>	<i>40</i>
<i>Outils de découvrabilité.....</i>	<i>41</i>
<i>Des outils holistiques : de la recherche à l'écriture.....</i>	<i>42</i>
<i>Enjeux critiques des solutions privées.....</i>	<i>43</i>
Des initiatives publiques.....	44
<i>HAL : améliorer le traitement des données.....</i>	<i>44</i>
<i>L'ABES.....</i>	<i>44</i>
<i>La BnF.....</i>	<i>45</i>
<i>ILaaS : une infrastructure mutualisée pour l'enseignement supérieur et la recherche.....</i>	<i>45</i>
Quelles expertises des bibliothécaires pour comprendre ces technologies ?	46
<i>Évaluer l'évolution des pratiques de consultation d'information.....</i>	<i>46</i>
<i>Évaluer méthodiquement la qualité des résultats fournis par l'IA générative</i>	<i>47</i>
<i>Être partie prenante des solutions technologiques.....</i>	<i>49</i>
ENCADREMENTS LÉGAUX, NORMATIFS ET PRATIQUES.....	52
Un cadre normatif en cours de constitution.....	52
<i>La loi.....</i>	<i>52</i>
<i>Préconisations d'acteurs variés sans coordination complète.....</i>	<i>54</i>

<i>Cas concrets : les chartes universitaires.....</i>	<i>56</i>
<i>La place des bibliothèques dans ces réflexions.....</i>	<i>57</i>
Questions de fond : la responsabilité scientifique dans l’usage des IA	
généralité.....	59
<i>Tendances concernant le signalement.....</i>	<i>59</i>
<i>Questions latentes.....</i>	<i>61</i>
QUELLES ÉVOLUTIONS POUR LES PROFESSIONNELS DES	
BIBLIOTHÈQUES ?.....	63
Des compétences pertinentes, une légitimité à construire.....	63
<i>Regard critique et capacité de médiation.....</i>	<i>64</i>
<i>Une reconnaissance et une identité professionnelle fragiles sur ces questions</i>	<i>65</i>
Une profession qui s’adapte pour rester pertinente.....	66
<i>Une variété de positionnements face à la technologie.....</i>	<i>66</i>
<i>Comment se former ?.....</i>	<i>67</i>
Les missions des professionnels des bibliothèques face à l’IA générative....	70
<i>Interactions avec les éditeurs.....</i>	<i>70</i>
<i>Déterminer le périmètre de la politique documentaire.....</i>	<i>71</i>
<i>Accompagner les chercheurs.....</i>	<i>72</i>
<i>Défendre l’accès ouvert.....</i>	<i>73</i>
CONCLUSION.....	75
DÉCLARATION D’USAGE.....	77
BIBLIOGRAPHIE.....	79
IA générative, généralités.....	79
Services d’IA génératives.....	81
Autour des questions d’open access.....	82
Ecosystème de la publication scientifique.....	84
Missions des bibliothèques.....	89
Chartes et guides.....	93
Lois et règlements.....	94
ANNEXES.....	95
TABLE DES MATIÈRES.....	103

Sigles et abréviations

ADBU : Association des directeurs de bibliothèques universitaires
API : application programming interface
APC: Article Processing Charges
COAR : Confederation of open access repositories
COPE : Committee on Publication Ethics
DDIS : Délégation à la déontologie et à l'intégrité scientifique.
DemoES : Démonstrateur numérique dans l'enseignement supérieur
DGESIP : Direction générale de l'enseignement supérieur et de l'insertion professionnelle
DGRI : Direction générale de la recherche et de l'innovation
ESR : Enseignement supérieur et recherche
IA : intelligence artificielle
IA générative : intelligence artificielle générative
IST : Information scientifique et technique
MESR : Ministère de l'Enseignement supérieur et de la Recherche
Ofis : Office français de l'intégrité scientifique
PNSO : Plan national pour la science ouverte
RAG : Retrieval Augmented Generation
RIA : Règlement sur l'intelligence artificielle
SCD : Service commun de la documentation
SPARC : Scholarly Publishing and Academic Resources Coalition
TEI-XML : Text encoding initiative – Extensive markup language
TDM : Text Data Mining
VP : vice-président

INTRODUCTION

Les pratiques de publication scientifique ont connu des mutations importantes ces dernières années, avec un passage massif du papier au numérique pour la diffusion des résultats de la recherche. Cela a entraîné une recomposition du champ éditorial dans son ensemble et les bibliothèques universitaires (BU) ont dû répondre à de nouvelles problématiques dans les politiques documentaires, en intégrant les formats numériques dans les ressources documentaires. La part de la documentation électronique dans les dépenses documentaires des BU françaises a notamment largement augmenté, pour devenir, souvent le plus gros poste de dépense¹. En 2024, d'après l'ESGBU, les ressources numériques correspondaient à 70 % des dépenses d'acquisition².

Le gain d'importance de la documentation électronique s'est accompagné d'un mouvement vers l'ouverture de l'information scientifique et technique (IST). En témoigne notamment le premier Plan National pour la Science ouverte, lancé en 2018, puis le second, en 2022, cherchant à assurer la possibilité, pour les lecteurs, d'accéder librement et gratuitement aux données et résultats de la recherche³. Ces initiatives pour la science ouverte comprennent un engagement pour l'*open access*, c'est-à-dire le libre accès aux publications scientifiques. Il s'agit d'une littérature :

mise à disposition sur l'internet public, permettant à tout usager de lire, télécharger, copier, diffuser et imprimer ces articles, de lancer une recherche dans ces articles, de créer un lien vers leur texte intégral, de les compiler pour les indexer, de les convertir en données pour traitement logiciel, et de les utiliser à toute autre fin légale, sans barrières financières, juridiques ou techniques autres que celles de l'accès à l'internet lui-même.⁴

Les professionnels des bibliothèques sont particulièrement engagés dans la défense de l'accès ouvert. Tout d'abord, ils jouent un rôle de médiation et de sensibilisation à l'accès ouvert auprès des chercheurs. Par ailleurs, ils peuvent participer à financer les publications en libre accès, en particulier lorsque celles-ci le sont par la voie diamant⁵.

Dans cet écosystème, de nouveaux acteurs ont fait irruption à partir de 2022 : les fournisseurs d'outils d'intelligence artificielle générative (ou IA génératives), mis à disposition du grand public. La date de 2022 correspond au développement de GPT-4 et du lancement du modèle d'agent conversationnel ChatGPT par OpenAI. Ces technologies promettent des possibilités nouvelles, assistant les tâches humaines et permettant des gains de productivité. Ces changements correspondent à des innovations technologiques, les grands modèles de langage (en anglais Large Language Models, LLM), des architectures neuronales avancées, entraînées sur des corpus de données massifs, qui parviennent à traiter et à générer

¹ D'après l'enquête ADBU de 2014 la documentation électronique devient le principal poste de dépense des dépenses documentaires à partir de 2009. ADBU, « Evolution des budgets d'acquisition des bibliothèques universitaires », 2014.

² 2024 : les structures documentaires, piliers de la communauté universitaire et de recherche, ESGBU, 2025

³ « Le Plan national pour la science ouverte 2021-2024 : vers une généralisation de la science ouverte en France », sur *enseignementsup-recherche.gouv.fr*, s. d. (en ligne : <https://www.enseignementsup-recherche.gouv.fr/fr/le-plan-national-pour-la-science-ouverte-2021-2024-vers-une-generalisation-de-la-science-ouverte-en-48525> ; consulté le 5 janvier 2026)

⁴ P. SUBER, « 1. Une définition », dans *Qu'est-ce que l'accès ouvert ?*, Marseille, OpenEdition Press, 2016, p. 23-44

⁵ C. PERRET, « LibGuides: Science ouverte et données de la recherche en bibliothèque: Comprendre la publication en open access », s. d. (en ligne : <https://enssib.libguides.com/c.php?g=705396&p=5083597> ; consulté le 4 janvier 2026)

du langage naturel avec un niveau de sophistication proche de celui des humains⁶. D'après le baromètre du numérique, près de la moitié de la population française utilise l'IA générative, et un tiers des utilisateurs en a un usage quotidien⁷. Les technologies d'intelligence artificielles existent néanmoins avant 2022 : *stricto sensu*, l'IA existe et est utilisée depuis la fin des années 1950⁸. La différence, en 2022 est le caractère « génératif » des nouveaux outils : ceux-ci sont capables, à partir de leur entraînement sur des larges jeux de données, de produire des contenus originaux, avec un degré de cohérence élevé et une pertinence contextuelle⁹. Ils sont, par ailleurs, accessible au grand public dans des versions gratuites.

Ces technologies reposent sur des modèles dont l'entraînement requiert des masses de données volumineuses pour générer un apprentissage automatique. La documentation électronique, en particulier celle en accès ouvert, entre dans le périmètre de ces données. Dans la mesure où la qualité des données d'entraînement influe sur la qualité des modèles d'IA générative, la qualité de la documentation électronique participe du bon fonctionnement de ces technologies – dans une proportion difficile à évaluer, notamment, car les données d'entraînement des LLM sont souvent mal connues. L'interaction entre IA générative et documentation électronique ne s'arrête néanmoins pas là : les outils d'IA générative peuvent être utilisés dans le cadre de recherches scientifiques, *a minima* comme aides à la rédaction ou à la correction de texte. Les cas d'usages sont très divers, du fait de la versatilité même des outils, capables de produire des contenus multimodaux : il est possible d'utiliser non seulement l'IA générative pour la production du contenu scientifique, mais également pour évaluer celui-ci, ou pour le diffuser ou la rendre accessible. Pour les moteurs de recherche ou les bases de données bibliographiques, l'IA générative ouvre, en effet, la potentialité de passer d'une recherche par mots-clés à une recherche sémantique, ce qui change les modalités de la recherche documentaire dans un cadre académique.

Les potentiels permis par l'IA générative se répercutent donc sur des points très divers du champ de la documentation électronique. T. Bergstrom et D. Ruediger, dans leur article, « A Third Transformation? Generative AI and Scholarly Publishing » affirment ainsi que cette technologie recompose radicalement l'écosystème de la publication scientifique, en altérant nombre des maillons de cet écosystème¹⁰. Ce mémoire se propose de faire le point sur les éléments saillants de cette recomposition, en s'intéressant notamment aux conséquences de ces évolutions sur le monde des bibliothèques, et à la vision de la profession sur ces questions, en portant une attention particulière à la manière dont ces bouleversements affectent les engagements pour l'accès ouvert. Face à ce bouleversement, il convient de s'interroger sur la manière dont l'écosystème réagit

⁶ M. BERGADAA et P. PEIXOTO, *Réinventer l'intégrité académique à l'ère de l'intelligence artificielle*, Caen, France, Éditions EMS, Management & société, 2025, p. 34

⁷ PÔLE SOCIÉTÉ DU CREDOC, *Baromètre du numérique*, 2026

⁸ J. HENNO, « 1956 : et l'intelligence artificielle devint une science », sur *Les Echos*, 21 août 2017 (en ligne : <https://www.lesechos.fr/2017/08/1956-et-lintelligence-artificielle-devint-une-science-181042> ; consulté le 9 janvier 2026)

⁹ L'IA Act du 13 juin 2024 définit ainsi un « système IA » : un système automatisé qui est conçu pour fonctionner à différents niveaux d'autonomie et peut faire preuve d'une capacité d'adaptation après son déploiement, et qui, pour des objectifs explicites ou implicites, déduit, à partir des entrées qu'il reçoit, la manière de générer des sorties telles que des prédictions, du contenu, des recommandations ou des décisions qui peuvent influencer les environnements physiques ou virtuels. »

¹⁰ T. BERGSTROM et D. RUEDIGER, *A Third Transformation? Generative AI and Scholarly Publishing*, Ithaca S+R, 2024

et s'adapte à une technologie qui, en partie, dépend des données de la science pour sa bonne qualité. En effet, l'intelligence artificielle générative s'appuie sur les données de la science ouverte tout en menaçant potentiellement leur qualité et leur accessibilité. Comment les professionnels des bibliothèques travaillant sur la documentation électronique ont-ils appréhendé les technologies d'IA générative à partir de 2022 ? Quelles nouvelles questions se posent à eux, dans cet écosystème recomposé ? Quelles adaptations dans leurs tâches et leurs compétences métier ont semblé nécessaires, et quelles sont les forces et faiblesses de la profession dans ce contexte nouveau ? En proposant un état des lieux des enjeux et des questionnements qui touchent aux bibliothèques, il s'agit donc d'identifier la manière dont les missions et les pratiques des bibliothécaires s'adaptent dans le champ de la documentation électronique bouleversé par l'IA générative.

Le présent travail s'appuie sur des recherches documentaires, et sur une veille des publications et communications touchant aux technologies d'IA générative, à leur impact sur l'ouverture des données et des publications. Au cours des dernières années, l'impact de l'IA générative a été proposé comme thème de nombreuses journées d'études professionnelles, que ce soit parmi les associations de bibliothécaires, ou chez les éditeurs et les chercheurs. Cela atteste de l'importance des bouleversements entraînés par les technologies d'IA générative. Le foisonnement des publications, et la nouveauté du sujet, ont eu des conséquences sur la méthode du présent travail : il n'était pas possible de viser l'exhaustivité dans le traitement de la documentation pouvant intéresser cette question. La bibliographie a donc été constituée avec un effort de représentativité. Les recherches documentaires ont été complétées par une série d'entretiens semi-directifs menés auprès de différents acteurs dans le champ de la documentation électronique : les professionnels des bibliothèques, bien évidemment, mais aussi des éditeurs et des formateurs. Le but de ces entretiens était d'obtenir une vue d'ensemble des positionnements des acteurs composant le paysage de la documentation électronique. La majorité des entretiens a été menée auprès de bibliothécaires travaillant soit dans le domaine de la documentation électronique, soit dans l'appui à la recherche. Certains des répondants sont, par ailleurs, membres d'organisations nationales (ABES, Couperin). Les entretiens cherchaient à rendre compte à la fois de la manière dont ces questions étaient appréhendées aussi bien dans des services de SCD assurant leur fonctionnement au quotidien que dans des instances adoptant des visions stratégiques notamment sur le temps plus long. La grille d'entretien, présente en annexe, questionnait la vision que les professionnels portent sur les mutations présentes et à venir dans le champ de la documentation électronique à l'aune des bouleversements technologiques en s'intéressant à quatre enjeux importants identifiés autour de cette question : les évolutions des outils de documentation électronique, l'impact de l'utilisation de l'IA générative sur les questions d'intégrité scientifique, les évolutions institutionnelles et, finalement, les évolutions des compétences nécessaires au bon exercice du métier de bibliothécaire. Cette méthodologie permet de proposer une vision d'ensemble du champ de la documentation électronique entre mars 2025 et janvier 2026. Force est de constater que, dans le domaine des technologies de l'IA générative, les projets et innovations se multiplient depuis 2022, et que le paysage technologique évolue très rapidement. Ce travail n'est donc qu'un instantané, qui, s'il tente d'envisager, les pistes prospectives concernant ces enjeux, reste myope quant aux possibilités et évolutions futures liées à ces technologies.

Dans un premier temps, il s'agit de dresser un panorama de l'écosystème de la publication scientifique et des changements entraînés par l'IA générative. Après avoir

rappelé les caractéristiques de cet écosystème et les mutations qu’il connaissait déjà avant 2022, nous analyserons les bouleversements spécifiques induits par l’IA générative sur l’ensemble de la chaîne de publication scientifique. La deuxième partie se concentre sur les potentiels technologiques dans les pratiques de la documentation électronique et de l’accès ouvert, et sur les conséquences des nouveaux outils d’IA générative. La troisième partie examine la présence et l’adaptation des normes et cadres légaux face à ces bouleversements technologiques. La quatrième partie se concentre sur les modifications des compétences et missions des professionnels des bibliothèques à l’aune de ces évolutions.

L'ÉCOSYSTÈME DE LA PUBLICATION SCIENTIFIQUE BOULEVERSÉ PAR LES IA GÉNÉRATIVES

L'irruption des technologies d'IA générative dans le paysage de la recherche scientifique doit se comprendre à la lumière des transformations profondes qui traversent, depuis plusieurs décennies, l'écosystème de la publication scientifique. Les mutations du modèle économique des revues, l'essor du mouvement pour la science ouverte et la reconfiguration des rapports de force entre chercheurs, institutions et éditeurs constituent autant de dynamiques qu'il convient d'appréhender pour saisir la portée des bouleversements en cours.

Cette partie propose donc, dans un premier temps, un panorama de l'écosystème de la publication scientifique et des grandes tendances qui le structurent aujourd'hui. Elle s'attachera ensuite à mettre en évidence les changements spécifiques qu'introduisent les systèmes d'IA générative dans cet environnement. Enfin, un troisième temps permettra d'explicitier les dynamiques d'interdépendance et les fragilités qui caractérisent désormais les relations entre ces technologies et l'écosystème de la communication scientifique.

CONTEXTE GÉNÉRAL

Les publications scientifiques sont un élément important du fonctionnement de la science. Elles assurent la circulation des idées, le partage des protocoles et des méthodes de recherche, et leur lecture critique représente une part importante du travail des chercheurs. La qualité de l'information scientifique et technique (IST) repose sur des usages bien établis : la présence de comités éditoriaux composés de chercheurs et l'exigence d'une validation par les pairs pour toute publication dans une revue scientifique garantissent le respect de la démarche scientifique et la qualité des résultats publiés. Il existe ainsi un écosystème complexe autour de la publication scientifique, dans lequel interviennent de nombreux acteurs pour assurer la rédaction, l'évaluation, la publication, la diffusion et l'accès aux publications scientifiques.

Vision d'ensemble

L'écosystème de la publication scientifique repose sur l'interaction de plusieurs catégories d'acteurs aux logiques et aux besoins différents. Les chercheurs occupent une position centrale en tant que producteurs du contenu autour duquel s'organise cet écosystème informationnel, puisqu'ils sont les rédacteurs et auteurs des publications scientifiques, que ce soit dans des articles de revues, des chapitres d'ouvrages collectifs ou des monographies. Ils sont, également, les garants de la qualité scientifique des publications lorsqu'ils interviennent en tant qu'évaluateurs lors du processus d'évaluation par les pairs. Cette double fonction les place au cœur du système de validation des connaissances dans la mesure où ils sont partie prenante à la fois de la production et de la diffusion des savoirs. Par ailleurs, les chercheurs sont également situés du côté de la réception : ils sont les destinataires principaux des publications scientifiques – même si celles-ci sont également lues par les étudiants, et peuvent

intéresser d'autres acteurs de la société. Les besoins, pour les lecteurs, d'accès à une documentation scientifique de qualité, à jour et la plus exhaustive possible est l'un des enjeux fort de la publication scientifique.

Les éditeurs, qu'ils soient commerciaux ou à but non lucratif, assurent la diffusion et la mise en forme des publications. Ils gèrent les processus éditoriaux, organisent l'évaluation par les pairs, et fournissent les infrastructures techniques nécessaires à la publication et à l'indexation des publications. Leur modèle économique, particulièrement celui des grands éditeurs commerciaux, repose sur la vente d'accès aux contenus produits par les chercheurs. Ces derniers ne sont ni rémunérés pour l'écriture, ni pour les évaluations des articles.

Les bibliothèques et institutions publiques jouent un rôle d'intermédiaire essentiel entre les éditeurs et les lecteurs finaux. Elles assurent notamment l'acquisition et la mise à disposition des publications. En France, le consortium Couperin, créé en 1999, mutualise les négociations avec les éditeurs pour optimiser les conditions d'accès aux ressources électroniques, principalement pour les bibliothèques universitaires. Les bibliothèques exercent également des missions de formation et d'accompagnement des chercheurs dans leurs pratiques documentaires. Elles les aident à se repérer dans cet écosystème et peuvent les conseiller sur les enjeux entourant leurs publications.

Le schéma ci-dessous, extrait de l'article « The political economy of academic publishing », publié par Puehringer et al. présente une vue simplifiée des différentes interactions entre les acteurs évoqués précédemment, en mettant de côté l'existence de lecteurs non-chercheurs¹¹.

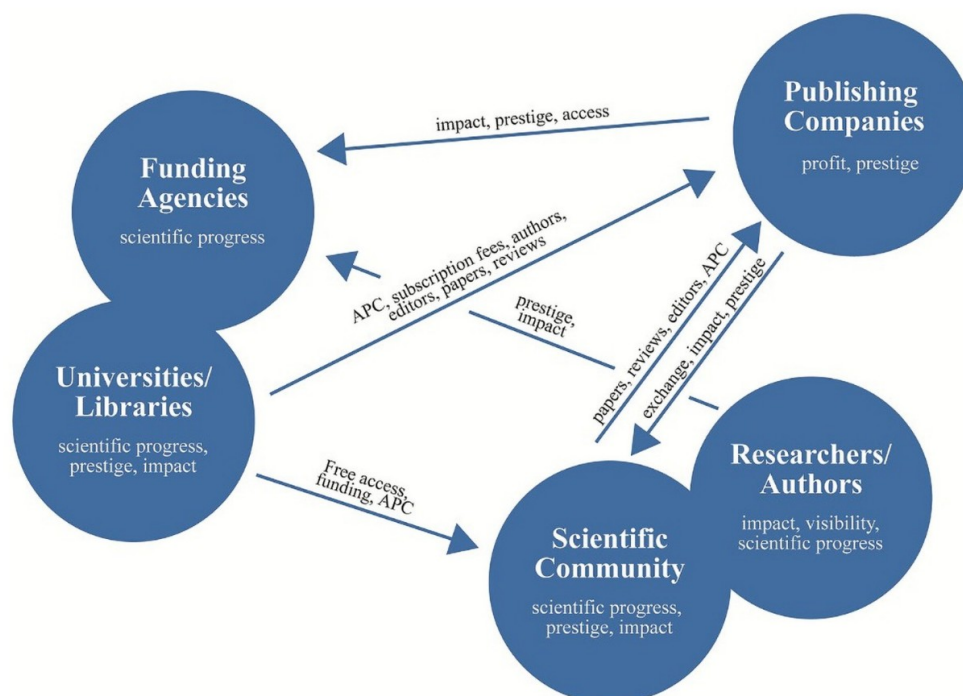


Fig 1. Stylized constellation of main actors in the field of academic publishing.

Ce schéma illustre la complexité des interactions entre les acteurs et met bien en lumière qu'il s'agit d'un écosystème dont le fonctionnement repose sur un

¹¹ S. Puehringer, J. Rath et T. Griesebner, « The political economy of academic publishing : On the commodification of a public good », PLOS ONE, vol. 16, no 6, Public Library of Science, 17 juin 2021 (DOI : 10.1371/journal.pone.0253226 consulté le 26 janvier 2026), p.5

équilibre holistique plus que sur une circulation linéaire de l'information. Cette représentation nécessite néanmoins d'être complétée par la mention d'archives numériques ouvertes. Elles permettent aux chercheurs de déposer et de diffuser leurs publications de manière autonome, en contournant partiellement les circuits éditoriaux traditionnels. Citons l'archive numérique à vocation pérenne HAL, en France, ou ArXiv, beaucoup utilisé dans certaines disciplines comme les mathématiques, la physique et l'informatique pour le dépôt de *preprint*. Ces plateformes garantissent un accès libre et gratuit aux publications, sans barrière financière ni pour les auteurs ni pour les lecteurs. Cependant, les publications qui s'y trouvent ne sont pas soumises au processus d'évaluation par les pairs, ce qui pose des questions quant à la fiabilité des résultats de la recherche qui y sont présentés.

Ce système a connu une croissance importante ces dernières années, dont témoigne notamment l'augmentation du nombre de chercheurs. D'après le Rapport sur la science de 2021 publié par l'UNESCO, le nombre de chercheurs a augmenté de 13,7 % entre 2014 et 2018, trois fois plus que la population mondiale. Néanmoins, la répartition est inégale en fonction des régions du monde. Ainsi, en 2018, 14 % de la population mondiale vivait en Afrique subsaharienne, mais seulement 0,7 % des chercheurs y étaient localisés. Cette augmentation du nombre de chercheurs a été alimentée par la hausse de 19,2 % des dépenses publiques des États entre 2014 et 2018¹².

Du côté des éditeurs, le paysage présente des caractéristiques particulières. S'il existe de nombreux journaux en accès ouvert – le Directory of Open Access Journals hébergeant en janvier 2026 plus de 22 000 journaux – le marché est fortement concentré autour de grands acteurs commerciaux : Springer Nature, Elsevier, Taylor & Francis, Wiley et Sage. Entre 2018 et 2024, les cinq plus grands éditeurs ont vu leur part de marché croître de 9 %, passant de 52 % à 61 % du marché total¹³, renforçant une dynamique déjà établie précédemment. Cette concentration crée des risques de comportements monopolistiques de la part des principaux éditeurs¹⁴, dans un contexte où les coûts des services proposés par ces éditeurs sont déjà très élevés. En 2021, on estimait à 475 milliards d'euros les dépenses annuelles consacrées aux abonnements à des revues scientifiques auprès de cinq grands éditeurs, montant pris en charge par les deniers publics, à 72 % par les budgets des universités¹⁵.

Le mouvement pour la science ouverte et l'accès ouvert

Le paysage de la publication scientifique connaît depuis le début des années 2000 un changement de paradigme important avec l'émergence puis la consolidation du mouvement pour la science ouverte. Parmi les impulsions fondatrices de ce mouvement, citons la *Budapest Open Access Initiative*, lancée en 2001 par l'Open Society. Cette déclaration a posé les bases d'une réflexion collective mondiale sur la nécessité de garantir un accès libre et gratuit aux résultats de la recherche scientifique. En janvier 2026, elle comptait 7 681 signataires individuels et plus de 2 000 institutions et

¹² United Nations Educational, Scientific and Cultural Organization, UNESCO Science Report 2021: The Race Against Time for Smarter Development, United Nations, 2021, p.34-35

¹³ D. CROTTY, « 2025 Update: Quantifying Consolidation in the Scholarly Journals Market », sur *The Scholarly Kitchen*, 20 août 2025 (en ligne : <https://scholarlykitchen.sspnet.org/2025/08/20/2025-update-quantifying-consolidation-in-the-scholarly-journals-market/> ; consulté le 26 janvier 2026)

¹⁴ S. PUEHRINGER, J. RATH et T. GRIESEBNER, « The political economy of academic publishing », *op. cit.* p.2

¹⁵ UNITED NATIONS EDUCATIONAL, SCIENTIFIC AND CULTURAL ORGANIZATION, UNESCO Science Report 2021, *op. cit.*, p.13

organisations¹⁶. La possibilité de consulter une publication scientifique sans barrière technique ou financière est comprise dans les engagements pour la science ouverte.

Pour la communauté scientifique, faciliter et élargir la diffusion des résultats de recherche répond à un objectif d'efficacité. Cela permet d'accélérer la circulation des idées et d'améliorer le fonctionnement global de la recherche. La suppression des barrières financières à l'accès favorise la collaboration entre chercheurs et institutions, indépendamment de leurs moyens économiques. Ce modèle permet aux chercheurs du monde entier, en particulier dans les pays en développement, d'accéder aux dernières découvertes sans coûts prohibitifs, facilitant ainsi une collaboration et un partage des connaissances accru au-delà des frontières¹⁷. Pour la société dans son ensemble, l'accès aux résultats d'une recherche largement financée par des fonds publics répond à un principe d'équité et contribue à maintenir la confiance du public dans la science. La science ouverte vise ainsi à conjuguer une plus grande efficacité de la recherche avec une démocratisation de l'accès aux connaissances et aux savoirs.

En France, cette dynamique s'est traduite par l'adoption d'un premier plan national pour la science ouverte (PNSO) entre 2018 et 2021, puis d'un second entre 2021 et 2024. Cette stratégie comprend un versant sur la publication scientifique, visant une généralisation de l'accès ouvert. Pour les articles de revues dont au moins un auteur est affilié à une institution française, l'objectif est un taux d'accès ouvert de 100 % en 2030. Le Baromètre pour la science ouverte mesure les progrès dans ce domaine. En 2024, il estimait que le taux d'accès ouvert des publications scientifiques françaises dotées d'un DOI Crossref et parues durant l'année 2023 s'élevait à 65 %, soit une augmentation de 26,9 points depuis 2018¹⁸.

Cela interroge le rapport des établissements de l'ESR à l'édition scientifique et remet en question les modèles économiques qui structuraient jusqu'alors la circulation des savoirs. On distingue ainsi plusieurs modèles de financements des publications en accès ouvert. La voie dorée repose sur le paiement de frais de publication par les auteurs, nommés *Article Processing Charges*, (APC). Ces frais ont triplé entre 2013 et 2020, pour atteindre 30 millions d'euros à cette date en France, concernant alors 25 % des articles publiés dans le pays¹⁹. Le modèle supprime la barrière à l'accès pour les lecteurs, mais il en crée une nouvelle pour les auteurs. Cela reproduit les inégalités entre les institutions et les chercheurs en fonction de leurs moyens, les barrières ne se situant plus sur les possibilités de lecture mais sur les possibilités d'accès aux publications. La voie diamant propose une alternative où ni les lecteurs ni les auteurs ne supportent les coûts de publication, le financement provenant d'institutions, de consortiums ou de subventions publiques. L'approche élimine les barrières financières tant pour la lecture que pour la publication, et favorise ainsi un accès équitable. Sa viabilité dépend toutefois de sources de financement régulières et d'un soutien institutionnel stable. La voie verte, quant à elle, repose sur le dépôt par les auteurs

¹⁶ « Budapest Open Access Initiative », s. d. (en ligne : <https://www.budapestopenaccessinitiative.org/read/> ; consulté le 10 juin 2025)

¹⁷ J. YOON, H. KU et E. CHUNG, « The road to sustainability: Examining key drivers in open access diamond journal publishing », *Learned Publishing*, vol. 37, n° 3, 2024, p. 1611 ; L. S. LO, « Generative AI and Open Access Publishing: A New Economic Paradigm », *Library Trends*, vol. 73, no 3, février 2025, p. 164

¹⁸ « Baromètre français de la Science Ouverte », (en ligne : <https://barometredelascienceouverte.esr.gouv.fr/> ; consulté le 26 janvier 2026)

¹⁹ A. BLANCHARD, D. THIERRY et M. van der GRAAF, *Retrospective and prospective study of the evolution of APC costs and electronic subscriptions for French institutions*, Comité pour la science ouverte, 2022

de leurs publications dans des répertoires et des archives numériques en libre accès comme HAL en France. Cette solution limite les coûts pour les auteurs et les institutions tout en assurant une large diffusion. Elle pose néanmoins des questions sur la qualité et la complétude des versions déposées, souvent des *preprints* qui peuvent différer substantiellement de la version finale validée par les pairs.

En France, le consortium Couperin joue un rôle de mutualisation et de négociation collective. Dans le contexte de la transition vers l'accès ouvert, il négocie notamment des accords transformants qui visent à réorienter les dépenses d'abonnement vers le soutien à la publication en accès ouvert²⁰. Comme l'explique Lo, ces accords « représentent une stratégie transitoire, facilitant le passage de modèles basés sur l'abonnement vers des modèles d'accès ouvert²¹ ». Ces négociations incarnent la volonté des bibliothèques de peser sur l'évolution du système de publication scientifique, en utilisant leur pouvoir d'achat collectif pour promouvoir des modèles plus équitables et durables.

L'évaluation précise des coûts de la transition vers l'accès ouvert à l'échelle européenne demeure difficile, comme le souligne l'étude de la Commission européenne sur la publication scientifique en Europe²². Cette complexité s'explique en partie par la diversité des modèles adoptés et par les variations importantes selon les disciplines. En effet, les pratiques d'accès ouvert varient d'un champ disciplinaire à l'autre. Alors que les APC élevés sont courants en biomédecine et dans les autres sciences dites exactes, ils sont absents dans les sciences humaines et les arts. Les sciences sociales se situent entre ces deux pôles²³. Cette hétérogénéité reflète des cultures disciplinaires distinctes, des modèles économiques différenciés et des rapports variables aux infrastructures éditoriales commerciales.

Inflation des publications scientifiques

Le système de publication scientifique fait face à une tension croissante entre volume et qualité. D'après Ophélie Fraiser-Vannier, entre 2015 et 2025, le nombre d'articles publiés dans des revues et actes de conférence a augmenté de 75 %, passant de 10 000 articles par jour à 17 500 par jour²⁴. Cette croissance, qui semble exponentielle, s'explique notamment par l'existence d'incitations à la publication à la fois pour les chercheurs et les éditeurs commerciaux. Les chercheurs ne sont pas rémunérés pour leurs publications ou leur travail d'évaluation par les pairs, mais ces activités sont essentielles à leur reconnaissance professionnelle. Elles assurent leur crédibilité et leur légitimité, et sont l'un des éléments de l'évaluation des carrières pour les institutions qui accordent des financements. Le rôle central de la publication dans l'évaluation et la carrière des chercheurs est synthétisé par le slogan « *publish or perish* » (publier ou disparaître). De leur côté, les éditeurs ont un intérêt financier direct à augmenter le nombre de publications, puisque cela accroît leurs bénéfices, que ce soit par le biais des abonnements ou des APC.

La tendance à l'augmentation des publications pose néanmoins des questions de viabilité. Hanson et al. considèrent ainsi que l'écosystème subi une « pression

²⁰ F. ROUSSEAU-HANS *et al.*, « La négociation d'accords transformants », février 2023 (en ligne : <https://hal.science/hal-04009964> ; consulté le 23 janvier 2026)

²¹ L. S. LO, « Generative AI and Open Access Publishing », *op. cit.*, p. 166. Je traduis

²² EUROPEAN COMMISSION. DIRECTORATE GENERAL FOR RESEARCH AND INNOVATION, « Study on scientific publishing in Europe: development, diversity, and transparency of costs », Publications Office, 2024

²³ S. PUEHRINGER, J. RATH et T. GRIESEBNER, « The political economy of academic publishing », *op. cit.*, p.3

²⁴ O. FRAISIER-VANNIER, « Mésusages des LLM dans la publication scientifique », 30 juillet 2024

insoutenable » à l'heure actuelle : le nombre d'articles évalués par les pairs augmente plus rapidement que le nombre de chercheurs capables de mener ces évaluations²⁵. Cette asymétrie met en péril l'un des piliers fondamentaux de la qualité scientifique puisque les besoins d'évaluation dépassent les forces à même d'y répondre. L'arbitrage entre volumétrie et qualité des publications est d'autant plus problématique que la qualité reste difficile à définir de manière objective et que les mécanismes de contrôle montrent des signes de saturation.

La dynamique est, par ailleurs, intensifiée par des pratiques prédatrices qui participent de l'augmentation du nombre de publications ou de citations tout en mettant en danger la bonne marche de la recherche scientifique. A été ainsi observé le développement de revues prédatrices, qui miment les attributs de légitimité des revues scientifiques mais privilégient l'intérêt et le gain financier au détriment de la rigueur scientifique. Elles publient des articles sans que ceux-ci soient rigoureusement évalués, mais font néanmoins payer des APC. En 2021, elles étaient estimées à 14 000²⁶. Une autre pratique de prédation se développe, non plus du côté de la diffusion de la recherche, mais du côté des résultats de la recherche : la multiplication de faux articles, issus de *paper mills*, publiés dans des revues scientifiquement intègres. En 2023, plus de 10 000 articles ont été retirés de la publication car ils ont été jugés inexacts ou frauduleux *a posteriori*. Les articles avaient passé les processus de vérification préalables à la publication, en particulier l'évaluation par les pairs, mais il s'est, par la suite, avéré qu'ils ne répondaient pas aux exigences de l'intégrité scientifique. Ce chiffre est considéré par les experts comme la « partie émergée de l'iceberg », ce qui témoigne de la large diffusion des pratiques de publication prédatrices²⁷. Les travaux sur ce type de publications mettent non seulement en évidence l'étendue du phénomène, mais également la fragilité des mécanismes de vérification de la qualité et de l'intégrité scientifique. En effet, les articles frauduleux sont parfois difficiles à détecter et ne sont pas toujours rétractés même lorsqu'ils sont identifiés²⁸. Par ailleurs, de nouvelles formes de prédation émergent, encore plus inquiétantes. Les *review mills* mobilisent de « faux reviewers » qui réalisent des évaluations de piètre qualité et indiquent dans leur rapport que l'auteur devrait citer des articles bien précis dans le but d'augmenter leur nombre de citations. Les *citation mills*, quant à eux, manipulent et vendent des citations pour augmenter l'impact apparent de certains chercheurs²⁹. Les éléments d'évaluation de la recherche sont détournés pour grossir des indicateurs de réussite sans qu'ils ne correspondent à des travaux de qualité.

Ces pratiques ont des conséquences délétères pour l'écosystème de la publication scientifique. Elles augmentent d'abord la charge de travail des relecteurs et des éditeurs, qui doivent consacrer du temps à identifier et traiter des soumissions frauduleuses plutôt qu'à évaluer des recherches authentiques. Cette charge supplémentaire renforce le déséquilibre préexistant entre les besoins et les capacités d'évaluation. Plus fondamentalement, elles érodent la confiance dans les

²⁵ M. A. HANSON *et al.*, « The strain on scientific publishing », *Quantitative Science Studies*, vol. 5, n° 4, 1^{er} novembre 2024, p. 823-843

²⁶ C. BOUKACEM-ZEGHMOURI, « La fin de la publication scientifique ? Une analyse entre légitimité, prédation et automatisation », *Histoire de la recherche contemporaine. La revue du Comité pour l'histoire du CNRS*, Tome X-n°2, Comité pour l'histoire du CNRS, 31 décembre 2021 (DOI : 10.4000/hrc.6184 consulté le 17 juillet 2025) p.9

²⁷R. VAN NOORDEN, « More than 10,000 research papers were retracted in 2023 — a new record », *Nature*, vol. 624, n° 7992, 12 décembre 2023, p. 479-481

²⁸ G. CABANAC et C. LABBÉ, « Prevalence of nonsensical algorithmically generated papers in the scientific literature », *Journal of the Association for Information Science and Technology*, vol. 72, n° 12, 2021, p. 1461-1476

²⁹« Les paper mills, review mills et citation mills », sur *Open science : évolutions, enjeux et pratiques*, 7 juin 2024 (en ligne : <https://www.enssib.fr/bibliotheque-numerique/documents/72200-les-paper-mills-review-mills-et-citation-mills.pdf> ; consulté le 6 mai 2025)

publications scientifiques en fragilisant le système en place censé garantir la qualité des publications. L'érosion de la confiance menace l'ensemble de l'édifice de l'IST, qui repose sur la crédibilité des résultats publiés.

Remise en question du modèle de la revue et de l'article

L'écosystème de la publication scientifique connaît une autre évolution qui peut changer les équilibres actuellement en place : la perte d'importance de l'article et de la revue. Dans le modèle exposé précédemment, la revue joue un rôle structurant dans la recherche scientifique. Elle assure simultanément plusieurs fonctions essentielles : la diffusion des résultats, leur validation par le processus d'évaluation par les pairs, leur conservation dans la durée. En outre, elle garantit la paternité d'un résultat pour son auteur³⁰. L'espace de la revue constitue également un lieu de conversation scientifique, avec notamment la publication de comptes rendus critiques portant sur les travaux d'autres chercheurs. Ce modèle existe depuis le XVII^e siècle, date à laquelle naissent les premières revues savantes. Il perdure avec la transition vers le numérique. Néanmoins, il connaît progressivement une transformation profonde : les métadonnées, les références bibliographiques, les données expérimentales peuvent être échangées, quantifiées, reliées, si bien que l'article et la revue ne sont plus des unités de savoir fondamentales³¹.

Ces transformations n'affaiblissent cependant pas le modèle dominé par les grands acteurs à but lucratif. Les grands acteurs privés s'adaptent aux évolutions des modes de publication, qui leur permet de maintenir, voire de renforcer, leur position dominante dans l'écosystème. Les éditeurs ont ainsi investi dans de maillons de la chaîne de valeur : Elsevier a acquis Mendeley, un outil de gestion de bases documentaires, ainsi qu'une base de *preprints* en sciences sociales. Nature Springer a racheté Dimensions, qui produit des indicateurs alternatifs et gère également une base de données d'articles. Taylor & Francis a absorbé des éditeurs alternatifs comme F1000 ou PeerJ. La stratégie de diversification et d'intégration verticale renforce le contrôle de ces acteurs sur l'ensemble du processus de production et de diffusion des savoirs scientifiques. Un rapport de la Scholarly Publishing and Academic Resources Coalition (SPARC) de 2019 constatait ainsi une progressive transformation de certaines maisons d'édition privées en « *data analytics company* », développant de nouveaux services commerciaux en proposant des analyses quantitatives des données scientifiques contenues dans les articles qu'elles publient³².

Ce portrait, à grands traits, des équilibres et tendances qui structurent l'écosystème de la publication scientifique à l'ère du numérique, est nécessaire pour mesurer les effets de l'arrivée des technologies d'IA générative. Celles-ci, en effet, créent des changements et ouvrent des potentiels nouveaux, qui interagissent avec les dynamiques déjà complexes préexistantes et sont susceptibles d'en modifier profondément les équilibres.

³⁰I. BIZOS, « Big deals et open access : quelle stratégie numérique pour les bibliothèques universitaires ? », 2020, p.22

³¹C. BOUKACEM-ZEGHMOURI, « La fin de la publication scientifique ? », *op. cit.* p.10.

³²C. ASPESI *et al.*, « SPARC Landscape Analysis: The Changing Academic Publishing Industry – Implications for Academic Institutions », LIS Scholarship Archive, 2019, p.7

UN ÉCOSYSTÈME CHANGÉ PAR L'IA GÉNÉRATIVE

Les technologies d'intelligence artificielle générative s'inscrivent dans ce que Sejnowski qualifie de « révolution du deep learning³³ » : polymorphes, elles ont des potentiels d'application dans un très large champ d'activités humaines. Sans surprise, elles sont susceptibles d'être mobilisées dans les processus de recherche et de publication scientifique. Il importe donc de déterminer dans quelle mesure elles peuvent être convoquées, et si ces nouvelles technologies recomposent, ou non l'écosystème décrit précédemment.

Un potentiel d'utilisation dans la totalité des étapes de la publication

Les IA génératives peuvent être utilisées pour un ensemble de tâches dans le processus de publication scientifique. Ces technologies sont extrêmement polymorphes grâce à leur double capacité à traiter d'informations de manière qui donne une impression de compréhension et à générer, en réaction, des réponses qui semblent elles aussi cohérentes. Précisant les usages possibles pour les chercheurs, Andersen et al. proposent une typologie qui différencie trois types de pratiques :

- « *GenAI as a work horse* » : automatisation de tâches répétitives, par exemple dans le traitement d'un jeu de données
- « *GenAI as a language assistant only* » aide à la rédaction de publications ou d'appels à projets, correction de textes, traduction
- « *GenAI as a research accelerator* »³⁴ : l'IA générative peut par exemple être utilisée pour faire un état de l'art, générer des hypothèses ou élaborer des protocoles de recherche

Cette liste, très générale, pourrait être complétée par des cas d'usages plus spécifiques en fonction des disciplines et des différents processus de recherche. Elle met en évidence la grande variété possible d'usages de ces technologies. D'après Hosseini et Resnik les possibilités d'application de l'IA générative dans la recherche sont « sans limites » et vont altérer profondément le processus de découverte scientifique dans un futur proche. Ces usages n'ont pas les mêmes implications sur le fonctionnement même de la recherche et sur l'écosystème de la publication. On conçoit bien que l'utilisation des IA générative pour la traduction perturbe moins le processus de publication qu'une utilisation pour générer des hypothèses.

L'IA générative peut également être mobilisée pour assister le processus d'évaluation par les pairs, les technologies étant très efficaces pour générer des synthèses³⁵. La puissance de synthèse et de reformulation peut aussi inciter à l'usage de ces outils pour aider à la lecture et à la compréhension des publications scientifiques, en particulier pour des lecteurs non familiers d'un domaine et du vocabulaire technique associé. Finalement, les IA génératives peuvent être utilisées

³³T. J. SEJNOWSKI, *The Deep Learning Revolution*, Cambridge, The MIT Press, 2018

³⁴ J. P. ANDERSEN *et al.*, « Generative Artificial Intelligence (GenAI) in the research process – A survey of researchers' practices and perceptions », *Technology in Society*, vol. 81, 1^{er} juin 2025, p. 102813

³⁵W. LIANG *et al.*, « Monitoring AI-Modified Content at Scale: A Case Study on the Impact of ChatGPT on AI Conference Peer Reviews », *ICML'24: Proceedings of the 41st International Conference on Machine Learning*, 15 juin 2024, p. 29575-29620

pour mener des recherches bibliographiques. Les technologies à usage généraliste ne sont pas très douées sur cette tâche, fournissant souvent des réponses erronées, nommées hallucinations, mais aujourd'hui des outils de recherche bibliographiques « augmentés » du type SciSpace se développent.

On le voit, les potentiels d'usage concernent autant la découverte, la rédaction, que l'évaluation et la lecture de publications. Les IA génératives peuvent donc intervenir dans la totalité de la chaîne de création de savoir scientifique. Elles présentent un potentiel d'optimisation et de gain d'efficacité significatifs, pouvant entraîner une réduction des coûts dans les processus de publication³⁶.

Les usages potentiels correspondent à des pratiques réelles de certains chercheurs. Aujourd'hui, les données quantitatives sur l'utilisation de l'IA générative dans les publications scientifiques restent fragmentaires et il est difficile d'en avoir une vision d'ensemble. L'initiative AcademAI, fondée par Alex Glynn, recense manuellement les usages non déclarés de ChatGPT dans les publications scientifiques. Au 20 janvier 2026, elle comptabilisait 668 articles de revues et 40 chapitres d'ouvrages. Une telle initiative ne permet pas d'avoir une vue d'ensemble de l'utilisation des IA générative, mais elle atteste d'une pratique réelle d'utilisation de l'IA générative dans la recherche, et d'une sensibilité de la communauté scientifique à la question. Des études plus systématiques commencent à émerger, même si elles ne donnent que des aperçus sur des échantillons disciplinaires. Kobak et al., dans une étude sur la publication biomédicale menée en 2024 estiment qu'au moins 13,5 % des résumés de leur échantillon ont été traités avec des LLM³⁷. En 2024, une étude sur les évaluations par les pairs effectuées lors de quelques conférences sur l'IA (ICLR 2024, NeurIPS 2023, CoRL 2023 et EMNLP 2023) estimait qu'entre 6,5 % et 16,9 % des textes soumis comme compte rendus d'évaluation pouvaient avoir été substantiellement modifiés par des LLM, au-delà de simples corrections orthographiques ou mineures³⁸. On le voit, ces usages ne sont pas des épiphénomènes. Dès lors, il importe d'approfondir les questions qu'ils posent dans les processus de publication scientifique.

La pertinence de ces usages en question

La multiplicité des usages potentiels n'implique pas que les usages soient pertinents. Comme le note Aline Bouchard dans son article de veille sur ces technologies : « Ce ne sont pas des outils de connaissances ; ce ne sont pas des outils de recherche d'information ; ce ne sont pas des outils de calcul ; ce ne sont pas des outils de logique³⁹ ». Ces technologies donnent l'impression de pouvoir accomplir des tâches très variées, alors qu'elles ne sont pas toujours efficaces. Les références bibliographiques hallucinées ou les erreurs sur des calculs très simples sont des preuves des limites de ces technologies. Or, cette inadéquation entre perception de compétence et compétence effective pose problème dans un écosystème où la fiabilité est cruciale : les chercheurs peuvent être tentés d'utiliser ces outils pour des tâches auxquelles ils ne sont pas adaptés.

³⁶L. S. LO, « Generative AI and Open Access Publishing », *op. cit.* p.161

³⁷D. KOBAK *et al.*, « Delving into LLM-assisted writing in biomedical publications through excess vocabulary », *Science Advances*, vol. 11, n° 27, 4 juillet 2025 (DOI : 10.1126/sciadv.adt3813 consulté le 20 janvier 2026)

³⁸W. LIANG *et al.*, « Monitoring AI-Modified Content at Scale », *op. cit.*

³⁹A. BOUCHARD, « #WorkInProgress : IA générative et outils de recherche de littérature académique », URFISTinfo, 2025, p.4

Par ailleurs, au-delà des limites techniques, les IA génératives peuvent simplifier et accélérer des pratiques problématiques. Dans un écosystème où la course à la publication était déjà identifiée comme un problème majeur, l'arrivée d'une technologie permettant des gains de temps peut entraîner un emballement dangereux. Paradoxalement, alors que les IA génératives sont souvent perçues comme des technologies novatrices, elles semblent d'abord intensifier les tendances préalables de l'écosystème en les poussant à bout. Ces technologies peuvent ainsi être perçues comme un catalyseur venant exacerber un problème systémique préexistant⁴⁰. Cette intensification des dérives contribue à l'épuisement général du système : surcharge des évaluateurs, inflation des publications, difficultés croissantes à distinguer contenus légitimes et frauduleux. Dans un article du *Monde* de juillet 2025, David Larousserie décrit le monde des revues scientifiques comme « au bord de l'asphyxie⁴¹ ».

Raluca Pierrot, de l'ABES, exprime en entretien une inquiétude marquée confirmant que les acteurs de la publication scientifique ont conscience de cette dynamique : « Il y a deux types de dérives. D'abord, des évaluations qui sont censées être faites par les pairs, mais dont on a l'impression qu'elles sont générées par des IA. Cela signifie que l'évaluation par les pairs, déjà parfois décriée pour son manque de sérieux, devient une évaluation de facilité faite par l'IA. Ensuite, la génération même des articles⁴². » Ces dérives sont confirmées par les travaux sur les pratiques scientifiques frauduleuses comme par les expériences des acteurs de l'écosystème. Nathalie Clément, d'EDP Sciences, mentionne ainsi que les premiers contacts de la maison d'édition avec l'IA générative se sont faits par le biais de la détection d'articles rédigés entièrement de manière artificielle⁴³. Si les IA génératives peuvent accélérer les processus de recherche légitimes, elles facilitent également les pratiques frauduleuses tout en les rendant plus difficiles à détecter⁴⁴.

La possibilité d'utiliser l'IA générative dans la totalité des maillons de la chaîne de documentation électronique peut aboutir à des articles rédigés, évalués par les pairs et lus par des IA génératives, engendrant un risque de perte de sens de la publication scientifique. L'un des problèmes actuels concernant l'intégrité scientifique concerne l'existence d'articles sur lesquels, de manière visible, aucun œil humain ne s'est penché dans le processus de rédaction, de relecture et d'édition. L'un des exemples les plus nets est la présence de prompts ou de textes génériques de ChatGPT dans le corps même du texte⁴⁵. De tels phénomènes révèlent des failles profondes dans la totalité des processus de publication, remettant en cause l'intégrité scientifique dans sa globalité, mais également les fondements même sur lesquels repose cet écosystème.

⁴⁰C. Ro et J. Leeming, « Authorship for sale: Nature investigates how paper mills work », *Nature*, vol. 642, no 8068, Nature Publishing Group, 9 juin 2025, p. 826

⁴¹D. LAROUSSE, « Le monde des revues scientifiques au bord de l'asphyxie », *e*, 7 juillet 2025 (en ligne : https://www.lemonde.fr/sciences/article/2025/07/07/le-monde-des-revues-scientifiques-au-bord-de-l-asphyxie_6619660_1650684.html ; consulté le 17 juillet 2025)

⁴²R. PIERROT, « Entretien », 10 décembre 2025

⁴³S. LA BARBERA et N. CLÉMENT, « Entretien », 11 septembre 2025

⁴⁴D. SINGH CHAWLA, « Fake AI images will cause headaches for journals », *Nature*, 27 mai 2025 (DOI : 10.1038/d41586-025-01488-z consulté le 6 juin 2025) ; C. Ro et J. LEEMING, « Authorship for sale: Nature investigates how paper mills work », *op. cit.*

⁴⁵O. FRAISIER-VANNIER, « Mésusages des LLM dans la publication scientifique », *op. cit.*

Des questionnements de natures multiples

Les usages des IA génératives dans l'écosystème de la publication scientifique soulèvent une multiplicité de questions, de nature juridique, éthique et épistémique. Sur le plan juridique, l'utilisation de l'IA générative dans les processus scientifiques pose des problèmes de protection des données personnelles et de confidentialité. Lorsque des manuscrits ou des évaluations sont versés dans des IA génératives généralistes, par exemple dans le cadre de processus d'évaluation assistés par IA, cela peut violer la confidentialité de la publication et exposer des données sensibles. Les principaux éditeurs scientifiques ont d'ailleurs établi un consensus interdisant aux évaluateurs de téléverser tout ou partie d'un manuscrit dans un outil d'IA générative accessible au public, l'utilisation des IA dans les processus d'évaluation étant strictement interdite⁴⁶.

Se posent également des questions relatives au maintien de l'intégrité scientifique. Elles concernent à la fois la détection de pratiques frauduleuses et les modalités de signalement d'usage de l'IA générative dans un cadre considéré comme acceptable. Avec l'assistance de l'IA, la fabrication et la falsification de données sont devenues plus sophistiquées et difficiles à détecter : les modèles génératifs peuvent produire des ensembles de données, des images scientifiques ou des textes académiquement plausibles mais entièrement fictifs. Face à ces risques, les éditeurs adoptent des outils de détection, tandis que les politiques de signalement varient selon les revues, allant de l'interdiction complète du texte généré par IA à des usages permissifs avec exigences de transparence.

Du refus à l'encadrement

Face à ces questionnements, les acteurs de l'édition scientifique ont réagi selon plusieurs logiques successives, qui dessinent une recomposition progressive du marché.

La première réaction des éditeurs a été celle d'une protection de leurs contenus, adoptant des postures de blocages. À partir de 2024, certains éditeurs ont inclus des clauses contractuelles restrictives interdisant l'usage de l'IA générative dans les licences proposées aux bibliothèques, visant à protéger les contenus mis à disposition⁴⁷. Ces clauses ont soulevé des questions juridiques, notamment au regard de l'exception TDM (*text and data mining*) prévue par la directive européenne sur le droit d'auteur (DANUM), et ont nécessité une mobilisation des bibliothécaires pour en clarifier la portée.

Cette posture de protection d'usage des contenus s'est doublée d'une réflexion sur les pratiques de rédaction et de publication pour les articles. Les éditeurs ont, sur ce versant, adopté une posture d'encadrement en élaborant des politiques d'usage, avec des approches variant de l'interdiction complète à des usages permissifs sous condition de transparence. Le consensus actuel repose sur deux principes : le maintien de la responsabilité humaine dans l'auctorialité et l'interdiction de considérer une IA générative comme auteur⁴⁸.

⁴⁶ « AI Policies in Academic Publishing 2025: Guide & Checklist - Thesify - Ethical AI Tools That Improve Your Academic Writing », s. d. (en ligne : <https://www.thesify.ai/blog/ai-policies-academic-publishing-2025> ; consulté le 28 janvier 2026)

⁴⁷ S. PERRIN, « Entretien », s. d. ; R. PIERROT, « Entretien », *op. cit.*

⁴⁸ « AI Policies in Academic Publishing 2025 », *op. cit.*

Parallèlement à cet encadrement normatif, les éditeurs se sont engagés dans une course technologique. Comme le note Nathalie Clément (EDP Sciences), plutôt que de chercher à empêcher l'utilisation de ces technologies, il s'agit de les mobiliser pour limiter les usages frauduleux et accélérer les processus éditoriaux⁴⁹. Cette stratégie se déploie sur deux fronts. D'une part, le développement d'outils de détection : les plateformes et éditeurs conçoivent des systèmes automatisés pour repérer les images ou données manipulées, dans une logique où l'IA sert à contrer les dérives qu'elle rend possibles. D'autre part, l'intégration d'outils d'assistance au *peer review*, processus central mais fragilisé par l'inflation du nombre de publications. En entretien, un conservateur observe ainsi que les éditeurs proposent désormais aux chercheurs des « outils natifs pour faire le *peer review* où vous êtes accompagné par de l'IA ».

On le voit, les éditeurs scientifiques deviennent des acteurs clés dans la recomposition, car leur position dans l'écosystème leur permet d'organiser les relations contractuelles autour de l'IA générative. Tout d'abord, ils édictent des normes d'utilisation dans le processus de publication et sont donc prescripteurs en ce qui concerne les pratiques d'usage et de signalement par les auteurs. Par ailleurs, ils travaillent à l'intégration de ces technologies dans les infrastructures et solutions technologiques sur lesquelles reposent les publications scientifiques. Finalement, ils peuvent négocier avec les entreprises d'IA générative des accès aux contenus qu'ils gèrent. Cette position de pivot leur confère un pouvoir structurant dans l'évolution de l'écosystème.

De nouvelles opportunités commerciales

Les IA génératives ouvrent de nouvelles opportunités commerciales pour les éditeurs, dans la continuité de la monétisation croissante des données de la recherche déjà évoquées.

De nombreux éditeurs signent désormais des accords avec des entreprises d'IA générative pour fournir des données d'entraînement aux modèles de langage. L'initiative Ithaka S+R « Generative AI Licensing Agreement Tracker » assure une veille sur l'ensemble des accords passés entre éditeurs et entreprises d'IA. On y observe principalement des accords consistant à fournir des données pour entraîner les LLM. La dernière mise à jour de ces données recensait des accords passés par Wiley, Taylor et Francis, John Hopkins University Press, Cambridge University Press, Sage, De Gruyter Brill, en autres. Ces accords ne se limitent pas au monde académique, comme en témoigne l'exemple du Monde avec OpenAI⁵⁰.

Les intérêts monétaires sont évidents dans de tels accords : l'accord entre Taylor & Francis et Microsoft a ainsi été estimé à 10 millions de dollars « pour un accès initial aux données ». Il n'existe pas de pratique unifiée, et beaucoup de secret entoure le périmètre de ces accords ainsi que leur montant financier. À l'heure actuelle, une certaine opacité prévaut dans les clauses et les conditions d'utilisation des données de la recherche entre éditeurs et entreprises d'IA générative. Les implications varient en fonction des modèles économiques des éditeurs. Comme le note Schonfeld, un éditeur fonctionnant sur abonnement peut

⁴⁹ S. LA BARBERA et N. CLÉMENT, « Entretien », *op. cit.*

⁵⁰ L. DREYFUS et J. FENOGLIO, « Intelligence artificielle : un accord de partenariat entre “Le Monde” et OpenAI », *Le Monde*, 13 novembre 2024 (en ligne : https://www.lemonde.fr/le-monde-et-vous/article/2024/03/13/intelligence-artificielle-un-accord-de-partenariat-entre-le-monde-et-openai_6221836_6065879.html ; consulté le 27 janvier 2026)

craindre que l'utilisation de ses données par un LLM réduise la consultation finale des textes par des lecteurs, et donc ses sources de revenus directs. L'inquiétude n'est pas la même pour éditeur publiant en *open access* en voie dorée, qui conserve ses revenus par APC même en cas de baisse de la consultation directe des publications. L'enjeu, dans le cas de la voie dorée, réside donc plus dans le maintien de qualité des données⁵¹.

Certaines approches attestent d'un souci de maintenir l'équilibre et la confiance entre les différents acteurs. Cambridge University Press, bien qu'estimant que les clauses des contrats classiques pouvaient couvrir des accords passés avec des fournisseurs d'IA générative, a fait le choix d'ajouter une clause nouvelle donnant l'opportunité aux auteurs d'accepter ou de refuser que leurs publications soient utilisées dans le cadre d'accords entre l'éditeur et des compagnies privées pour l'entraînement des LLM ou pour des usages en RAG⁵². Autre exemple : Bloomsbury a donné à l'été 2025 l'opportunité aux auteurs de recevoir 20 % de royalties en cas de contrats avec des entreprises d'IA générative⁵³.

Par ailleurs, les éditeurs développent des nouveaux services qui intègrent également des fonctionnalités d'IA directement dans leurs offres d'abonnement. Ils sont de nature variable : outils de recherche augmentée, recommandations personnalisées, assistance à la synthèse, assistance à l'évaluation. Outre les potentiels technologiques, ces développements constituent, pour les éditeurs, un argument de négociation. Isabelle Filliâtre note ainsi que l'IA représente « un levier pour les éditeurs notamment au niveau des négociations [...] pour augmenter les prix⁵⁴ ». Ces évolutions font craindre l'émergence d'un système d'accès à l'information différencié. Lo souligne que les services d'analyse pilotés par l'IA, les recommandations personnalisées et les capacités de recherche avancée pourraient devenir des services premium, créant une recherche à plusieurs vitesses où seules les institutions disposant de moyens importants accèderaient aux fonctionnalités de découvrabilité et de synthèse⁵⁵.

Plus largement, le développement de systèmes d'IA avancés nécessite des investissements considérables, ce qui favorise les éditeurs disposant de ressources importantes. Cela pourrait maintenir la concentration forte du marché, affectant la diversité de l'écosystème éditorial et exacerbant les inégalités existantes dans la production et la dissémination mondiale des connaissances. Ici encore, les changements induits par l'IA renforcent des tendances préexistantes.

Ces évolutions technologiques s'inscrivent dans le prolongement de dynamiques antérieures, notamment la concentration du marché, la monétisation des données, les tensions autour de l'accès. Par rapport à l'effet de ces technologies sur l'écosystème de la publication scientifique, Adeline Régé note : « il y a un effet de masse et ça rend en fait visible des choses qui se produisent mais qui étaient déjà là⁵⁶ ». Cependant, à un niveau supérieur, elles participent désormais d'un bouleversement plus fondamental qui affecte la viabilité même de l'écosystème de la publication scientifique.

⁵¹R. C. SCHONFELD, « Tracking the Licensing of Scholarly Content to LLMs », sur *The Scholarly Kitchen*, 15 octobre 2024 (en ligne : <https://scholarlykitchen.sspnet.org/2024/10/15/licensing-scholarly-content-llms/> ; consulté le 4 septembre 2025)

⁵²D. HANSEN, « AI Licensing: An Interview with Ben Denne of Cambridge University Press », sur *Authors Alliance*, 17 mars 2025 (en ligne : <https://www.authorsalliance.org/2025/03/17/ai-licensing-an-interview-with-ben-denne-of-cambridge-university-press/> ; consulté le 2 octobre 2025)

⁵³« Generative AI Licensing Agreement Tracker », sur *Ithaka S+R*, s. d. (en ligne : <https://sr.ithaka.org/our-work/generative-ai-licensing-agreement-tracker/> ; consulté le 4 septembre 2025)

⁵⁴I. FILLIÂTRE, « Entretien », 23 octobre 2025

⁵⁵L. S. LO, « Generative AI and Open Access Publishing », *op. cit.*, p. 161

⁵⁶A. RÉGÉ, « Entretien », 10 décembre 2025

UN ÉCOSYSTÈME PERMETTANT AUX IA GÉNÉRATIVES D’EXISTER, MAIS FRAGILISÉ PAR LES IA GÉNÉRATIVES

Henri Farrel et al., dans un article publié dans *Science*, insistent sur l’importance de considérer les LLM comme des technologies culturelles et sociales, qui s’ajoutent aux systèmes préalables permettant de traiter de l’information de manière à permettre une coordination à grande échelle⁵⁷. Ces technologies rendent plus évidentes les fragilités préexistantes de l’écosystème de la publication scientifique, mais elles entraînent également un bouleversement de fond dans le traitement des savoirs. Les IA génératives questionnent les processus de création, de validation et de dissémination de l’IST, et ce alors qu’elles reposent, dans leur fonctionnement même, sur l’exploitation de ces systèmes informationnels. Contrairement aux outils de recherche documentaire classiques, l’IA générative ne se contente pas de faciliter l’accès à la connaissance : elle y contribue, remettant potentiellement en cause les paradigmes établis. Selina La Barbera et Nathalie Clément, d’EDP Sciences, parlent d’ « édition scientifique en mutation accélérée⁵⁸ » du fait de la combinaison des évolutions de l’IA générative avec l’*open access*. Il existe, en effet, une synergie forte entre le mouvement d’*open access* et le développement de l’IA générative : ces technologies reposent sur l’accès libre à des données de qualité pour l’entraînement des modèles. Cette interdépendance crée néanmoins un paradoxe : si l’*open access* a rendu possible l’émergence des IA générative, celles-ci fragilisent en retour les fondements mêmes de l’écosystème qui les a fait naître.

Les données en libre accès comme source d’entraînement pour les IA génératives

Les IA génératives ont besoin, pour fonctionner, de vastes jeux de données. Celles-ci sont particulièrement importantes dans l’entraînement des LLM. Les modèles reposent sur des processus d’apprentissage qui nécessitent l’exposition à des volumes considérables de données textuelles : les corpus sont soumis à des analyses statistiques qui permettent de dégager des schémas récurrents, réutilisés, ensuite, pour générer des réponses.

Comme le souligne Paul Keller, ces modèles sont construits en utilisant des données publiques et en libre accès sur Internet. Celles-ci sont de nature diverse : le contenu quotidien des sites web, les informations en libre accès dans des initiatives comme Wikipedia, mais aussi des savoirs hébergés par des institutions publiques, parmi lesquelles les publications scientifiques⁵⁹. La masse critique de données librement accessibles constitue le socle technique sur lequel reposent les modèles de langage. Les ressources en libre accès représentent ainsi, pour reprendre les termes de Molly White, un véritable « approvisionnement de nourriture » pour les entreprises d’IA générative⁶⁰. La performance d’une IA

⁵⁷H. FARRELL *et al.*, « Large AI models are cultural and social technologies », *Science*, American Association for the Advancement of Science, 14 mars 2025 (DOI : 10.1126/science.adt9819 consulté le 19 janvier 2026)

⁵⁸S. LA BARBERA et N. CLÉMENT, « Entretien », *op. cit.*

⁵⁹P. KELLER, « Beyond AI and copyright. Funding a sustainable information ecosystem », juin 2025

⁶⁰M. WHITE, « “Wait, not like that”: Free and open access in the age of generative AI », sur *Citation Needed*, 14 mars 2025 (en ligne : <https://www.citationneeded.news/free-and-open-access-in-the-age-of-generative-ai/> ; consulté le 22 janvier 2026)

généraliste dépend ainsi de la qualité, de la variété et du volume des données textuelles qui ont nourri son apprentissage. L'appel de Yann Le Cun, alors scientifique en chef IA chez Meta, illustre cet enjeu :

Every institution, library, foundation, cultural group, and government around the world that possesses cultural content should make it available for training free and open AI foundation models. Free and open AI systems will constitute the repository of all human knowledge and culture⁶¹.

Toutes les institutions, bibliothèques, fondations, groupes culturels et gouvernements du monde qui possèdent des contenus culturels devraient les mettre à disposition pour l'entraînement de modèles d'IA libres et ouverts. Les systèmes d'IA libres et ouverts constitueront le référentiel de toutes les connaissances et de toute la culture humaines.

Si cette auto-promotion a suscité de vives critiques, elle met néanmoins en lumière un enjeu réel : les IA génératives, pour s'entraîner efficacement, requièrent des contenus de qualité. À ce titre, les informations conservées notamment dans les institutions scientifiques et culturelles, y compris les bibliothèques, sont une ressource très précieuse.

Ainsi, les engagements en faveur de la libre circulation des contenus et l'ouverture des données ont contribué à l'essor des IA génératives : les vastes corpus de textes librement accessibles sur le web constituent le matériau permettant l'entraînement des modèles. Bernard Jacquemin, dans un entretien, observe que la bibliothèque est une ressource précieuse pour les entreprises d'IA générative souhaitant entraîner leurs modèles :

la bibliothèque, avec un grand B, est évidemment une source de ces données. Ce n'était pas différent dans les années 2000, quand Google a commencé à lancer ses programmes de numérisation de livres pour créer Google Books. Et ils se sont déjà retournés vers les bibliothèques en essayant de passer des partenariats ou en essayant de tordre la main au gouvernement de n'importe quel pays pour essayer de contourner, dans une certaine mesure, le droit d'auteur.

Ici, il y a des éditeurs d'intelligence artificielle qui développent exactement le même type de stratégie, en essayant de faire en sorte que tout puisse devenir données alimentant les systèmes d'apprentissage à moindre coût⁶².

Les publications scientifiques en accès ouvert sont, à cet égard, des corpus particulièrement recherchés, car les processus de vérification et de révision par les pairs assurent qu'il s'agit de données de très bonne qualité. HAL est ainsi sollicitée par des équipes de recherche qui souhaitent entraîner leurs modèles de langage sur des corpus scientifiques⁶³. Notons néanmoins que les méthodes utilisées pour obtenir ces grands corpus de données, ou les contenus exacts de ces corpus, ne sont pas rendus publics par les entreprises qui entraînent des LLM. À titre d'exemple, voici les informations sur les données d'entraînement utilisées par OpenAI, l'entreprise qui développe ChatGPT :

⁶¹ P. KELLER, « Beyond AI and copyright. Funding a sustainable information ecosystem », *op. cit.*, p. 2. Je traduis.

⁶² B. JACQUEMIN, « Entretien », 21 octobre 2025

⁶³ N. FARGIER et Y. BARBORINI, « AP2024 | HAL et l'IA », 21 novembre 2024 (DOI : 10.60527/S5HF-QQ55 consulté le 26 mai 2025)

(1) les données accessibles publiquement sur Internet, (2) les données obtenues via des partenaires tiers, et (3) les données fournies ou générées par nos utilisateurs, notre personnel en charge de l'entraînement des modèles et nos chercheurs.⁶⁴

Le champ précis des données utilisées n'est pas précisé. Ainsi, alors même qu'elles en ont besoin pour fonctionner, les entreprises d'IA générative ne pratiquent elles-mêmes pas le principe d'ouverture des données. Les principes d'ouverture, conçus initialement pour démocratiser l'accès aux savoirs, ont été mis au service du développement d'outils commerciaux dont les logiques échappent largement aux intentions initiales du mouvement pour l'accès ouvert.

Une remise en cause du contrat social autour de l'open access

L'une des transformations majeures induites par les IA génératives réside dans l'apparition d'une couche supplémentaire de médiation dans l'accès à l'information, contrôlée par de nouveaux acteurs. Or, les entreprises qui développent des IA génératives sont, largement, des entreprises privées, avec des objectifs commerciaux. Comme le souligne Keller, ces acteurs ne sont pas des gardiens impartiaux de l'information et leurs motivations financières peuvent avoir des conséquences dans le traitement de l'information⁶⁵. Les motivations des entreprises privées peuvent, en effet, contrevenir aux exigences de neutralité et de pluralité de la diffusion de l'information.

Par ailleurs, l'utilisation des ressources en libre accès par des entreprises privées pour entraîner des LLM crée un déséquilibre fondamental dans l'écosystème de l'accès ouvert : il donne lieu à une captation de la valeur générée par les ressources en accès ouvert, sans réciprocité. En effet, les IA génératives sont entraînées sur des données en accès ouvert, mais elles ne réorientent pas leurs usagers vers ces contenus, ou et elles ne les citent même pas. White identifie ainsi un sentiment de détournement pour les acteurs s'engageant dans l'accès ouvert⁶⁶. La libre circulation des savoirs et des connaissances, est exploitée par des acteurs privés pour développer des produits monétisés. Cela génère un dilemme fondamental : pour rester performants, les systèmes d'IA s'appuient sur une importante masse de données publiques. Pourtant, les producteurs et gestionnaires de ces informations bénéficient rarement des revenus générés par leur exploitation. Qui plus est, les cadres juridiques permettant l'utilisation de contenus pour l'IA ne prévoient aucun mécanisme de soutien économique aux producteurs d'information, brisant ainsi la relation de confiance entre ceux qui alimentent l'écosystème de la connaissance et ceux qui en profitent financièrement⁶⁷. Ce décalage risque d'aboutir à ce que Hardinges et al. qualifient de « rupture du contrat social » sur lequel repose l'*open access*. Jean-Philippe Moreux, responsable de la mission IA de la BnF illustre ce potentiel de dérèglement pour les institutions du patrimoine :

On peut dessiner des scénarios pessimistes où les opérateurs du patrimoine produisent de la donnée qui est valorisée par les entreprises privées. Et les opérateurs producteurs de données ne rencontrent plus leur public, puisque

⁶⁴« Processus de développement de ChatGPT et de nos modèles de fondation », sur *Openai.com*, s. d. (en ligne : <https://openai.com/fr-FR/policies/how-chatgpt-and-our-foundation-models-are-developed/> ; consulté le 28 janvier 2026)

⁶⁵P. KELLER, « Beyond AI and copyright. Funding a sustainable information ecosystem », *op. cit.*, p. 4

⁶⁶M. WHITE, « “Wait, not like that” », *op. cit.*

⁶⁷P. KELLER, « Beyond AI and copyright. Funding a sustainable information ecosystem », *op. cit.*, p. 8

celui-ci est capturé par les plateformes en question. Cela devient un cercle vicieux qui questionne les fondements de ces opérateurs du patrimoine : pourquoi l'État financerait-il une activité qui ne rencontre pas ses utilisateurs ?⁶⁸

La captation des informations par les entreprises d'IA générative peut créer des incitations négatives à l'entretien et l'alimentation d'institutions et de plateformes des savoirs en accès ouvert. Or, nous l'avons vu, les technologies d'IA générative reposent, pour leur bon fonctionnement, sur ce contrat social. Le déséquilibre risque donc, à terme, de mettre en danger les fondements mêmes ayant permis le développement de cet écosystème⁶⁹.

Le passage d'un accès direct et ouvert aux sources scientifiques vers un accès médiatisé par des systèmes d'IA propriétaires constitue ainsi un bouleversement structurel dont les implications se manifestent, notamment sur un point technique : la sursollicitation des infrastructures ouvertes. Le moissonnage intensif des contenus en libre accès pose des défis techniques considérables pour les infrastructures qui hébergent ces contenus. D'après Arnaud Laborderie⁷⁰, 10 % du trafic sur Gallica est aujourd'hui le fait de moissonnage depuis internet. Les « siphonnages » ont été facilités par l'existence d'API ouvertes développées dans une logique de science ouverte, notamment pour permettre à des chercheurs de faire de la fouille de texte. Les pratiques de moissonnage ont profité de ces politiques de développement des API en détournant leur finalité. Or, ces moissonnages intensifs ont des impacts négatifs sur le bon fonctionnement de la plateforme : en sollicitant à l'excès les serveurs, ils entraînent un ralentissement du service et ont, par le passé, pu causer une mise à l'arrêt complète du site de Gallica. Nathalie Clément, d'EDP Sciences, note également : « Un usage excessif de nos contenus vient impacter la fluidité d'accès pour nos propres abonnés. Nous sommes donc obligés de nous protéger en identifiant les adresses IP de consommateurs excessifs et en bloquant ces adresses IP, pour permettre à nos véritables abonnés et clients d'accéder au contenu⁷¹ ». L'ampleur du phénomène est confirmée par une enquête menée par la COAR (coalition de dépôts numériques) sur les conséquences du moissonnage pour les infrastructures de ces dépôts. 80 % des répondants indique que les moissonnages et robots ont entraîné des perturbations de services, de ralentissements à des arrêts de services pouvant parfois durer plusieurs jours⁷².

Ces pratiques ont entraîné la mise en place de nouvelles mesures de régulation et de contrôle pour réguler les accès. L'évolution des modalités d'accès au contenu ouvert témoigne de la fin d'une phrase d'optimisme. Il existe des moyens technologiques pour se protéger du moissonnage : firewalls, blocage d'adresses IP spécifiques des robots IA identifiés. Leur mise en place a néanmoins des conséquences. Tout d'abord, un temps humain important doit être dédié à l'adaptation des systèmes. Le rapport de la COAR mentionne ainsi une moyenne de 10 % d'un ETP consacré à ces tâches, allant pour certaines plateformes jusqu'à 50 %⁷³. Par ailleurs, ces blocages n'ont qu'une capacité limitée à discriminer entre les différents types de robots et peuvent aussi limiter des initiatives vertueuses. Certains dépôts ont commencé à maintenir des « listes blanches » pour éviter de bloquer l'accès à de telles initiatives.

⁶⁸ J.-P. MOREUX, « Entretien », 28 novembre 2025

⁶⁹ J. HARDINGES, S. PEARSON et R. ROSS, « From Human Content to Machine Data », 2025 « In putting the commons at risk, large AI models are 'drilling away at their own foundations.' »

⁷⁰ A. LABORDERIE, « Entretien », 3 décembre 2025

⁷¹ S. LA BARBERA et N. CLÉMENT, « Entretien », *op. cit.*

⁷² K. SHEARER et P. WALK, *The impact of AI bots and crawlers on open repositories: Results of a COAR survey, April 2025*, Confederation of Open Access Repositories, 2025, p. 4

⁷³ *Ibid.*, p. 5

En parallèle de la sursollicitation par des robots, on observe une moindre fréquentation directe de certaines sources. Les sources primaires ne sont plus consultées directement par des usagers humains, ceux-ci se limitant aux synthèses et agrégations d'informations fournies par les IA génératives. Hardinges et al. citent en exemple la baisse de 25 % du nombre de visites sur le second site d'Histoire, World History Encyclopedia, depuis l'introduction de l'outil de Google AI Overviews en 2024⁷⁴. Les lecteurs se contentent de consulter les résumés et synthèses qui leur sont fournis par les IA génératives. Cela pose des questions de soutenabilité sur le long terme. La double dynamique de sursollicitation technique par les robots d'un côté, et de désaffection des usagers humains, de l'autre, menace directement la viabilité des infrastructures d'accès ouvert.

Elle crée, notamment des désincitations à générer de nouveaux savoirs originaux de qualité, et entraîne une dégradation de la qualité des données et des savoirs publiés. L'omniprésence potentielle de l'IA générative dans tous les maillons de la publication scientifique peut aboutir à une situation dans laquelle l'humain ne participe plus au processus, avec des publications écrites, évaluées, découvertes et lues par des IA génératives. Cette perspective engendre, finalement, un risque d'« empoisonnement » des données. Outre la question des impacts de ces savoirs erronés pour les lecteurs humains, cela compromet la qualité des corpus d'entraînement futurs pour les technologies d'IA génératives. Les effets sont multiples : renforcement des biais et appauvrissement de la diversité des connaissances et idées, accumulation d'erreurs. Leur accumulation risque, à terme, de causer un « effondrement des modèles⁷⁵ ». Dans le cas de la publication scientifique, l'inflation d'articles frauduleux interroge particulièrement. Si les pratiques de rétractation se multiplient afin de limiter la diffusion de la fake science, elles sont, notamment, insuffisantes pour faire disparaître les publications prédatrices d'internet, comme le note Kosmopoulos : « Malgré les rétractions des articles, les références et les citations restent présentes sur le web, dans d'autres articles, dans les résultats des robots conversationnels... et même dans des moteurs de recherche *a priori* sérieux comme Google Scholar⁷⁶. » La présence de publications de faible valeur scientifique, ou ne respectant pas les règles d'intégrité scientifique, est particulièrement dangereuse dans la mesure où ces publications font partie de la masse de données sur lesquelles s'entraînent les IA génératives, créant ainsi une boucle de dégradation progressive.

Tentation de la fermeture et ses limites

Face aux déséquilibres constatés, la tentation de réinstaurer des verrous et de protéger les données émerge. L'impression d'être lésé, de voir son travail détourné, en particulier de voir les principes de l'accès ouvert récupérés par des entreprises privées ayant des motivations commerciales, suscite une volonté de récupérer du contrôle. White résume cette réaction par la formule « *wait, no, not like that* », qui capture le décalage entre les intentions initiales de l'*open access* et

⁷⁴J. HARDINGES, S. PEARSON et R. ROSS, « From Human Content to Machine Data », *op. cit.*, p. 9

⁷⁵I. SHUMAILOV *et al.*, « The Curse of Recursion: Training on Generated Data Makes Models Forget », *ArXiv*, 27 mai 2023

⁷⁶C. KOSMOPOULOS, « Entre prédation et désincarnation, l'édition scientifique à l'épreuve de l'IA », *Cybergeography: European Journal of Geography*, CNRS-UMR Géographie-cités 8504, 13 février 2025 (en ligne : <https://journals.openedition.org/cybergeography/41699> ; consulté le 10 juin 2025)

les usages commerciaux actuels⁷⁷. Ces réactions individuelles sont le symptôme d'un dérèglement structurel de l'équilibre de l'écosystème informationnel, réactivant des réflexes qui vont à l'encontre des principes de science ouverte. Parmi les acteurs qui se trouvent lésés par ce dilemme, Keller cite notamment les scientifiques qui publient leurs recherches en accès libre.

Cette tension entre ouverture et protection se manifeste également au niveau institutionnel. Un conservateur témoigne des réticences exprimées par les vice-présidents recherche de son établissement lors de la refonte de la charte de science ouverte. Alors que la précédente version ne comportait pas d'engagements concrets, la nouvelle mouture devait inclure dix engagements précis. Parmi ceux-ci, la formulation « favoriser le dépôt systématique des publications dans HAL » a suscité des résistances : les VP recherche ont demandé à retirer le terme « systématique ». L'émergence des IA génératives grand public semble avoir intensifié cette frilosité préexistante vis-à-vis de la science ouverte, conduisant à l'ajout d'un engagement spécifique invitant à « réfléchir à la problématique de l'IA vis-à-vis de la science ouverte ». La formulation traduit davantage une inquiétude qu'une orientation claire.

Pourtant, ce réflexe protectionniste présente des limites importantes. Comme le soulignent Hardinges et al., il ne protégera pas nécessairement les créateurs de contenus. Nombre d'entre eux ayant cédé leurs droits à des plateformes de diffusion ou à des éditeurs commerciaux, la limitation de l'accès ouvert et le renforcement des protections par le droit d'auteur risquent de bénéficier principalement à ces intermédiaires plutôt qu'aux créateurs eux-mêmes⁷⁸. White exprime également cette tension : outre le problème individuel de la captation sans réciprocité, c'est l'infrastructure dans son ensemble qui est en jeu. La fermeture risquerait ainsi de mettre en péril le web ouvert dans sa globalité. Le réflexe de fermeture briserait la dynamique vertueuse de l'accès ouvert pour la recherche.

Adaptations

Face à ces tensions, il est nécessaire de repenser les conditions du partage de l'information et ce « contrat social » sur lequel repose la science ouverte. L'ADBU adopte une position claire en faveur du maintien de l'ouverture. Dans son argumentaire « Ouvrir la science à l'heure de l'intelligence artificielle : pourquoi et comment ? », l'association rappelle que la fermeture ne constitue pas une protection efficace : dès lors qu'un auteur a cédé ses droits, l'éditeur peut choisir d'exploiter la publication pour entraîner une IA générative ou vendre des corpus à des tiers. Le retour à l'accès fermé ne résout donc pas les difficultés posées par l'exploitation des publications scientifiques par les IA génératives. L'ADBU soutient au contraire que la poursuite de l'ouverture de la science offre les meilleures garanties pour la transparence et la reproductibilité de la recherche, tout en permettant la diversité des solutions d'IA générative et leur adaptation aux spécificités de la recherche⁷⁹.

Face aux déséquilibres engendrés par l'exploitation massive des contenus en libre accès par les systèmes d'intelligence artificielle, plusieurs initiatives émergent pour restaurer les conditions d'un contrat social équitable. Une première voie consiste à développer des infrastructures publiques éthiques garantissant une juste rétribution des

⁷⁷M. WHITE, « “Wait, not like that” », *op. cit.*

⁷⁸J. HARDINGES, S. PEARSON et R. ROSS, « From Human Content to Machine Data », *op. cit.*, p. 18

⁷⁹ADBU, « Ouvrir la science à l'heure de l'intelligence artificielle : pourquoi et comment ? », s. d.

pratiques de publication en accès ouvert. Paul Keller, de l'Open Future Foundation, propose ainsi l'instauration d'un modèle de rémunération déclenché lors de la commercialisation de modèles d'IA entraînés sur des données en libre accès, permettant ensuite une redistribution des ressources vers les producteurs de contenus⁸⁰. Wikimedia Enterprise a mis en œuvre un dispositif de cette nature : les entreprises d'IA peuvent utiliser les données de l'encyclopédie collaborative, mais en empruntant des canaux de circulation dédiés. Ce système présente un double avantage : il préserve la fluidité des services proposés au grand public tout en faisant participer financièrement les entreprises aux coûts des infrastructures qui rendent possible l'entraînement de leurs modèles⁸¹. En France, la Bibliothèque nationale de France a lancé Data IA, un service imposant aux acteurs commerciaux le paiement d'une redevance d'utilisation pour contribuer à l'effort de numérisation patrimoniale, tout en maintenant la gratuité d'accès pour les projets de recherche⁸².

Une autre approche consiste à repenser les métadonnées pour les adapter aux usages des systèmes d'IA générative. L'association Creative Commons développe ainsi le projet « CC Signals », un système de signalement qui viendrait compléter les déclarations de droits existantes afin de permettre aux auteurs de spécifier quels usages d'entraînement ils autorisent. L'enjeu est d'introduire une granularité dans les droits d'utilisation, en distinguant par exemple la fouille de textes et de données, l'entraînement d'IA non génératives ou de grands modèles de langage. Il s'agit également de chercher à établir des règles de réciprocité propres à maintenir l'équilibre de l'écosystème de l'information en libre accès⁸³.

Ces différentes initiatives témoignent d'une capacité d'adaptation de l'écosystème de la science ouverte, qui cherche à concilier le développement des technologies d'IA générative avec le maintien des conditions de réciprocité sur lesquelles repose le partage des connaissances.

On le voit, les bouleversements induits par les technologies de l'IA générative dans l'écosystème de la publication scientifique sont multiples et de natures extrêmement variées : ils sont techniques, juridiques, économiques, épistémiques et éthiques, et menacent l'équilibre entier de l'écosystème. Tout l'enjeu réside, face à ces questions, dans les rééquilibrages et dynamiques à l'œuvre afin de maintenir la confiance et la viabilité de l'écosystème de l'information scientifique. Ceci repose à la fois sur une prise en compte des possibilités technologiques et sur les efforts de normalisation. Une question se pose ensuite : quel rôle peuvent jouer les bibliothécaires et les institutions documentaires pour préserver les acquis de l'Open Access tout en s'adaptant aux réalités imposées par l'IA ? Comment peuvent-ils contribuer à maintenir le contrat

⁸⁰P. KELLER, « Beyond AI and copyright. Funding a sustainable information ecosystem », *op. cit.*, p. 11

⁸¹M. WHITE, « "Wait, not like that" », *op. cit.*

⁸²« L'offre de services de la BnF pour les acteurs de l'intelligence artificielle (IA) », sur *BnF - Site institutionnel*, s. d. (en ligne : <https://www.bnf.fr/fr/loffre-de-services-de-la-bnf-pour-les-acteurs-de-lintelligence-artificielle-ia> ; consulté le 2 février 2026)

⁸³J. HARDINGES, S. PEARSON et R. ROSS, « From Human Content to Machine Data », *op. cit.* p. 25

social de la science ouverte dans ce contexte de recomposition profonde ? C'est à ces questions que nous nous attacherons désormais.

LES ÉVOLUTIONS TECHNOLOGIQUES DANS LA DOCUMENTATION ÉLECTRONIQUE

Après ce premier temps cherchant à décrire le champ de la documentation électronique et ses recompositions suite à l'arrivée des technologies d'IA générative, le deuxième temps de ce travail vise à approfondir le fonctionnement des solutions technologiques pour expliciter leurs possibilités et limitations. Y sont notamment déclinées quelques-unes des solutions développées par des acteurs privés et publics, qui témoignent de logiques différentes. Un dernier temps interroge la place des professionnels des bibliothèques face à ces questions. Il s'agit donc de montrer que les innovations techniques s'articulent également avec des considérations économiques et informationnelles qui posent des questions concernant le travail autour de l'IST.

AUX FONDEMENTS TECHNOLOGIQUES DE L'IA GÉNÉRATIVE

La capacité impressionnante à émuler des actions humaines permise par les outils d'IA générative doit faire l'objet d'une analyse critique. Il convient d'approfondir le fonctionnement de l'IA générative pour comprendre la nature de ces outils et des résultats qu'ils permettent. La compréhension des fonctionnements techniques est nécessaire afin de mesurer les avantages et les inconvénients de tels outils et leurs capacités effectives.

Des outils statistiques ayant des limites structurelles

Les LLM qui sous-tendent des systèmes d'IA générative, que ce soit des chatbots généralistes ou des outils plus spécialisés, reposent sur des modèles statistiques et probabilistes d'analyse du langage humain. Les LLM sont entraînés sur des ensembles de données massifs, avec des réseaux de neurones qui apprennent eux-mêmes à classer les données, et construisent des liens probabilistes à partir de l'analyse des données d'entraînement : il s'agit du *deep learning*⁸⁴. Ce fonctionnement a des conséquences importantes pour l'évaluation des tâches réalisées par les IA génératives. La dimension d'auto-entraînement au cœur des LLM implique un effet « boîte noire » : les modèles statistiques sont construits à partir de traitement de données qui dépassent l'entendement humain. Les processus d'apprentissage des modèles relèvent en grande partie de méthodes non supervisées, ce qui signifie que les humains n'ont pas de visibilité complète sur la manière dont les systèmes construisent leurs inférences : l'explicabilité des contenus générés par ces technologies demeure limitée. Cela n'implique pas une absence totale d'intervention humaine, celle-ci intervenant après l'apprentissage profond, pour opérer des opérations d'ajustement, dites de *fine-tuning*. Il s'agit cependant plus d'ajustements que d'interventions sur le cœur même des modèles probabilistes. Les traitements statistiques sont appliqués à deux niveaux complémentaires dans le cas des IA générative, comme le synthétise Bouchard :

⁸⁴F. TÉTART, « L'intelligence artificielle appliquée aux bibliothèques : une opportunité plutôt qu'une menace », *I2D - Information, données & documents*, n° 1, A.D.B.S., 3 septembre 2025, p. 81

« la “compréhension” du texte et la génération de contenu⁸⁵ ». La capacité d'évaluation et de regard critique humain sur ces outils est donc, en partie, limitée. Une telle opacité est importante à garder en tête dans le contexte de la recherche scientifique, où la traçabilité des raisonnements et la transparence méthodologique participent de la qualité du travail. Ces processus de *deep learning* sont, par ailleurs, extrêmement coûteux en énergie, les besoins de calcul et de stockage de ces processus augmentant de manière exponentielle⁸⁶. Ils posent donc également des questions de sobriété énergétique.

De plus, la nature probabiliste des LLM a des implications sur les capacités de traitement des informations proposées par ces technologies. Si celles-ci miment des propos qui semblent avoir une qualité humaine, il n'y a, en fait, pas de compréhension effective ni des informations qui sont traitées, ni des résultats qui sont générés. Bernard Jacquemin souligne cette réalité fondamentale : l'évolution technologique, aussi « époustouflante » soit-elle, « n'amène en réalité aucun progrès dans la capacité de la machine à réellement comprendre les éléments d'information ou la donnée qu'on lui confie⁸⁷ ». La métaphore du « perroquet stochastique⁸⁸ », utilisée pour qualifier le fonctionnement de ces technologies, insiste sur le fait que les productions des LLM sont ainsi de purs produits statistiques, et non des raisonnements au sens épistémique du terme. Pour le dire autrement, il n'y a, en fait, pas d'intelligence réelle dans l'intelligence artificielle, il n'y a pas de compréhension par les outils mais uniquement une réponse statistique. Ainsi, les réponses générées par l'IA ne sont pas construites à partir de critères de véracité ou d'exactitude. Il s'agit exclusivement de réponses qui paraissent pertinentes et logiques au regard des données d'entraînement. Cela se manifeste par des réponses fausses, mais qui ont une apparence de cohérence, communément appelées « hallucinations ». « Ce qui est important, c'est qu'il y ait une réponse quelle qu'elle soit et que cette réponse ressemble le plus possible à une formulation humaine » note Jacquemin⁸⁹. L'un des éléments marquants dans les hallucinations est qu'elles ont l'apparence d'une réponse correcte et vraie, mais qu'elles ne correspondent pas à la réalité. Un exemple de plus en plus fréquent, en bibliothèque, est un lecteur ou une lectrice cherchant des références bibliographiques trouvées grâce à un outil d'IA générative, mais n'existant pas. Un outil comme ChatGPT est ainsi capable d'halluciner des titres d'ouvrages, voire de fournir des ISBN fictifs, et donne des réponses qui paraissent si convaincantes qu'un usager non averti ne discernera pas l'erreur⁹⁰. L'usage massif de ces outils comme substituts aux moteurs de recherche revient, selon Jacquemin à « utiliser un marteau pour enfoncer une vis⁹¹ ».

⁸⁵A. BOUCHARD, « ”#WorkInProgress », *op. cit.*, p. 4

⁸⁶A.-L. LIGOZAT *et al.*, « Unraveling the Hidden Environmental Impacts of AI Solutions for Environment Life Cycle Assessment of AI Solutions », *Sustainability*, vol. 14, n° 9, publisher, 25 avril 2022 (DOI : 10.3390/su14095172 consulté le 6 février 2026)

⁸⁷B. JACQUEMIN, « Entretien », *op. cit.*

⁸⁸E. M. BENDER *et al.*, « On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? », dans *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, Virtual Event Canada, ACM, 2021, p. 610-623 (en ligne : <https://dl.acm.org/doi/10.1145/3442188.3445922> ; consulté le 18 février 2026)

⁸⁹B. JACQUEMIN, « Entretien », *op. cit.*

⁹⁰G. BRANDY, « Aux Etats-Unis, des bibliothécaires désemparés face aux demandes de livres inventés par l'IA », *Le Monde*, 24 janvier 2026 (en ligne : https://www.lemonde.fr/culture/article/2026/01/24/aux-etats-unis-des-bibliothecaires-desempares-face-aux-demandes-de-livres-inventes-par-l-ia_6663922_3246.html ; consulté le 6 février 2026)

⁹¹B. JACQUEMIN, « Entretien », *op. cit.*

Ajoutons également que la nature des données utilisées pour entraîner les LLM est déterminante pour le fonctionnement de ces technologies. À ce titre, elles sont très dépendantes de jeux de données de qualité pour leur entraînement, comme l'illustre l'expression *garbage in, garbage out* : si les modèles sont entraînés à partir de contenus biaisés, ou limités, ils reproduisent mécaniquement les biais ou les limitations de leurs données initiales. Cela pose des questions importantes quant aux choix de limiter les entraînements sur certains corpus : par exemple, si les LLM ne sont pas entraînés sur des corpus français, cela entraîne des angles morts dans les résultats fournis par les IA génératives.

Un autre élément est à prendre en considération concernant les données d'entraînement : la capacité à les situer en contexte et à identifier les différences de ton, par exemple en fonction des situations d'énonciation. Arnaud Laborderie observe ainsi que ces outils ne disposent pas d'une compréhension contextuelle des documents : ils ne saisissent pas les spécificités d'un contexte historique et énonciatif, ni les degrés d'ironie ou de nuance des textes⁹². Cette lacune est significative lorsqu'il s'agit de traiter des corpus scientifiques, où la signification d'un énoncé dépend souvent de son inscription dans un champ disciplinaire, un débat académique ou un cadre méthodologique spécifique.

La technologie RAG : réduire les biais par la sélection documentaire

Malgré les limites de ces modèles, leur potentiel technologique est considérable. George souligne que ces technologies « peuvent automatiser la génération de métadonnées, améliorer les systèmes de recommandation de contenus et renforcer la découvrabilité des ressources⁹³ », avec des conséquences significatives sur la manière dont les bibliothèques gèrent et traitent leurs collections. Une des réponses potentielles aux limites de ces technologies est de développer des outils ciblés, conçus de manière à limiter les biais et à renforcer les capacités d'évaluation des réponses apportés.

Parmi les développements les plus prometteurs, la technologie RAG (*Retrieval Augmented Generation*) ouvre des perspectives pour construire des solutions technologiques qui gagnent notamment en transparence et limitent les réponses hallucinées. Les systèmes RAG sont entraînés sur un modèle LLM préexistant, mais dans la constitution d'une réponse, elles fonctionnent en deux temps. La première phase est celle de la récupération (*retrieval*), au cours de laquelle le système recherche des documents pertinents en fonction de la requête de l'utilisateur. Le système cherche dans une base de connaissances externe fournie, plus réduite que les données d'entraînement, donc plus facile à contrôler, et potentiellement plus à jour. Dans un second temps, la phase de génération, le RAG génère une réponse fondée uniquement sur les documents récupérés. Ce mécanisme permet d'extraire des réponses à partir d'un corpus spécifique, de sourcer les réponses produites, et l'outil peut refuser de répondre en cas d'absence d'information pertinente dans le corpus. Cela réduit ainsi les hallucinations et les biais de réponse en se focalisant sur un type d'information particulier, le corpus restreint permettant une évaluation par un observateur humain et une vérification. Le traitement d'une quantité de données limitées peut, par ailleurs, participer d'une réduction des coûts de calcul de ces solutions technologiques. Bernard Jacquemin souligne que

⁹²A. LABORDERIE, « Entretien », *op. cit.*

⁹³S. GEORGE, « Impacts of Open Access and Open Repositories on Acquisitions and Collection Development », *International Journal of Social Science Humanity & Management Research*, vol. 3, n° 08, 23 août 2024, p. 1123 (DOI : 10.58806/ijsshr.2024.v3i8n15 consulté le 20 mai 2025)

l'« enfermement dans une base de connaissances plus réduite et plus vérifiée permet d'obtenir des réponses dont le niveau de qualité est supérieur⁹⁴ ». Cette observation met en lumière une convergence possible entre les exigences documentaires des bibliothèques et les atouts techniques des systèmes RAG.

On le voit à travers ces explications, le fonctionnement technique des outils a des conséquences sur les résultats qu'ils proposent. Les fonctionnements statistiques, fondés sur des grandes quantités de données, sans prise en compte de critères de véracité ou de contextualisation pointent vers des limites et témoignent de l'importance de comprendre les spécificités des outils pour adapter les utilisations. Or, comme nous allons le voir maintenant, les champs des possibles sont très larges en ce qui concerne l'utilisation des IA pour accompagner la production de savoir. Elles sont développées par des entreprises privées comme par des initiatives publiques.

UN CONTINUUM DE SOLUTIONS PRIVÉES

L'objectif ici n'est pas de dresser un relevé exhaustif des outils disponibles développés par des entreprises privées. Une telle tâche ne serait pas possible du fait à la fois du grand nombre de solutions proposées sur le marché et de l'évolution constante de l'offre technologique. Il importe, néanmoins, de décrire quelques-unes des solutions aujourd'hui existantes, afin d'illustrer la variété des usages potentiels, mais également les acteurs qui développent de telles solutions. En effet, les outils et les domaines d'application de l'IA générative pour la production de savoirs scientifiques sont multiples, les technologies pouvant être utilisées de manière ponctuelle ou, au contraire, de manière très large. Des questions différentes se posent en fonction de leur champ d'application.

Outils d'aides à la rédaction et à la soumission

Une première catégorie d'outils permet d'accélérer et d'améliorer les processus de mise en forme et de rédaction. Ces outils proposent des corrections formelles, orthographiques et grammaticales, mais également des interventions sur le fond, en proposant des changements visant à améliorer la fluidité et l'efficacité des formulations. Un service comme Grammarly propose ainsi à la fois de la relecture ortho-grammaticale et un accompagnement pour adapter le ton d'un texte ou sa clarté⁹⁵. Service similaire d'aide à la rédaction, Prowriting Aid propose, outre la relecture, des analyses et conseils pour aider l'utilisateur humain à améliorer son style et sa prose⁹⁶. Cet outil met en outre l'accent sur la protection des écrits versés par l'utilisateur, tout comme Trinka AI, autre outil concurrent⁹⁷.

⁹⁴B. JACQUEMIN, « Entretien », *op. cit.*

⁹⁵« Our Features | Grammarly », s. d. (en ligne : <https://www.grammarly.com/features> ; consulté le 6 février 2026)

⁹⁶« Great writing made easy with ProWritingAid », s. d. (en ligne : <https://prowritingaid.com/> ; consulté le 6 février 2026)

⁹⁷« Trinka: AI Writing and Grammar Checker Tool », s. d. (en ligne : <https://www.trinka.ai> ; consulté le 6 février 2026)

D'autres plateformes, comme UNSILO, utilisent l'apprentissage automatique pour fluidifier les processus de dépôt de manuscrit, la sélection des évaluateurs et la vérification des contenus⁹⁸. Ces outils interviennent donc en amont de la publication, dans des étapes traditionnellement chronophages du circuit éditorial, à savoir la rédaction ou le contrôle éditorial.

Déjà, on le voit, la frontière entre les usages pour des améliorations de forme et des améliorations de fond devient poreuse. Ces outils ne se limitent pas à la vérification grammaticale : ils proposent également des suggestions en matière de clarté, de concision et de conformité aux styles d'écriture académique. Les présupposés de la rhétorique classique qui lient fond et forme sont questionnés. Pensons en particulier à l'idée, résumée par l'adage de Nicolas Boileau que « ce qui se conçoit bien s'énonce clairement/ et les mots pour le dire viennent aisément⁹⁹ ». Il ne semble plus si véridique à une époque où des outils d'IA générative peuvent venir altérer le processus même d'énonciation et proposer des actions pour améliorer leur clarté, sans que l'idée de l'auteur ne soit, en elle-même, changée.

Outils de découvrabilité

Une seconde catégorie d'outils concerne les aides à découverte des publications scientifiques. De tels outils ont vocation à assister les chercheurs, en particulier dans l'étape de revue de littérature. Comme le note Bouchard, deux types d'acteurs existent dans le paysage de la recherche documentaire : les acteurs déjà existant dans le paysage de la recherche académique, qui développent de nouvelles fonctionnalités, boostées à l'IA, et les *start ups* créant de nouveaux services, et cherchant à trouver leur place dans les chaînes de découverte¹⁰⁰. Encore une fois, les technologies d'IA sont implémentées à des degrés variables. Il peut s'agir de simples améliorations des fonctionnalités usuelles des outils de recherche documentaire, mais aussi de changements plus profonds dans les mécanismes de recherche, avec par exemple l'intégration de la recherche sémantique, qui traite des requêtes en langage naturel plutôt que des requêtes formées à partir d'une sélection de mot-clés. D'autres outils proposent une expérience de recherche augmentée, fournissant en sus des résultats bibliographiques, des résumés et des synthèses des travaux trouvés. Ces solutions sont plus ou moins transparentes et compréhensibles, et plus ou moins orientées vers la valorisation de l'accès ouvert.

Parmi les nouvelles solutions de découverte, trois nouveaux venus, développés par des start ups, peuvent servir d'exemple : Consensus, Elicit et Scispace. En janvier 2026, Consensus, qui aspire à permettre une « revue de littérature approfondie » revendique un partenariat avec 170 bibliothèques universitaires, et plus de 10 millions d'utilisateurs¹⁰¹. Elicit présente son service comme un outil permettant d'être « 10 fois plus fondé sur des preuves¹⁰² » et annonce être utilisé par 5 millions d'utilisateurs. On ne trouve pas de données quantitatives du nombre d'utilisateurs sur le site de Scispace, outil qui se propose de « gérer les tâches quotidiennes de la recherche avec des résultats fiables et sourcés¹⁰³ ». Ces données d'utilisation attestent d'une adoption de ces outils par certains

⁹⁸L. S. LO, « Generative AI and Open Access Publishing », *op. cit.*, p. 169

⁹⁹N. BOILEAU DESPRÉAUX, *Oeuvres poétiques*, Paris, L'Imprimerie Générale, 1872, p. 218

¹⁰⁰A. BOUCHARD, « "#WorkInProgress », *op. cit.*, p. 14

¹⁰¹« Consensus: AI for Research », s. d. (en ligne : <https://consensus.app/search/> ; consulté le 10 février 2026)

¹⁰²« Elicit: AI for scientific research », s. d. (en ligne : <https://elicit.com/welcome> ; consulté le 10 février 2026)

¹⁰³« SciSpace AI Research Agent | 150+ Tools, 280 M Papers », sur *SciSpace*, s. d. (en ligne : <https://scispace.com> ; consulté le 10 février 2026)

acteurs du monde de la recherche, même si elles participent, aussi, de stratégies de communication et de marketing.

Pour ce qui est des outils développés par des entreprises déjà établies dans l'écosystème de la publication scientifique, nous choisirons l'exemple de Cairn. La plateforme mène des projets d'IA, notamment SophIA, présenté comme un « moteur de recherche expérimental utilisant l'intelligence artificielle pour faciliter la découverte des publications et des auteurs sur Cairn.info ». Cairn annonce clairement des règles déontologiques : ne pas transmettre les contenus de ses partenaires aux LLM exploités par les grands acteurs de l'IA, protéger ces contenus technologiquement et juridiquement, utiliser l'IA générative pour faciliter l'accès au corpus scientifique plutôt que comme alternative à la lecture des textes originaux, associer ses partenaires dans une approche transparente et collaborative, et privilégier la sobriété numérique. On observe ici une tentative de répondre à un certain nombre de questions que posent les IA génératives, en termes juridiques, de sobriété écologique et de rapport à l'information. Le choix de « privilégier l'accès au corpus scientifique » témoigne notamment d'un souci de préserver les pratiques de bonne recherche¹⁰⁴.

L'affichage de lignes déontologiques, du nombre d'utilisateurs ou d'institutions partenaires témoignent du besoin qu'ont les entreprises de démontrer la pertinence de leur service et de gagner la confiance des utilisateurs. La communication de ces outils est bien rodée : ils se proposent d'accélérer et de faciliter le travail de recherche, sans mentionner la manière dont ils peuvent, potentiellement, venir l'altérer et en changer les fondamentaux. Ces bouleversements, que l'on pourrait encore considérer comme minimes ou marginaux dans le cas d'outils de recherche documentaire, deviennent cependant plus visibles lorsque l'on considère un troisième type d'outils, qui proposent d'accompagner le chercheur dans l'ensemble des étapes de son travail.

Des outils holistiques : de la recherche à l'écriture

Des outils intégrés proposent d'accompagner l'ensemble du processus de recherche, de la découverte bibliographique à la rédaction finale. Ces solutions technologiques ne se présentent plus comme des aides à la découverte ou à la rédaction, mais comme de véritables assistants de recherche automatisant un ensemble de tâches diverses, que cela soit la revue scientifique, la rédaction, ou la mise en forme. Les offres sont, de nouveau, développées à la fois par des acteurs déjà établis et des nouveaux venus dans le monde de la publication scientifique.

Prism, lancé par OpenAI en janvier 2026, se présente comme « conçu pour accélérer le travail scientifique au quotidien¹⁰⁵ », en proposant une relecture assistée par l'IA, des fonctionnalités de recherche de références et d'automatisation de la mise en forme. La promesse est celle d'un gain de temps sur la forme permettant de se concentrer sur le fond. Scopus AI, développé par Elsevier, propose également une variété de services avec la même promesse de faciliter le travail de recherche. Parmi celles-ci, notons la composition de résumés

¹⁰⁴ « Comment utiliser SophIA ? - Aide Cairn.info », s. d. (en ligne : <https://aide.cairn.info/article/378-comment-utiliser-sophia> ; consulté le 5 février 2026)

¹⁰⁵ « Prism », s. d. (en ligne : <https://openai.com/fr-FR/prism/> ; consulté le 4 février 2026)

thématiques sur des thèmes de recherche spécifiques, la génération de cartes mentales pour aider à la visualisation de certains concepts techniques, et un agent conversationnel qui permettrait de « formuler ses propres questions de recherche en langage du quotidien¹⁰⁶ ». L'entreprise adopte une stratégie de communication qui cherche à répondre aux questionnements du monde académique en se présentant comme un outil « *not for the lazy* » et en énonçant des principes de transparence et de fiabilité. L'outil met également en avant des garanties en matière de protection des données, incluant la conformité au RGPD européen et un engagement à ce que les requêtes des utilisateurs ne soient ni stockées ni utilisées pour entraîner les LLM sous-jacents. Ces stratégies communicationnelles, si elles témoignent d'une conscience des enjeux, doivent néanmoins être examinées avec discernement.

Enjeux critiques des solutions privées

Les quelques exemples proposés ci-avant illustrent le foisonnement des propositions technologiques. Point notable, alors que certains outils semblent se concentrer sur une tâche précise et développer des outils ciblés, d'autres ont l'ambition de centraliser en un service unique une variété de tâches. On note, par ailleurs, la variété des acteurs commerciaux : tandis que certains viennent du monde de l'édition scientifique, à l'instar d'Elsevier, d'autres viennent du monde des nouvelles technologies. La variété des acteurs témoigne d'une série d'opportunités ouvertes par les nouvelles technologies, et d'un marché à la fois en rapide évolution et plein d'opportunités.

Ces outils cherchent à répondre aux besoins des communautés de chercheurs et adoptent des stratégies de communication qui tentent de désamorcer les inquiétudes du monde de la recherche en ce qui concerne le respect des données, du droit d'auteur, ou de l'intégrité scientifique. Qu'il s'agisse d'acteurs déjà connus ou de nouveaux venus, il convient de souligner qu'il s'agit, pour l'essentiel, de services développés par des entreprises privées à but lucratif. Cela pose la question de l'autonomie et de la neutralité de la recherche. Le possible développement d'une situation de dépendance vis-à-vis d'acteurs ayant des intérêts commerciaux peut aller à l'encontre des principes d'intégrité de la recherche.

Les coûts de ces solutions, par ailleurs, interrogent. Bien souvent, une hiérarchie s'établit entre des services gratuits, plus limités, et des services payants offrant des fonctionnalités avancées. Les questions autour de telles structures tarifaires sont encore plus fortes dans les contextes budgétaires actuels. Les établissements publics qui financent la recherche sont, aujourd'hui dans des situations financières souvent précaires. Elles ont peu de marges de manœuvre financières et l'abonnement à des services d'IA payants peut nécessiter des arbitrages de dépenses et des réallocations de crédits. L'apparition de nouveaux outils monétisés peut, en outre, renforcer les inégalités d'accès entre les institutions et les pays d'accueil des chercheurs, en fonction de leurs moyens.

¹⁰⁶ « Scopus AI | Elsevier », sur www.elsevier.com, s. d. (en ligne : <https://www.elsevier.com/products/scopus/scopus-ai> ; consulté le 4 février 2026)

DES INITIATIVES PUBLIQUES

En parallèle des solutions technologiques privées, un certain nombre d'initiatives publiques tente d'intégrer les technologies d'IA. Ceci témoigne de l'appropriation de ces outils dans une logique de service public et d'ouverture des savoirs. Tout comme dans le cas des solutions privées, il existe un grand nombre de projets et d'initiatives. L'évocation qui suit n'est donc pas exhaustive, mais permet de donner un aperçu de la variété des possibilités d'implémentation des technologies d'IA générative, en identifiant quelques acteurs qui s'emparent de ces projets.

HAL : améliorer le traitement des données

L'archive ouverte HAL intègre le modèle Grobid, un outil d'IA qui extrait des informations structurées depuis les PDF déposés sur la plateforme et les transforme en TEI XML. Cette transformation rend les contenus lisibles par les machines, permettant d'automatiser des traitements de données, d'améliorer la découvrabilité et d'identifier plus précisément les droits associés à une publication¹⁰⁷. Les technologies d'IA générative sont donc mises à profit pour améliorer le fonctionnement de la plateforme et augmenter les potentiels de recherche. La technologie est mise au service des objectifs initiaux de la plateforme, à savoir récupérer les dépôts, les archiver de manière pérenne, diffuser les métadonnées, diffuser et ouvrir la connaissance à tous. Ces avancées ouvrent des perspectives variées, telles que le projet ANR MatOS de traduction automatique ou les pistes d'aide à la modération et de détection de doublons. La technologie est donc mise au service d'une plus grande efficacité d'un service préexistant, sans changement de nature informationnelle.

L'ABES

Raluca Pierrot rappelle que l'ABES emploie l'IA non générative depuis longtemps pour le fonctionnement du Sudoc, notamment pour enrichir les données et établir des liens automatiques entre la description des ouvrages et l'identification fine des auteurs¹⁰⁸. Plus récemment, l'ABES a développé des expérimentations d'usage de l'IA générative, notamment comme aide à l'indexation, présentée lors du Printemps Couperin 2025 par Yann Nicolas¹⁰⁹.

Raluca Pierrot précise les ambitions de l'ABES en termes d'IA : « au sein de l'écosystème des bibliothèques universitaires françaises, nous portons comme projet de développer des outils d'intelligence artificielle pour utiliser l'IA au quotidien dans l'enrichissement des outils, notamment les outils d'informatique documentaire, les outils qui permettent la découverte des contenus par les utilisateurs¹¹⁰ ». Il s'agit donc, on le voit, d'une approche essentiellement technique, ancrée dans les pratiques professionnelles existantes et visant à les améliorer plutôt qu'à les remplacer. Le travail sur l'indexation augmentée à l'IA se

¹⁰⁷N. FARGIER et Y. BARBORINI, « AP2024 | HAL et l'IA », *op. cit.*

¹⁰⁸R. PIERROT, « Entretien », *op. cit.*

¹⁰⁹« Printemps Couperin - Sciencesconf.org », s. d. (en ligne : <https://printempscoup.sciencesconf.org/resource/page/id/1> ; consulté le 11 février 2026)

¹¹⁰R. PIERROT, « Entretien », *op. cit.*

fait en collaboration avec des professionnels des bibliothèques : premiers usagers, ils sont sollicités dans l'évaluation de la pertinence des résultats.

La BnF

La Bibliothèque nationale de France développe également des projets d'IA, en particulier pour le traitement de ses ressources numériques. Les travaux en cours portent notamment sur le développement d'un moteur de recherche pour la plateforme Gallica qui permettrait de faire de la recherche sémantique¹¹¹. Un autre projet vise à construire un outil de repérage d'images qui permettrait de nouvelles modalités de recherche, et donc d'accès aux contenus graphiques déposés sur Gallica¹¹². La BnF collabore également avec le consortium ArGiMi pour mener à bien un projet d'industrialisation, financé par la Banque publique d'investissement. L'ambition est de travailler sur des cas d'usages pour le patrimoine et de spécialiser des modèles de LLM afin de produire des modèles avec des méthodologies, des modèles en open source. La BnF participe en fournissant des jeux de données d'entraînement à partir de ses corpus. Un cas d'usage concret est la correction des OCR produits par la BnF depuis une dizaine d'années. Jean-Philippe Moreux précise l'orientation stratégique de la BnF dans l'intégration des outils d'IA : il s'agit d'utiliser ces technologies « pour mieux remplir [les] missions [de la BnF]. C'est vraiment l'IA technique dans divers secteurs : par exemple la numérisation, le catalogage ou la diffusion des collections avec Gallica en particulier¹¹³ ». La technologie est donc utilisée pour faciliter des processus déjà en place, sans changer les orientations stratégiques.

ILaaS : une infrastructure mutualisée pour l'enseignement supérieur et la recherche

La fédération ILaaS est une infrastructure d'IA mutualisée qui « vise à fournir des services d'intelligence artificielle générative de confiance, robuste, éthique et sobre pour l'enseignement supérieur et la recherche¹¹⁴ ». Elle est, à l'heure actuelle, dans sa première phase, et compte six partenaires : les universités de Reims, de Rennes, de Lille, de Lorraine, Panthéon Sorbonne et Centrale Supélec. L'ambition est, dans un second temps, d'étendre le projet à plus d'une vingtaine de partenaires. Cette fédération vise à construire une infrastructure d'IA mutualisée conçue en prenant en compte les impératifs de confiance, de sobriété numérique, et de résilience. Elle est soutenue par le Ministère de l'Enseignement Supérieur et de la recherche (MESR), avec un suivi conjoint de la Direction générale de l'enseignement supérieur et de l'insertion professionnelle (DGESIP) et la Direction générale de la recherche et de l'innovation (DGRI).

En regroupant les ressources, ILaaS promeut une culture de partage de compétences et d'innovation, tout en répondant à trois enjeux structurants : la maîtrise des coûts par la mutualisation des investissements en infrastructures, la soutenabilité

¹¹¹Mentionné par Arnaud Laborderie au cours d'un entretien.

¹¹²BIBLIOTHÈQUE NATIONALE ET UNIVERSITAIRE (BNU), *Le projet Gallica Images. L'intelligence artificielle pour l'accès aux grands corpus d'images*, 18 juin 2024, 27:23 (en ligne : <https://www.youtube.com/watch?v=IVMfpR96jrk> ; consulté le 14 février 2026)

¹¹³J.-P. MOREUX, « Entretien », *op. cit.*

¹¹⁴ILaaS, « Communiqué de presse », 18 juin 2025, p. 1 (en ligne : <https://www.ilaas.fr/wp-content/uploads/2025/07/20250605.CommuniqueDePresse.pdf> ; consulté le 3 février 2026)

environnementale par l'optimisation de l'utilisation des serveurs, et la souveraineté des données par le renforcement du contrôle sur leur traitement dans un environnement de confiance. Cette initiative répond au besoin de développer des solutions technologiques qui, pour être menées à bien, nécessitent une taille critique. Cela témoigne d'une forme d'autonomie institutionnelle, mais aussi de l'importance de fonctionner en réseau et de mutualiser les coûts, en adéquation avec les moyens nécessaires au développement de ces technologies. À la différence des autres services évoqués, l'outil est nouveau et change la manière dont les informations sont traitées et générées. Paradoxalement, il semble qu'il s'agisse, de tous les services évoqués, celui dans lequel les professionnels des bibliothèques sont le moins impliqués.

À l'exception d'ILaas, les différents exemples mentionnés semblent principalement relever d'une prolongation de missions préexistantes des bibliothèques, archives ouvertes, institutions du patrimoine. Contrairement aux initiatives privées, les discours semblent beaucoup plus précautionneux. Il s'agit de faire des expérimentations, dans lesquelles les chercheurs et usagers sont associés, avec une évaluation transparente des fonctionnements des outils. Cela contraste avec les solutions développées par des initiatives privées, produits présentés comme finis, sans exposition claire de leurs biais ou limitations. Qu'il s'agisse d'initiatives privées ou publiques, l'évaluation des fonctionnements de ces outils reste importante, et les bibliothécaires peuvent intervenir dans cette étape, comme nous allons le voir maintenant.

QUELLES EXPERTISES DES BIBLIOTHÉCAIRES POUR COMPRENDRE CES TECHNOLOGIES ?

On l'a vu, les cas d'usage des IA génératives sont variés. Ils concernent autant le traitement de la documentation mise en ligne que sa découvrabilité, ou même sa rédaction. À la nature polymorphe de ces technologies s'ajoute une évolution technologique, et des développements par différents types d'acteurs ayant des intérêts variables. Se pose dès lors la question des expertises que peuvent mobiliser les professionnels des bibliothèques pour comprendre et évaluer les implications de l'usage de ces technologies dans l'écosystème de l'IST.

Évaluer l'évolution des pratiques de consultation d'information

L'un des rôles des bibliothécaires est de porter un regard critique sur les évolutions des pratiques de consultation de l'information. Or, les modèles d'IA générative qui proposent des synthèses et construisent des réponses à partir de sources changent fondamentalement les voies d'accès à l'information : les usagers humains, lorsqu'ils utilisent des IA génératives, ne consultent plus les sources primaires déposées dans les bibliothèques numériques, plateformes des éditeurs ou archives ouvertes, mais se contentent de consulter les réponses synthétiques proposées par les IA génératives qu'ils utilisent. Dans une telle approche, les sources de la documentation électronique restent importantes, mais principalement

en tant qu'elles alimentent les réponses des systèmes automatisés, et non en tant qu'elles sont consultées par des humains.

Dans une telle hypothèse, il est important de ne pas laisser uniquement des initiatives privées et fermées s'emparer de ces outils. La sauvegarde d'une neutralité dans l'usage des outils d'accès à l'information touche au cœur des métiers des bibliothèques, et à des problématiques déjà rencontrées par ces institutions dans le cadre du passage au numérique. Comme le rappelle Bertrand Calenge dans ses notes, face aux ressources numériques, « la véritable question [...] est au fond celle de l'accès à ces ressources. Elles existent, elles sont théoriquement accessibles, mais elles sont noyées pour l'utilisateur dans un maelström de moteurs de recherche, interfaces propriétaires, etc.¹¹⁵ » La politique documentaire consiste alors à « ménager les portails, points d'accès, réseaux de circulation, qui autoriseront et faciliteront l'appropriation des ressources utiles par l'utilisateur final ». Dans un contexte où les IA génératives ajoutent une couche d'intermédiation supplémentaire, ce rôle de médiation prend une importance renouvelée : les professionnels des bibliothèques sont en mesure de cerner les évolutions des pratiques de consultation des informations, de porter un regard critique sur les conséquences de ces nouvelles technologies, et de proposer des adaptations ou guides de bonnes pratiques face à ces nouvelles technologies. Le regard critique passe également par une évaluation des possibilités et limites de ces outils.

Évaluer méthodiquement la qualité des résultats fournis par l'IA générative

Valentine Favel-Kapoian et Susan Kovacs, dans leur travail consacré à l'accompagnement des usagers vers l'IA par les professionnels des bibliothèques soulignent « l'inaccessibilité épistémologique et cognitive de l'apprentissage automatique¹¹⁶ ». Cette inaccessibilité correspond, au fonctionnement en « boîte noire » des LLM et du *deep learning* : les entraînements statistiques et probabilistes sur des jeux de données massifs ne sont pas intelligibles pour un humain. Il y a une tension forte entre la lacune épistémologique et la puissance apparente de ces outils qui, de prime abord, semblent pouvoir aider les usagers sur un large panel de tâches. Le rôle des bibliothécaires apparaît dès lors important : il s'agit de cerner les enjeux, de formuler cette inaccessibilité, mais également de la mitiger en trouvant des moyens pour composer avec elle, afin de maintenir un esprit critique et des pratiques informationnelles et scientifiques viables. Cela peut, notamment, passer par le fait de ne pas prendre pour argent comptant les déclarations d'entreprises qui ont bien compris que les IA génératives ouvrent un nouveau marché et qui cherchent, avant tout, à vendre des services à des clients potentiels. Par ailleurs, les bibliothécaires doivent développer des habitudes d'évaluation de la qualité des réponses fournies par les outils mobilisant de l'IA générative, en prêtant non seulement attention à la possibilité d'hallucinations, mais aussi aux biais et limitations dans les réponses fournies, à partir d'une analyse des données d'entraînement.

En s'appuyant sur leurs compétences d'évaluation des outils de découverte de ressources, et sur leur compréhension des enjeux pour assurer la bonne circulation de

¹¹⁵B. CALENGE, « Permanence bibliothécaire », sur *Carnet de notes*, 21 décembre 2015 (en ligne : <https://bccn.wordpress.com/> ; consulté le 23 janvier 2026)

¹¹⁶V. FAVEL-KAPOIAN et S. KOVACS, « Accompagner les usagers vers l'IA : défis et motivations des bibliothécaires-“explorateurs” de l'enseignement supérieur », dans *Information et intelligence artificielle*, s. l., De Boeck Supérieur, 2025, p. 235

l'information, les professionnels des bibliothèques peuvent développer des méthodes d'évaluation des solutions technologiques d'IA générative. C'est notamment le travail que fait le bibliothécaire singapourien Aaron Tay, au travers de ses billets de blog « *Musings about librarianship* » Il énonce l'importance des principes de reproductibilité et d'interprétabilité pour évaluer la qualité des résultats fournis par les moteurs de recherche. Dans le cas où ceux-ci fonctionnent sur des technologies d'IA générative, il insiste sur l'importance de comprendre les mécanismes logiques sur lesquels reposent les technologies, et de les doubler de tests pratiques, cruciaux pour évaluer la reproductibilité d'un résultat. Aline Bouchard exprime ainsi :

Beaucoup plus que dans le cas de la recherche par mots-clés, il est somme toute difficile d'évaluer et de contrôler correctement les résultats obtenus. Sans compter que ces outils, dans la droite ligne de ChatGPT et des LLM, fournissent des réponses aléatoires, interrogeant la reproductibilité d'une requête au sein d'un même outil. Elicit reconnaît ainsi une « imprévisibilité inhérente » dans leurs réponses et invite à considérer l'outil comme un assistant humain, donnant des réponses différentes selon le moment...¹¹⁷

Au regard de la rapidité d'évolution de ces technologies et de leur adoption par les usagers, les bibliothécaires doivent développer des méthodes de test rapides. Ce travail est crucial pour guider et accompagner les utilisateurs dans le bon usage d'outils dont ils s'emparent, quoi qu'il arrive. Il est complémentaire de celui des chercheurs qui mènent des analyses plus rigoureuses et plus lentes¹¹⁸. Plusieurs articles du blog de Tay exposent de telles méthodologies de test, et analysent les variations de reproductibilité et d'interprétabilité des résultats en fonction de choix technologiques et des corpus documentaires avec lesquels sont conçues ces technologies. Si les corpus d'entraînement initiaux des LLM sont trop massifs pour être évalués par un œil humain, l'expertise documentaire redevient cruciale dans le cas d'outils fonctionnant sur des technologies de RAG. Celles-ci, en effet, fondent leurs réponses sur des ensembles documentaires plus limités, ce qui augmente la qualité de leurs réponses et introduit la possibilité de citer des sources. Par ailleurs, si puissants soient-ils, ces outils de recherche ne couvrent pas la totalité des publications scientifiques et il est crucial de porter un regard critique et d'informer les usagers sur les corpus consultés par ces outils pour comprendre les biais des réponses qu'ils génèrent. Par exemple Scopus AI ne propose pas de travaux antérieurs à 2003 dans ses recherches¹¹⁹.

Delhay et Perrin tentent de poser les exigences d'une IA au service des savoirs. Ils soulignent que « l'outil n'est rien sans le corpus, et le corpus n'est rien sans un cadre éthique de construction ». Ils identifient une série d'exigences précises : transparence sur les corpus utilisés, explicitation des chaînes de transformation des données, lisibilité cognitive et sémantique des réponses, explicitation des critères de sélection activés par le modèle, et droit de regard sur les modèles sollicités. À la suite de Tay, ils proposent trois critères d'évaluation des outils : le discernement cognitif, l'enracinement éthique et l'utilité située pour

¹¹⁷A. BOUCHARD, « "#WorkInProgress », *op. cit.*, p. 19

¹¹⁸A. TAY, « Testing AI Academic Search Engines - What to find out and how to test (2) », sur *Musings about librarianship*, mai 2025 (en ligne : <https://musingsaboutlibrarianship.blogspot.com/2025/05/testing-ai-academic-search-engines-what.html> ; consulté le 31 juillet 2025) « While academic researchers have rigorous (and time-consuming) methods for evaluating information retrieval (like TREC), librarians need practical approaches. We need to understand these tools well enough to guide users and make informed decisions, without needing weeks of formal testing. »

¹¹⁹A. BOUCHARD, « "#WorkInProgress », *op. cit.*, p. 17

les usagers. Le discernement cognitif exige que l'outil rende lisibles ses mécanismes, permettant de « retracer les corpus sollicités, connaître les critères de réponse activés, et identifier les biais éventuels ». L'enracinement éthique interroge l'origine, la validation et la labellisation des données mobilisées, ainsi que le respect des droits d'auteur et des principes FAIR. L'utilité située pour l'utilisateur suppose que l'outil renforce l'autonomie critique plutôt que de proposer un « prêt-à-penser algorithmique », dans une logique où « l'IA devient un partenaire dialogique, et non une boîte noire autoritaire¹²⁰ ».

La capacité à développer une métaréflexion sur les outils de diffusion de l'information, à construire des protocoles d'évaluation, puis à les transmettre, est particulièrement précieuse dans un environnement informationnel en mutation profonde. Comme le souligne Delhay à propos des compétences importantes pour un « bibliothécaire open accès », « la curiosité des agents¹²¹ » est un atout fondamental. Elle est, également, essentielle face à des technologies dont les potentialités se multiplient alors même que l'écosystème n'est pas stabilisé. Outre les capacités d'analyse et d'évaluation des outils, il semble finalement que les professionnels des bibliothèques ont des expertises techniques qu'il peut être pertinent de mobiliser dans le bon développement de ces technologies.

Être partie prenante des solutions technologiques

Delhay et Perrin identifient parmi les zones de tension autour du développement des technologies d'IA générative une « mutation silencieuse des interfaces » : « les interfaces de consultation et les portails des bibliothèques sont déjà concurrencés par les moteurs de recherche généralistes, qui risquent eux-mêmes d'être de plus en plus invisibilisés par des interfaces d'IA venant d'acteurs de la tech ». Si les bibliothécaires ne participent pas à la conception de ces outils, ils risquent d'être « relégués au rôle de simples accompagnateurs, voire invisibilisés¹²² ». Or, les professionnels des bibliothèques ont des compétences qui peuvent être précieuses pour l'implémentation de technologies d'IA générative dans l'écosystème de la publication scientifique. Leur capacité à formaliser des règles et des normes dans la circulation des savoirs peut aider à la conception d'IA dont les résultats soient traçables, situés et discutables. L'une des forces des bibliothèques réside dans leur maîtrise des données et des métadonnées. Celles-ci sont, en effet, cruciales pour l'entraînement des modèles. Fort de ce constat, Geoffroy invite à une évolution du point de vue des rapports entre bibliothèques et IA générative. Plutôt que de considérer les IA génératives comme des outils à mobiliser, ou un maillon supplémentaire de la chaîne documentaire il invite à « repenser le rôle des bibliothèques comme fournisseuses de données dans l'écosystème de l'IA générative¹²³ ». Dans la même ligne, Delhay et Perrin défendent l'idée que les professionnels des bibliothèques peuvent intervenir, « non pas comme experts en algorithmie, ce qu'[ils] ne sont pas, mais comme “architectes de confiance documentaire”, s'appuyant sur des corpus constitués¹²⁴ ». À la capacité d'identification de corpus de qualité s'ajoute, par ailleurs, une autre expertise précieuse des professionnels des bibliothèques : la maîtrise des éléments historiques et la

¹²⁰M. DELHAYE et S. PERRIN, « Bibliothèque de recherche et IA : vers de nouveaux espaces documentaires ? », 2 juillet 2025, p. 23 (en ligne : <https://bbf.enssib.fr/consulter/bbf-2025-00-0000-013> ; consulté le 17 juillet 2025)

¹²¹M. DELHAYE, « Portrait métier # 1 : bibliothécaire open access », *Bulletin des bibliothèques de France*, n° 13, 1^{er} janvier 2017, p. 20-21

¹²²M. DELHAYE et S. PERRIN, « Bibliothèque de recherche et IA », *op. cit.*

¹²³G. Geoffroy, « Pour une approche décomplexée de l'IA », *Arabesques*, no 107, Agence bibliographique de l'enseignement supérieur, 12 octobre 2022, p. 22

¹²⁴M. DELHAYE et S. PERRIN, « Bibliothèque de recherche et IA », *op. cit.*

compréhension des spécificités documentaires permettant de contextualiser un corpus. Cette expertise s'avère d'autant plus cruciale que les LLM traitent les savoirs de manière neutre, sans capacité d'adapter les analyses aux contextes de production. Les professionnels des bibliothèques détiennent une capacité à situer les savoirs qui fait précisément défaut aux systèmes automatisés. Là où les LLM produisent des réponses décontextualisées, les bibliothécaires sont capables de restituer les contextes, de qualifier les sources et de fournir des réponses contextualisées. Leur expertise peut donc être mobilisée pour améliorer le traitement des données et, à terme, le fonctionnement des IA générative. Lo insiste ainsi sur l'intérêt d'une posture de collaboration à la fois avec les entreprises qui développent les technologies, les éditeurs, et les institutions académiques pour promouvoir des pratiques éthiques dans l'usage de l'IA, en particulier dans le cadre de la publication assistée par l'IA générative¹²⁵.

Par ailleurs, à la capacité d'évaluation et la connaissance des corpus s'ajoute une compréhension fine de l'écosystème de l'information et des équilibres sur lesquelles il repose. De nouveau, les professionnels des bibliothèques peuvent jouer un rôle déterminant, notamment, car ils échangent avec l'ensemble des acteurs, en particulier pour défendre des bonnes pratiques assurant la pérennité et la bonne santé de l'écosystème. Un exemple récent touche aux évolutions récentes de l'accès ouvert, en particulier concernant certaines données nécessaires pour le bon fonctionnement des moteurs de recherche. Ceux-ci dépendent, largement, de métadonnées et d'*abstracts* en accès ouvert. Tay l'exprime par le moyen d'une métaphore : « *If AI-powered discovery is the shiny sports car, open abstracts are the petrol*¹²⁶. » Or, dans le même billet de blog, il note une réduction drastique de la disponibilité en accès ouverts des *abstracts*, mentionnant un passage de 80 % à 25 % d'*abstracts* en accès ouvert pour les articles détenus par Springer et Elsevier à l'automne 2025. De telles évolutions peuvent limiter l'efficacité de la découverte des outils s'appuyant sur l'IA générative, mais aussi poser des problèmes pour les outils de découverte et de recherche plus classiques. Cette publication s'inscrit dans une pratique d'évaluation des conditions de la bonne circulation et découverte de l'information, propre aux compétences des professionnels de l'information. À ce rôle d'analyse, s'ajoute une seconde fonction, celle de la défense de bonnes pratiques envers l'ensemble des acteurs. Les bibliothèques sont ainsi à un point nodal de l'écosystème et doivent continuer à défendre l'accès ouvert comme un fondement crucial de la bonne santé de la circulation de l'IST, malgré les recompositions entraînées par les IA génératives.

L'interaction est donc double, et les bibliothèques peuvent trouver leur place dans le développement des technologies d'IA générative. Fournisseurs de données d'entraînement et observateur des modalités de circulation des connaissances, les professionnels des bibliothèques sont aussi évaluateurs des technologies. Cela implique une capacité à concilier exigences documentaires et potentialités technologiques. Les bibliothécaires, forts de leur tradition de médiation et de

¹²⁵L. S. Lo, « Generative AI and Open Access Publishing », *op. cit.*, p. 174 « By collaborating with technology companies, academic institutions, and publishers, libraries can drive innovation and promote ethical practices in AI-driven publishing. »

¹²⁶« Si la découverte augmentée à l'IA est la belle voiture de sport, les *abstracts* en accès ouvert en sont le carburant » A. Tay, « The Petrol Tank for AI Discovery Might be Running Dry as Publishers close access to scholarly content such as *abstracts* due to AI incentives », sur Aaron Tay's Musings about Librarianship, 27 septembre 2025 (en ligne : <https://aarontay.substack.com/p/the-petrol-tank-for-ai-discovery> ; consulté le 8 février 2026)

défense de l'accès ouvert, disposent des compétences nécessaires pour jouer un rôle de garant de la qualité des savoirs et des informations diffusées, en promouvant un regard critique sur des outils qui, souvent, se diffusent en se targuant d'un progrès en termes de résultat, mais tendent parfois à gommer, voire à minimiser leurs biais et leurs limites. Ils peuvent aussi jouer un rôle de conseil important dans un écosystème en recomposition, et porter des messages et des recommandations auprès des décideurs publics et des instances de normalisation qui seront évoquées dans la partie suivante.

ENCADREMENTS LÉGAUX, NORMATIFS ET PRATIQUES

Au-delà des questions techniques traitées précédemment, l'IA générative soulève des interrogations de natures juridiques, déontologiques et épistémologiques. Comme le souligne Morgan Blangeois, « l'émergence des systèmes d'IA pose de nouveaux défis à notre compréhension traditionnelle de la paternité intellectuelle et de la création de savoir¹²⁷ ». Or, la rapidité des bouleversements entraînés par ces technologies entraîne un décalage considérable entre les possibilités offertes par ces outils et les cadres normatifs censés aider les chercheurs à déterminer quelles lignes de conduite ils doivent respecter dans le cadre de leurs pratiques. Resnik et Hosseini insistent sur l'urgence de réduire ce décalage, au vu des questions que posent les technologies d'IA générative en ce qui concerne des fondements de la recherche, parmi lesquels « objectivité, reproductibilité, transparence, responsabilité, et confiance dans la science¹²⁸. » Ce décalage se manifeste dans le cadre même de rédaction du présent travail. Au moment où ces lignes sont écrites, l'Enssib n'a pas adopté de charte ni de cadre édictant des règles sur les usages considérés comme acceptables ou non de l'IA générative dans les travaux des élèves conservateurs. Je dois donc choisir seule que faire par rapport à mon usage de ces technologies. En adéquation avec les principes d'intégrité scientifique, ce travail comporte une déclaration d'usage qui détaille les outils mobilisés. Ce choix est motivé par un souci de transparence et d'honnêteté intellectuelle, et par un alignement sur les lignes de conduite et les recommandations étudiées dans la présente partie. Pour cerner la manière dont ces questionnements sont traités aujourd'hui, cette partie propose une analyse du cadre normatif en construction. Elle explicite d'abord les différentes instances qui édictent les lois, les normes et les règles encadrant ces pratiques, qu'elles soient préexistantes à l'arrivée de l'IA générative ou élaborées depuis. Elle montre ensuite que des questions de fond restent à traiter face aux mutations en cours.

UN CADRE NORMATIF EN COURS DE CONSTITUTION

La loi

Les questions de droit que posent les IA génératives se situent, principalement, à deux niveaux : le respect de la propriété intellectuelle, et le respect des données personnelles. L'un des points de tension majeurs dans le développement des technologies d'IA générative réside dans le fait qu'elles sont entraînées sur des données moissonnées à grande échelle sur Internet et sur les questions de respect des droits d'auteurs et des licences dans l'utilisation de ces contenus. Or, le moissonnage et le traitement de données ne sont pas le propre des

¹²⁷M. BERGADAÀ et P. PEIXOTO, *Réinventer l'intégrité académique à l'ère de l'intelligence artificielle*, Caen, France, Éditions EMS, Management & société, 2025, p. 44

¹²⁸D. B. RESNIK et M. HOSSEINI, « The ethics of using artificial intelligence in scientific research: new guidance needed for a new tool », *AI and Ethics*, vol. 5, n° 2, 1^{er} avril 2025, p. 1500. « [T] he need for guidance is urgent because using AI raises novel epistemological and ethical issues related to objectivity, reproducibility, transparency, accountability, responsibility, and trust in science. » Je traduis

entreprises privées développant de l'IA générative. Certains domaines de recherche scientifique traitent de tels jeux de données dans différents contextes d'expérimentation depuis plus d'une décennie. Ainsi, des dispositions encadrant les pratiques de moissonnage existaient avant 2022 et l'arrivée des IA générative grand public, développées par des entreprises privées. Les premières dispositions légales visaient à protéger ces pratiques en permettant aux chercheurs de pratiquer de la fouille de texte et de données. La directive européenne 2019/790 sur le droit d'auteur dans le marché unique numérique (DAMUN), transposée en droit français par l'ordonnance du 24 novembre 2021, établit ainsi une exception au droit d'auteur au bénéfice de la fouille de textes et de données dans l'article 3 :

Des copies ou reproductions numériques d'œuvres auxquelles il a été accédé de manière licite peuvent être réalisées sans autorisation des auteurs en vue de fouilles de textes et de données menées à bien aux seules fins de la recherche scientifique par les organismes de recherche, les bibliothèques accessibles au public, les musées, les services d'archives ou les institutions dépositaires du patrimoine¹²⁹.

La seule condition pour mener de la fouille de textes et de données pour les organismes de recherche et les institutions du patrimoine culturel est donc que l'accès aux œuvres soit licite, et les ayants-droits ne peuvent faire valoir leurs droits d'auteur pour empêcher cette fouille. Cette disposition est, usuellement, nommée l'exception TDM (Text data mining) au droit d'auteur. L'article 4, quant à lui, étend cette autorisation à toute personne, y compris les entreprises, toujours à condition que l'accès aux contenus soit fait de manière licite. Dans ce second cas, toutefois, l'ayant droit, c'est-à-dire le détenteur des droits d'auteurs peut refuser cette utilisation par le moyen d'une opposition explicite (« *opt-out* ») devant être signalée par des moyens lisibles par machine. La conséquence pratique de ce dispositif est considérable : sauf si l'éditeur ou l'auteur exprime un refus explicite, les publications scientifiques sont moissonnables par des systèmes d'IA par défaut.

Rappelons que cette disposition légale a été prise avant l'arrivée des technologies d'IA génératives grand public. Adopté à l'échelle européenne et entré en vigueur le 1^{er} août 2024, le Règlement sur l'intelligence artificielle (RIA), parfois aussi appelé IA Act, est venu préciser le cadre normatif en l'adaptant aux nouvelles technologies. L'article 53 stipule que les fournisseurs de systèmes d'IA ont pour obligation de « prendre des mesures visant à respecter le droit d'auteur et les droits voisins » et y ajoute une exigence de transparence, puisqu'ils doivent « élaborer et rendre publiquement disponible un résumé suffisamment détaillé des données ayant servi à l'entraînement de leur modèle¹³⁰ ». Ce résumé doit être assez précis pour permettre d'identifier l'utilisation potentielle d'une œuvre ou d'un contenu protégé, mais n'a pas à détailler la manière dont ce contenu a été utilisé, cette dimension étant protégée par le secret des affaires. Pour reprendre la métaphore employée dans le rapport de mission relative à la mise en œuvre du Règlement européen, les entreprises sont tenues de partager les ingrédients, mais pas la recette¹³¹. Le RIA pose un cadre essentiel pour la normalisation des pratiques d'IA générative et leur mise en adéquation avec le droit d'auteur et la protection des

¹²⁹ « Ordonnance n° 2021-1518 du 24 novembre 2021 complétant la transposition de la directive 2019/790 du Parlement européen et du Conseil du 17 avril 2019 sur le droit d'auteur et les droits voisins dans le marché unique numérique et modifiant les directives 96/9/CE et 2001/29/CE - Légifrance », s. d.

¹³⁰ « Règlement (UE) 2024/1689 du Parlement européen et du Conseil du 13 juin 2024 établissant des règles harmonisées concernant l'intelligence artificielle et modifiant les règlements (CE) n° 300/2008, (UE) n° 167/2013, (UE) n° 168/2013, (UE) 2018/858, (UE) 2018/1139 et (UE) 2019/2144 et les directives 2014/90/UE, (UE) 2016/797 et (UE) 2020/1828 (règlement sur l'intelligence artificielle) (Texte présentant de l'intérêt pour l'EEE) », 2024

¹³¹ A. BENSAMOUN et L. FEIRREIRA, *Rapport de mission relative à la mise en œuvre du règlement européen établissant des règles harmonisées sur l'intelligence artificielle*, Conseil supérieur de la propriété littéraire et artistique, 2024

données. Néanmoins, il ne rebat pas les cartes pour ce qui est de l'IA et de la recherche. Par ailleurs, si le résumé des contenus utilisés permet de poser un cadre de création d'un marché respectueux des droits d'auteur et des interactions commerciales, il ne traite pas d'une rétribution équitable pour les ressources en accès ouvert. Or, cette question, nous l'avons vu, est essentielle pour la viabilité de l'écosystème de la connaissance. Le RIA ne tranche pas non plus les questions de signalement de l'usage de l'IA pour la recherche scientifique. Sur ce dernier point, la référence est le Code de conduite européen pour l'intégrité en recherche. Il énonce quatre principes fondamentaux de l'intégrité scientifique : la fiabilité de la recherche, l'honnêteté dans le travail, le respect des collègues et collaborateurs, et la responsabilité à l'égard de la société en général. Depuis juin 2023, il a été mis à jour et comporte une précision sur l'utilisation de l'IA, en esquisant le cadre dans lequel elle peut être utilisée pour la recherche. Il inscrit parmi les exemples de pratiques inacceptables, contrevenant aux bonnes pratiques de recherche, le fait de « masquer l'utilisation d'outils d'intelligence artificielle ou d'outils automatisés dans la création de contenu ou la rédaction de publications ». Par ailleurs, il fixe des grandes lignes concernant le signalement de l'utilisation d'IA générative « les chercheurs font connaître leurs résultats et leurs méthodes, y compris l'utilisation de services externes, d'outils d'intelligence artificielle ou automatisés, sous une forme qui respecte les normes faisant consensus de la discipline, et qui facilite, le cas échéant, la vérification et la reproductibilité¹³² ». Ces recommandations restent, on le voit, assez larges et les « normes faisant consensus [d'une] discipline », sont, aujourd'hui, encore loin d'être stabilisées.

Préconisations d'acteurs variés sans coordination complète

Une multiplicité d'acteurs, à différentes échelles, s'efforcent de produire des normes, des préconisations et des guides de bonnes pratiques pour orienter la recherche. Cette effervescence témoigne de l'ampleur des recompositions en cours, mais aussi du désordre dans lequel elles s'effectuent. Comme le note la Commission européenne dans ses *Living Guidelines on the responsible use of generative AI in research*, le paysage informationnel est aujourd'hui complexe, du fait d'une prolifération de normes et de guides, ce qui crée une difficulté à s'orienter et choisir quelles règles choisir¹³³.

Les éditeurs scientifiques sont parmi les acteurs ayant réagi le plus rapidement. Leur nature d'acteurs privés, leurs moyens, et les enjeux financiers liés à leurs activités constituent autant d'incitations fortes à s'adapter et à clarifier les usages. Les éditeurs occupent, de surcroît, une position singulière : ils sont à la fois utilisateurs d'outils d'IA et détenteurs de droits portant sur des contenus pouvant servir pour l'entraînement des modèles. Les rencontres François Gèze de mars 2025, organisées par le Syndicat national de l'édition sur le thème « Édition scientifique et intelligence artificielle », attestent de la dynamique de la profession pour s'emparer de la question. La présentation de ces rencontres insiste sur la place centrale des éditeurs scientifiques dans l'écosystème : ils sont « à un poste

¹³²ALLEA, « Code de conduite européen pour l'intégrité en recherche », 2023, p. 7

¹³³EUROPEAN COMMISSION, « Living guidelines on the responsible use of generative AI in Research », 2025, p. 4

d'observation mais aussi moteur dans le cadre des développements de l'IA¹³⁴ ». Selina La Barbera, directrice de publication chez EDP Sciences, témoigne de cette prise de conscience dans un entretien et évoque la ligne adoptée par cette maison d'édition : « nous sommes en train d'établir, comme tous les autres éditeurs, des guides et des lignes directrices en ligne à destination de toutes les différentes parties prenantes concernant les revues scientifiques¹³⁵. » Aujourd'hui, la plupart des grands éditeurs ont adopté des lignes claires concernant l'utilisation et le signalement de l'IA générative dans les travaux qu'ils publient¹³⁶. Au-delà des guides, les éditeurs investissent également dans des activités annexes : publication de cas d'usages, organisation de masterclass et de séminaires destinés à inciter aux bonnes pratiques. La réactivité des éditeurs privés contraste avec la temporalité plus lente d'autres acteurs, en particulier ceux fonctionnant sur des fonds publics, et disposant d'une moindre force d'action. Un hébergeur comme OpenEdition, par exemple, est encore en cours de réflexion et n'a pas, à l'heure actuelle, adopté de ligne précisant qu'elle acceptait ou non en termes d'utilisation de l'IA. Ce décalage soulève une remarque : la réactivité des acteurs privés fait d'eux des prescripteurs, de fait, des normes d'utilisation et de signalement de l'IA. Néanmoins, leur légitimité en tant qu'instances de normalisation des bonnes pratiques scientifiques peut être mise en question, dans la mesure où ils ont des intérêts, notamment commerciaux, dans le sujet.

Plusieurs instances de référence, plus neutres de par leur nature institutionnelle, se sont saisies de la question. L'UNESCO et la Commission européenne ont ainsi publié des « guides », respectivement en 2023 et 2025, sur la question de l'utilisation de l'IA dans la recherche. Ces institutions adoptent une position de conseil sans normalisation contraignante¹³⁷. Le but des publications est d'aider les décideurs publics, à différentes échelles, à adopter des règles et des lois pour l'usage de l'IA générative dans l'éducation et la recherche. Pour ce qui est de la France, ces règles peuvent relever de la responsabilité de l'Ofis, département chargé de mettre en œuvre les missions en matière d'intégrité scientifique confiées à l'Hcéres. Elle a organisé en 2025 un colloque intitulé « Comment l'IA générative transforme les pratiques de recherche : nouveaux enjeux d'intégrité scientifique¹³⁸ ». De manière révélatrice, parmi les espaces thématiques du site web de l'Ofis, une section « Intégrité scientifique et intelligence artificielle » apparaît avec la mention « ouverture prochaine ». La question est identifiée, mais les recompositions sont à l'œuvre sans qu'une position claire soit identifiée.

L'IRAFPA, Institut de recherche et d'action sur la fraude et le plagiat académiques, association indépendante fondée en 2016 à Genève, a publié un communiqué sur la transgression des normes académiques. Il inclut « l'usage de l'IA sans y faire référence¹³⁹ » parmi les atteintes à l'éthique formelle et aux codes de conduite en matière de plagiat. La question des modes de signalement de l'usage de l'IA et des différents types d'utilisation de l'IA n'est néanmoins pas tranchée : l'implication n'est pas la même selon qu'il s'agit d'une détection de fautes d'orthographe, de la

¹³⁴ « Rencontres François Gèze sur l'édition scientifique et l'intelligence artificielle », sur *Syndicat national de l'édition*, s. d. (en ligne : https://www.sne.fr/evenement_sne/recontres-francois-geze-sur-ledition-scientifique-et-lintelligence-artificielle/ ; consulté le 12 février 2026)

¹³⁵ UNIVERSITÉ DE LILLE, « Charte des bonnes pratiques d'utilisation des intelligences artificielles génératives dans les travaux pédagogiques », s. d. ; S. LA BARBERA et N. CLÉMENT, « Entretien », *op. cit.*

¹³⁶ « AI Policies in Academic Publishing 2025 », *op. cit.* ; M. PERKINS et J. ROE, « Academic publisher guidelines on AI usage: A ChatGPT supported thematic analysis », *F1000Research*, vol. 12, 16 janvier 2024, p. 1398

¹³⁷ *Guidance for generative AI in education and research*, s. l., UNESCO, 2023 ; EUROPEAN COMMISSION, « Living guidelines on the responsible use of generative AI in Research », *op. cit.*

¹³⁸ « Colloque 2025 - Comment l'IA générative transforme les pratiques de recherche : nouveaux enjeux d'intégrité scientifique », sur *OFIS*, s. d. (en ligne : <https://www.ofis-france.fr/colloque/colloque-2025/> ; consulté le 6 juin 2025)

¹³⁹ « Les repères de l'intégrité académique – Responsable Academia », s. d. (en ligne : <https://irafpa.org/a-propos/regles-dintegrite-academique/> ; consulté le 11 septembre 2025)

rédaction d'un résumé, de la construction d'un plan ou de l'élaboration d'une hypothèse déléguée à une IA générative.

La Commission européenne souligne dans ses *Living Guidelines* l'importance d'une mobilisation collective des différents acteurs de la recherche, publique et privée, pour développer des habitudes d'utilisation de l'IA générative d'une manière appropriée et efficace¹⁴⁰. La normalisation et la création de règles ne sont pas suffisantes : ce qui importe, c'est aussi le développement d'une conversation sur ces outils et d'un consensus établi entre l'ensemble des acteurs du monde de la recherche. Des mécanismes de coordination préexistants peuvent aider à l'existence d'une telle conversation. Nathalie Clément, d'EDP Sciences, témoigne ainsi d'interconnexions entre le secteur de l'édition et les instances publiques : « La directrice générale d'EDP Sciences fait partie de différentes commissions organisées auprès du ministère de l'Enseignement supérieur et de la Recherche, notamment sur les publications en *open access* et les publications scientifiques. [...] Par ailleurs, comme EDP Sciences a signé un accord transformant avec Couperin et le ministère de l'Enseignement supérieur depuis 2017, nous avons des obligations : un acte d'engagement avec l'ABES qui nous demande de rendre compte de toutes ces informations¹⁴¹. » Ces coordinations et concertations peuvent, néanmoins, prendre du temps et la normalisation n'est, aujourd'hui, pas stabilisée. Les différentes chartes universitaires sur l'utilisation de l'IA en attestent. Un autre élément participe à la difficulté de la stabilisation : le fait que les technologies sont elles-mêmes très mouvantes, et que les innovations continuent de se multiplier. Tous les documents qui visent à poser des lignes directrices insistent par ailleurs sur un même point : les préconisations devront faire l'objet de mises à jour et d'adaptation en fonction des évolutions futures des possibilités et usages technologiques.

Cas concrets : les chartes universitaires

Les universités se dotent progressivement de chartes pour encadrer les usages de l'IA, à la fois dans les pratiques pédagogiques et dans les processus de recherche. Ces chartes ont des orientations et des périmètres variables. Nous en étudions quelques-unes ci-dessous pour mettre en évidence les grandes tendances, à la fois dans le fond et dans la forme de ces documents.

Les lauréats de l'AMI DemoES ont rédigé un document modèle pouvant être réutilisé par différents établissements. La proposition de charte se distingue par son ambition méthodologique : elle comprend une charte, qui précise les obligations et interdictions minimales dans l'utilisation de l'IA à la fois en recherche et en pédagogie, mais aussi un guide de bonnes pratiques, un guide réflexif sur le déploiement et une bibliographie. Son approche détaille « chaque étape du processus, de la préparation au déploiement, en passant par l'évaluation continue » de la charte et vise à « instaurer une véritable symbiose entre humain et technologie¹⁴² ». Il s'agit donc autant d'un modèle que d'un document cadre pour

¹⁴⁰EUROPEAN COMMISSION, « Living guidelines on the responsible use of generative AI in Research », *op. cit.*, p. 4

¹⁴¹S. LA BARBERA et N. CLÉMENT, « Entretien », *op. cit.*

¹⁴²AMI DEMOES, « IA générative & ESR : propositions de charte & guides », 2025

accompagner le développement et la mise à jour, par les établissements, de leur propre charte. Par ailleurs, notons que le document, dans sa dernière version, datant du 27 septembre 2025, est présenté comme un « projet » en cours de construction.

L'examen comparé de quelques chartes révèle des orientations et des périmètres variés. Les chartes des universités de Lille, d'Angers et de l'ISM-IAE de Versailles Saint-Quentin-en-Yvelines portent exclusivement sur les travaux pédagogiques et n'encadrent pas la recherche. Ce choix répond peut-être à une urgence dans le domaine de l'enseignement, rythmé par le calendrier des évaluations, à une plus grande latitude de décision des universités en matière pédagogique : elles sont garantes des pratiques d'enseignement et d'apprentissage, mais ne sont qu'un maillon dans l'écosystème de la recherche.

Certains établissements ont adopté des chartes qui règlent l'ensemble des usages de l'IA générative, pour le personnel, les chercheurs, et les étudiants. La charte de l'Université d'Orléans, adoptée le 18 octobre 2024, propose un encadrement complet. Elle expose le positionnement de l'université par rapport à ces technologies, présente un ensemble de consignes générales puis les décline de manière spécifique dans le cadre pédagogique, le cadre des activités de recherche et celui des activités administratives et techniques. Sa ligne directrice est claire : accepter l'usage de ces technologies pour l'assistance aux activités humaines, mais refuser que l'IA générative s'y substitue, selon trois principes : transparence, respect de la confidentialité et éthique¹⁴³. Le positionnement de cette charte est « volontariste mais prudent¹⁴⁴ ». La charte de l'Université de Toulouse offre une particularité notable : elle contient une déclaration d'usage des outils d'IA générative mobilisés lors de sa propre rédaction. C'est, de tous les documents consultés, la seule qui s'applique à elle-même les principes qu'elle énonce. Elle comprend un volet pour les étudiants, un pour les chercheurs et un pour les personnels administratifs. Pour les chercheurs, elle invite à « explore[r] les possibilités offertes par l'IA générative pour améliorer la recherche » tout en demandant de « veille[r] à la transparence dans l'utilisation de l'IA générative dans [les] publications¹⁴⁵ ».

La prolifération de chartes, aux périmètres et ambitions variables atteste du bouleversement que représentent les IA génératives pour le monde universitaire. Elles témoignent d'une situation actuelle qui se caractérise par une absence de réponse unifiée. Dans l'ensemble, une ligne directrice semble émerger : les usages sont permis, mais doivent être encadrés, et soumis à des principes préexistants pour maintenir la transparence et la responsabilité humaine dans la production des savoirs. Les exigences dans les déclarations d'usage sont variables et encore peu balisées. Notons, par ailleurs, que la majorité des établissements de l'ESR n'a pas encore adopté de charte ou de document de normalisation sur ces questions.

La place des bibliothèques dans ces réflexions

Les entretiens menés dans le cadre de ce mémoire ont fait apparaître un constat : si de nombreuses universités sont en train de réfléchir à l'adoption de chartes pour encadrer les usages de l'IA, les professionnels des bibliothèques ne sont pas systématiquement sollicités pour participer à ces réflexions. Les processus de rédaction

¹⁴³UNIVERSITÉ D'ORLÉANS, « Charte sur l'utilisation de l'Intelligence Artificielle », 2024, p.2

¹⁴⁴AMI DEMOES, « IA générative & ESR : propositions de charte & guides », *op. cit.*

¹⁴⁵UNIVERSITÉ DE TOULOUSE, « Charte du bon usage des IA génératives à l'UT », 2023, p.4

des chartes varient considérablement d'un établissement à l'autre, et l'implication des bibliothécaires l'est tout autant. Le SCD de l'Université d'Orléans semble avoir été identifié comme un interlocuteur essentiel sur la question de l'IA. À Lille, le SCD a également été sollicité, mais cela est surtout le fruit de la proactivité d'une professionnelle. À Rennes, les bibliothécaires n'ont pas été impliqués. Le facteur humain semble déterminant : la profession n'est pas, automatiquement, identifiée comme pouvant jouer un rôle dans les processus de rédaction de ces chartes au sein des bibliothèques.

Le rapport de l'IGESR « IA et Enseignement supérieur : formation, structuration et appropriation par la société » est à ce titre révélateur. Le rapport ne mentionne que très brièvement les bibliothèques, alors même que plusieurs membres de l'ADBU ont été sollicités dans le cadre de la mission. Cette absence est d'autant plus remarquable que l'une de ses recommandations est de « créer des tiers lieux, permettant la formation, l'appropriation et l'usage de l'IA, sur tout le territoire, en associant les acteurs locaux et assurer la formation des médiateurs¹⁴⁶ ». On peut légitimement se demander pourquoi les bibliothèques, qui exercent déjà une fonction de médiation et constituent des tiers lieux établis sur l'ensemble du territoire, ne sont pas identifiées comme des acteurs naturels de cette mission.

Ce déficit de reconnaissance contraste avec l'expertise démontrée par certains opérateurs du monde des bibliothèques. Couperin a travaillé rapidement à comprendre les implications légales des clauses de restriction d'usage que les éditeurs ont tenté d'introduire suite à l'arrivée des IA générative grands publics. Un groupe de travail dédié a pu s'appuyer sur l'expertise d'une agence comme l'ABES, dont la force réside dans une vision complète des implications des restrictions de l'usage de l'IA, nourrie par le développement interne de projets d'IA¹⁴⁷. Ce travail, mobilisant des compétences à la fois juridiques et techniques, et une compréhension des pratiques de la recherche, témoigne d'une force d'analyse et de réactivité dans la profession¹⁴⁸.

L'ADBU, pour sa part, a publié un texte intitulé « Ouvrir la science à l'heure de l'intelligence artificielle : pourquoi et comment ? » en janvier 2026¹⁴⁹. Le texte porte un regard critique sur les évolutions de l'écosystème de la science ouverte et défend l'importance de l'accès ouvert face aux inquiétudes soulevées par l'IA générative. Si ce document relève davantage de la posture de principe que d'un cadre strictement opérationnel, il témoigne du rôle de l'association nationale dans le positionnement de la profession et d'une compréhension stratégique des enjeux de ces technologies concernant l'écosystème informationnel.

Zhou et Soullière identifient trois éléments clés pour garantir la bonne adoption de l'IA dans l'écosystème de la publication scientifique : la confiance, la collaboration et la gouvernance. L'IA doit être transparente et explicable pour gagner la confiance des acteurs ; la collaboration doit être repensée comme une

¹⁴⁶F. Pascal et al., « IA et Enseignement Supérieur : Formation, Structuration et Appropriation par la Société », p.5

¹⁴⁷R. PIERROT, « Entretien », *op. cit.*

¹⁴⁸J. BAUDRY, « Enjeux juridiques de l'IA à l'Université : comment la loi encadre les usages scientifiques et documentaires », 15 octobre 2025 (en ligne : https://www.amue.fr/fileadmin/amue/RDV_Amue/2025/J_BAUDRY_Presentation_IA_-_AMUE_Aspects_juridiques_v2.pdf ; consulté le 22 janvier 2026)

¹⁴⁹ADBU, « Ouvrir la science à l'heure de l'intelligence artificielle : pourquoi et comment ? », *op. cit.*

interaction humain-IA où l'humain reste au centre de la décision ; et des politiques solides doivent garantir que les systèmes d'IA font ce pour quoi ils sont conçus, sans excéder leur rôle¹⁵⁰. Il y a donc un équilibre à trouver afin de préserver l'écosystème de la publication scientifique tout en maintenant les potentiels bénéfiques des technologies. Comme cela vient d'être montré, un ensemble d'acteurs s'empare de cette question et propose des encadrements, mais les recompositions sont en cours. Si des grandes tendances commencent à émerger, la normalisation n'est pas encore complètement arrêtée. Les technologies étant encore récentes et l'analyse de leur impact sur les pratiques de l'information scientifique et technique n'étant pas achevée, des questions de fond persistent malgré ces premiers encadrements.

QUESTIONS DE FOND : LA RESPONSABILITÉ SCIENTIFIQUE DANS L'USAGE DES IA GÉNÉRATIVE

Réfléchir aux conséquences épistémologiques de l'utilisation d'une technique et adapter les méthodologies et les pratiques constitue l'un des fondements de la recherche scientifique ; à ce titre, la communauté scientifique est outillée pour réfléchir aux conséquences de l'utilisation de l'IA générative pour faire de la recherche des technologies et adapter ses exigences afin de garantir la bonne marche de la science. La multiplication de publications étudiant l'adoption de ces outils et les conséquences de leur utilisation témoigne d'une méta-réflexion à l'œuvre. Si le paysage normatif n'est pas encore complètement stabilisé, de grandes tendances et de nouvelles questions émergent autour de l'usage de l'IA générative.

Tendances concernant le signalement

La capacité de l'IA générative à générer du texte a créé des interrogations sur l'auctorialité, et soulevé des questionnements quant au fait de signaler ces outils comme des auteurs ou co-auteurs¹⁵¹. Aujourd'hui un consensus autour de cette question semble avoir émergé : les systèmes d'IA ne peuvent être considérés comme des auteurs ou des co-auteurs de travaux scientifiques. Même en cas d'utilisation d'IA générative, les auteurs humains engagent leur entière responsabilité vis-à-vis des contenus publiés. Le Committee on Publication Ethics (COPE) a pris position dès février 2023 en soulignant que les outils d'IA ne peuvent engager leur responsabilité¹⁵². La Commission européenne reprend cette position sans ambiguïté : pour pouvoir prétendre à une auctorialité, une entité doit pouvoir déployer une puissance d'action et une responsabilité, deux conditions que ne remplissent pas les IA génératives. Cela place définitivement l'auctorialité du côté des humains¹⁵³.

¹⁵⁰H. ZHOU et M. SOULLIÈRE, « From Detection to Disclosure — Key Takeaways on AI Ethics from COPE's Forum », sur *The Scholarly Kitchen*, 25 août 2025 (en ligne : <https://scholarlykitchen.sspnet.org/2025/08/25/from-detection-to-disclosure-key-takeaways-on-ai-ethics-from-cope-forum/> ; consulté le 2 octobre 2025)

¹⁵¹C. STOKEL-WALKER, « ChatGPT listed as author on research papers: many scientists disapprove », *Nature*, vol. 613, n° 7945, 18 janvier 2023, p. 620-621

¹⁵²« Authorship and AI tools », sur *COPE: Committee on Publication Ethics*, 13 février 2023 (en ligne : <https://publicationethics.org/guidance/cope-position/authorship-and-ai-tools> ; consulté le 2 octobre 2025) « [T]hey cannot take responsibility for the submitted work. As non-legal entities, they cannot assert the presence or absence of conflicts of interest nor manage copyright and license agreements. »

¹⁵³EUROPEAN COMMISSION, « Living guidelines on the responsible use of generative AI in Research », *op. cit.* « AI systems are neither authors nor co-authors. Authorship implies agency and responsibility, so it lies with human researchers. » p.7

L'absence d'auctoralité ne dispense pas de signaler l'usage de ces technologies dans le cadre du travail de recherche. Les guides de conduite et chartes convergent, dans l'ensemble, vers l'idée d'une nécessaire transparence dans l'usage de l'IA générative. L'outil peut être utilisé, mais il faut préciser de quelle manière. La question de la forme que prend ce signalement, cependant, reste encore difficile à trancher. Ainsi, la charte de l'Université d'Orléans préconise de « préciser par un mécanisme de citations classiques » tout texte généré par une IA générative, tout en admettant que « cette obligation peut être levée dans certains types d'usage : simple traduction, amélioration du niveau de langue sans apport de changements sur le fond¹⁵⁴ ». Derrière l'apparente clarté de cette ligne de conduite, un flou persiste : quels sont les mécanismes de citation classiques pour l'usage des IA ? Par ailleurs, la frontière entre les usages avec ou sans apport sur le fond peuvent, parfois, être difficiles à définir.

Les organismes fédéraux de recherche au Canada (CRSH, CRSNG, IRSC) ont choisi une autre approche, dans leurs nouvelles directives adoptées en 2025. Ils ne demandent pas aux chercheurs de divulguer s'ils ont utilisé l'IA pour rédiger leurs demandes de subventions mais exigent en revanche que les chercheurs assument l'entière responsabilité du contenu soumis et des sources utilisées. Comme le commente Stéphane Vial : « La nuance est essentielle : il ne s'agit pas de traquer les outils, mais de rappeler que l'auteur est celui qui répond publiquement du texte. La machine n'efface pas l'esprit humain : c'est le choix affirmé de l'esprit qui fait disparaître la machine derrière lui.¹⁵⁵ » L'utilisation de l'IA ne pose donc pas de problème en soi, mais elle ne peut décharger la responsabilité humaine.

C'est bien, finalement, cette expertise humaine qu'il semble crucial de maintenir au cœur des processus de la recherche. Blangeois souligne que « les LLM nous obligent à dissocier l'originalité apparente d'un texte (sur le plan de la forme, du style, de la combinaison inédite d'éléments) de l'originalité réelle de la contribution scientifique (sur le plan des idées, des concepts, des résultats nouveaux)¹⁵⁶ ». Cela remet l'accent sur des éléments formels dans la méthode scientifique, parmi lesquels la rigueur méthodologique et la transparence, garantes de la validité des résultats scientifiques. Un texte généré par un LLM peut présenter une originalité formelle sans pour autant apporter une contribution scientifique authentique¹⁵⁷. Les outils d'IA générative présentent par ailleurs un risque de plagiat déguisé : ils réutilisent des idées provenant de sources existantes sans les citer, ce qui peut aboutir à la réappropriation, par un individu, d'idées issues de travaux antérieurs, sans que cela soit perceptible par l'utilisateur humain. Ces écueils remettent en évidence la nécessité cruciale de l'expertise humaine dans l'évaluation critique de la valeur des productions scientifiques. La charte du DemoES formule cette exigence en termes opérationnels : « valider et vérifier soi-même les contenus générés par IA générative après avoir procédé à une relecture critique et attentive », et « après relecture et validation, assumer la pleine

¹⁵⁴UNIVERSITÉ D'ORLÉANS, « Charte sur l'utilisation de l'Intelligence Artificielle », *op. cit.*, p. 3

¹⁵⁵S. VIAL, « La mort de l'auteur n'aura pas lieu à l'ère de l'IA », *Le Devoir*, 25 août 2025 (en ligne : <https://www.ledevoir.com/opinion/idees/911919/mort-auteur-aura-pas-lieu-ere-ia> ; consulté le 31 août 2025)

¹⁵⁶M. BERGADAÀ et P. PEIXOTO, *Réinventer l'intégrité académique à l'ère de l'intelligence artificielle*, *op. cit.*, p. 44

¹⁵⁷*Ibid.*, p. 33

responsabilité des contenus générés par IA générative même s'ils comportent des erreurs¹⁵⁸ ».

Si les préconisations sur le signalement et la transparence progressent, de vastes zones d'incertitude subsistent dès lors que l'on sort de l'aide à la rédaction pour considérer les usages de l'IA générative dans le processus de recherche lui-même, pour générer des hypothèses ou traiter des données. Resnik et Hosseini observent que les codes d'éthique de l'IA existants, tels que ceux de l'UNESCO ou du Bureau de la politique scientifique et technologique américain restent encore imprécis sur ces questions¹⁵⁹. Cette difficulté est renforcée par la rapidité d'évolution des technologies, qui rend tout cadre normatif susceptible d'obsolescence. Andersen et ses collègues suggèrent d'adopter des principes suffisamment flexibles pour être appliqués dans différents domaines de recherche, plutôt que des règles strictes vouées à être dépassées par les avancées technologiques¹⁶⁰.

À la difficulté de poser des règles strictes dans un contexte technologique mouvant s'ajoute la nécessité d'adapter les pratiques aux champs disciplinaires. En effet, en dépit de l'existence de grands principes de l'intégrité scientifique, les attendus méthodologies varient largement en fonction des disciplines. Chaque communauté disciplinaire ayant ses pratiques propres, l'usage de l'IA générative doit y être adapté. Une étude menée par Andersen et ses collègues sur l'utilisation de l'IA par des chercheurs révèle des différences marquées selon les disciplines. L'échantillon est circonscrit aux chercheurs danois, mais l'on peut supposer que les tendances qu'ils dégagent se retrouvent, au moins partiellement, dans l'ensemble du monde de la recherche. D'après eux, les chercheurs en sciences techniques, médicales et ceux ayant des approches quantitatives et expérimentales utilisent tendanciellement l'IA générative pour accélérer leurs recherches en automatisant des processus, quand les chercheurs en humanités, sciences sociales qualitatives et sciences théoriques l'utilisent principalement en tant qu'assistant pour des éléments langagiers¹⁶¹. La versatilité de l'outil se traduit ainsi par des usages différents en fonction des méthodes de production de savoir. Certains domaines soulèvent des enjeux particuliers : le droit et la médecine, en raison de la sensibilité de leurs objets ; les mathématiques et l'informatique, en raison de leur proximité technologique avec les outils eux-mêmes. Ces éléments complexifient encore les processus de normalisation des usages et des signalements. À ces différences disciplinaires s'ajoutent des différences générationnelles : l'étude de Andersen et al. montre ainsi que les jeunes chercheurs ont tendance à plus s'approprier les technologies d'IA générative, mais qu'il n'y a pas de différence en fonction des genres¹⁶².

Questions latentes

Les préconisations de signalement et de transparence peuvent permettre l'adoption d'une transparence nécessaire à la bonne marche de la science. Néanmoins, parmi les impensés du cadre normatif actuel figure le rapport entre l'IA générative et les systèmes d'évaluation de la recherche. Les préconisations de signalement et de transparence ne suffisent pas si le système d'évaluation lui-même pousse à la surproduction. Or, nous

¹⁵⁸AMI DEMOES, « IA générative & ESR : propositions de charte & guides », *op. cit.*, p. 3

¹⁵⁹D. B. RESNIK et M. HOSSEINI, « The ethics of using artificial intelligence in scientific research », *op. cit.*, p. 1500

¹⁶⁰J. P. ANDERSEN *et al.*, « Generative Artificial Intelligence (GenAI) in the research process – A survey of researchers' practices and perceptions », *op. cit.*, p. 10

¹⁶¹J. P. ANDERSEN *et al.*, « Generative Artificial Intelligence (GenAI) in the research process – A survey of researchers' practices and perceptions », *op. cit.*

¹⁶²*Ibid.*, p. 1

l'avons vu, c'est une tendance qui préexiste à l'arrivée de l'IA générative et que cette technologie risque d'amplifier.

La concurrence internationale entre chercheurs et l'inflation des publications scientifiques favorisent le maintien, voire le renforcement, du poids des indicateurs quantitatifs dans l'évaluation de la recherche. En France, des critères bibliométriques comme l'Impact Factor et le H-index sont utilisés par le CNU ou l'Hcéres pour évaluer à la fois les carrières individuelles et les productions des laboratoires. Or, l'arrivée d'outils facilitant la production de contenu scientifique remet en question la pertinence de ces évaluations quantitatives, déjà critiquées précédemment¹⁶³. Si produire devient plus aisé, les métriques fondées sur le volume perdent leur valeur discriminante. Au contraire, elles augmentent les incitations à des mésusages des IA générative pour produire des publications en quantité, sans apport scientifique de qualité¹⁶⁴. Ces évolutions sont lentes, car les pratiques scientifiques sont bien ancrées. Larousserie en 2025 montre ainsi que la communauté scientifique elle-même porte une part de responsabilité dans la perpétuation de ces modèles de publication et d'évaluation¹⁶⁵. Le journaliste rapporte ainsi des propos de chercheurs comme Denis Bourguet : « La réalité est que les collègues gardent leur réflexe de métrique et de marques... » On observe néanmoins, en France, des initiatives introduites pour réformer le système d'évaluation. La plus notable est celle engagée par le CNRS. L'institution majeure du monde de la recherche s'est désabonnée des bases de données bibliométriques commerciales, coupant d'abord l'accès à Scopus puis, en décembre 2025, au Web of Science et a annoncé une nouvelle prise en compte, dans ses évaluations, d'éléments qualitatifs, parmi lesquels l'ouverture des données de la recherche¹⁶⁶.

L'IA générative rend potentiellement obsolètes des mesures quantitatives strictes qui démontraient déjà leurs limites. Son arrivée pourrait ainsi accélérer une transformation des modèles d'évaluation que les appels à la réforme n'avaient pas suffi à enclencher. Dans ce contexte, le bibliothécaire peut trouver sa place en exerçant une fonction de médiation, en accompagnant les chercheurs dans la compréhension critique de ces enjeux et en proposant des alternatives aux modèles dominants. La construction d'un cadre normatif cohérent, capable de concilier innovation technologique, intégrité scientifique et accès ouvert au savoir, est en cours. Il s'agit de l'un des enjeux importants des années à venir pour le monde de la recherche. Les bibliothèques, en raison de leur expertise peuvent trouver leur place dans ce chantier important.

¹⁶³« San Francisco Declaration on Research Assessment », sur *DORA*, 16 décembre 2012 (en ligne : <https://sfedora.org/read/> ; consulté le 16 février 2026)

¹⁶⁴O. FRAISIER-VANNIER, « Mésusages des LLM dans la publication scientifique », *op. cit.*

¹⁶⁵D. LAROUSSE, « Le monde des revues scientifiques au bord de l'asphyxie », *op. cit.*

¹⁶⁶C.-I. PASTEUR, « Bases de données commerciales : après s'être désabonné de Scopus, le CNRS coupe l'accès au Web of Science », sur *Open science : évolutions, enjeux et pratiques*, 18 décembre 2025 (en ligne : <https://openscience.pasteur.fr/2025/12/18/bases-de-donnees-commerciales-apres-setre-desabonne-de-scopus-le-cnrs-coupe-lacces-au-web-of-science/> ; consulté le 14 janvier 2026)

QUELLES ÉVOLUTIONS POUR LES PROFESSIONNELS DES BIBLIOTHÈQUES ?

Keller note que la puissance des IA générative semble, à première vue, remettre en cause les institutions traditionnellement dépositaires du savoir, notamment les bibliothèques, les archives et les musées. Cela tient à la triple action des IA générative qui interviennent comme médiateurs dans l'accès au savoir, simulent une autorité dans ce rapport au savoir et génèrent du sens. Néanmoins Keller insiste : l'autorité que les IA génératives projettent relève d'un simulacre : derrière l'assertivité de leurs réponses il n'y a, en fait, aucun contrôle de la véracité des propos¹⁶⁷. Or, une observation fine de la situation met en évidence que, plus qu'un risque de disparition du métier, c'est un repositionnement de la profession qui s'impose. Comme le formule Adeline Regé : « Pour moi, ce n'est pas une révolution, ce sont des évolutions classiques qui révèlent en fait la pertinence des compétences qu'on a déjà et qu'on peut appliquer à ces nouvelles technologies. Je ne suis pas de ceux qui disent que c'est la mort du bibliothécaire, au contraire. Comme Internet n'était pas la mort du bibliothécaire non plus¹⁶⁸. » Il importe que les professionnels de l'information maintiennent leur expertise et préservent leur place face à ces technologies. Tout l'enjeu réside dans le fait de déterminer quelles sont les voies nécessaires pour assurer cette place, quelles sont celles qui doivent être renforcées. Cette partie a pour objectif de répondre à ces questions en faisant un état des lieux, tout d'abord, des compétences préexistantes de la profession qui peuvent être valorisées et remobilisées, quelles sont les adaptations enclenchées par rapport à ces technologies, et comment les évolutions se traduisent sur les missions des professionnels des bibliothèques.

DES COMPÉTENCES PERTINENTES, UNE LÉGITIMITÉ À CONSTRUIRE

Un workshop mené par MIT University Press en 2024 sur l'accès à la science et à la communication scientifique pointe l'importance d'infrastructures solides pour assurer la bonne diffusion du savoir et la nécessité d'une coopération entre les différents acteurs, publics et privés¹⁶⁹. Or, il est notable que les bibliothèques universitaires sont une institution dans une position privilégiée pour défendre la coopération et faire le lien entre les différents acteurs de l'écosystème. Les professionnels des bibliothèques sont, en effet, à un point nodal et interagissent avec des acteurs clés : ils négocient avec les éditeurs, mais également accompagnent et forment les chercheurs, et peuvent porter des préconisations envers les institutions publiques de normalisation pour maintenir la bonne santé de l'écosystème de la publication scientifique. Par ailleurs, les missions mêmes de bibliothèques, gardiennes de l'information et d'un savoir vérifié, peuvent contribuer à alimenter un climat de confiance et de coopération. Il importe néanmoins que la légitimité des bibliothèques sur ces questions soit reconnue.

¹⁶⁷P. KELLER, « Beyond AI and copyright. Funding a sustainable information ecosystem », *op. cit.*, p. 2

¹⁶⁸A. REGÉ, « Entretien », *op. cit.*

¹⁶⁹*Access to Science & Scholarship 2024 - Summary Report*, MIT Press Workshop, 2024

Regard critique et capacité de médiation

En tant que professionnels de l'information, les bibliothécaires occupent une position privilégiée pour observer les recompositions de l'écosystème de la publication et de la circulation des savoirs scientifiques. Ils sont, ainsi, en mesure de porter un regard critique sur les dynamiques qui entourent cette situation, sans avoir d'intérêts privés dans ces dynamiques. Comme le note Lo, les bibliothèques ont une position unique pour cultiver la maîtrise et la compréhension des outils d'IA générative auprès des chercheurs, des étudiants, mais également des autres acteurs du monde académique¹⁷⁰. Leur compréhension des questions de qualité de l'information est ainsi particulièrement précieuse dans un moment d'évolution rapide et de bouleversement entraîné par de nouvelles technologies.

Par ailleurs, par rapport aux autres acteurs, les bibliothèques bénéficient d'un avantage structurel : leur position de neutralité dans le traitement des savoirs. Nogueira et ses collègues le soulignent : dans un paysage caractérisé par les cycles de « hype » où l'IA peut être utilisée comme argument commercial, les bibliothèques ont une position unique pour mesurer, effectivement, l'efficacité et l'utilité de ces outils, sans être influencée par les effets modes ou des biais financiers¹⁷¹. Les mêmes auteurs appellent ainsi les bibliothèques à adopter un « scepticisme sobre » et le principe de précaution face aux innovations, c'est-à-dire à résister à la pression d'adopter des outils au seul motif qu'ils sont techniquement réalisables ou que d'autres les utilisent déjà¹⁷². Il ne s'agit ni de verser dans un « techno-solutionnisme » béat, ni de rejeter en bloc les IA génératives, mais de les évaluer de manière rationnelle et objective, en mobilisant les compétences d'analyse de l'information propres à la profession. Nous avons vu, dans la troisième partie, des exemples attestant du fait que les professionnels de l'information disposent de méthodes et d'outils d'évaluation qui peuvent, et doivent, être déployés pour évaluer le fonctionnement des IA génératives.

Nogueira et ses collègues soulignent que les professionnels des bibliothèques, ces dernières années, ont développé un rôle de médiation auprès des usagers pour les aider à s'approprier les technologies permettant de découvrir, d'explorer ou de consulter de l'information¹⁷³. Cette évolution fait des bibliothécaires des intermédiaires entre les usagers et l'information : ils forment et conseillent l'ensemble de la communauté universitaire. Graaf décrit la profession comme exerçant une « fonction d'intermédiaires » au regard des questions d'accès ouvert¹⁷⁴. Les compétences qu'il identifie sont directement transférables aux problématiques soulevées par l'IA : vue d'ensemble de l'environnement réglementaire et politique, de données FAIR et d'intégrité scientifique ; expertise technique en matière de publication numérique ; et soutien à la production scientifique sous toutes ses formes. Face aux mutations entraînées par l'IA générative, ce positionnement déjà établi, peut être remobilisé, sous réserve d'une montée en compétence pour les fonctionnements et enjeux documentaires propres à ces technologies notamment autour des questions concernant la science ouverte.

¹⁷⁰L. S. Lo, « Generative AI and Open Access Publishing », *op. cit.*, p. 172

¹⁷¹L. A. NOGUEIRA *et al.*, « Cutting through the noise: Assessing tools that employ artificial intelligence », *IFLA Journal*, vol. 51, n° 3, SAGE Publications Ltd, 1^{er} octobre 2025, p. 745

¹⁷²*Ibid.*, p. 737

¹⁷³L. A. NOGUEIRA *et al.*, « Cutting through the noise », *op. cit.*

¹⁷⁴M. VAN DER GRAAF, *Open science services by research libraries. Organisational perspectives*, LIBER et ADBU, 2023, p. 6

Néanmoins, si le positionnement et les compétences métier sont présents, il est nécessaire que l'institution bénéficie d'une reconnaissance et d'une légitimité reconnue par l'ensemble des acteurs pour pouvoir s'emparer de cette question. Or, celle-ci semble, aujourd'hui, fragile.

Une reconnaissance et une identité professionnelle fragiles sur ces questions

Les bibliothécaires doivent trouver une place et une légitimité autour de ces questions. Comment l'observe Madeleine Geroudet, « Il faut concevoir la diffusion de la science comme étant une chaîne en train d'évoluer dans ses modalités, dans son fonctionnement. [...] [L]a manière dont on va affirmer notre identité et la nature de notre expertise va jouer un rôle sur la manière dont on sera repositionnés ou non. » Or, il semble qu'aujourd'hui, la profession souffre d'un déficit de reconnaissance sur les questions liées à l'IA générative. Valentine Favel-Kapoian et Susan Kovacs notent la persistance d'interrogations sur la légitimité des bibliothécaires à se positionner sur ces questions, qu'elles considèrent « symptomatique de l'ancrage identitaire fragile » d'une profession « sans cesse réinterrogé[e] par le développement de nouvelles fonctions et de nouvelles missions¹⁷⁵ ».

Cette fragilité, déjà observée à propos des processus de normalisation, semble liée pour partie à un ethos professionnel, celui d'une profession « modeste », pour reprendre les analyses de Cécile Rabot. Cet ethos transparaît, notamment, dans les postures individuelles des professionnels par rapport à ces technologies, qui sont exposés plus tard dans cette partie. Il a été assez transparent, lorsque j'ai sollicité des professionnels pour mener mon enquête : la majorité des personnes que j'ai contactées ont commencé par me répondre qu'elles n'étaient pas spécialistes et être aujourd'hui principalement dans une posture d'interrogation vis-à-vis des outils d'IA générative. Cela est d'autant plus surprenant que, comme nous l'avons vu, ces professionnels possèdent bien des pratiques et des compétences leur permettant d'utiliser et d'évaluer ces outils qui ont un impact très concret sur l'écosystème informationnel. En entretien, une conservatrice évoque à propos des professionnels des bibliothèques « notre crainte de ne pas être légitimes si on s'aventure sur des terrains qui sont un peu plus mouvants que ceux sur lesquels on s'est investis jusque-là¹⁷⁶ ». Elle relate les réactions de collègues lorsqu'elle évoque les questions d'intégrité scientifique et les enjeux relatifs à la validation des contenus scientifiques : « on ne peut pas jouer ce rôle-là, on n'a pas de rôle dans l'évaluation de la qualité scientifique d'un contenu ». Ce à quoi elle répond que c'est « à la fois vrai et en même temps pas tout à fait exact, puisqu'on fait de la validation de contenu, en travaillant sur des paramètres qui ne sont pas le contenu en tant que tel, mais l'environnement et le contexte de production et de diffusion ». Ainsi, il semble que les professionnels des bibliothèques ont toute leur place dans l'accompagnement à l'utilisation, mais également au développement d'IA génératives de manière maîtrisée, éthique, sobre, intègre, et dans le bon respect à la fois des droits d'auteurs et de l'accès ouvert au savoir. La récente mise à jour du référentiel métier RECIF de l'ADBU témoigne ainsi d'une prise de conscience institutionnelle : une nouvelle compétence y a été inscrite, « Utiliser l'IA de façon éthique et responsable¹⁷⁷ ». Si cette inclusion souligne l'importance accordée à ces technologies dans l'évolution des métiers des

¹⁷⁵V. FAVEL-KAPOIAN et S. KOVACS, « Accompagner les usagers vers l'IA », *op. cit.*, p. 235

¹⁷⁶Entretien anonymisé.

¹⁷⁷ADBU, « Référentiel de compétences informationnelles », 2024

bibliothèques, la formulation reste assez indéfinie. Les notions de responsabilité et d'éthique, mobilisées sans être précisées, reflètent l'état actuel des réflexions de la profession : la nécessité d'agir est identifiée, mais les contours concrets de cette action demeurent encore à dessiner.

Paradoxalement, si la profession hésite sur sa propre légitimité, des acteurs extérieurs l'ont déjà identifiée. Certaines entreprises d'IA semblent ainsi considérer les bibliothécaires comme des interlocuteurs, pour lesquels ils façonnent une communication spécifique. On trouve ainsi, sur le site de Scispace, une section intitulée « Ressources for librarians », qui contient, notamment une page intitulée « The Modern Librarian's Role in supporting AI-Driven research ». Le texte insiste sur le rôle des professionnels des bibliothèques pour accompagner les chercheurs et les étudiants dans la maîtrise de ces nouveaux outils : « Aujourd'hui, les bibliothécaires universitaires ne sont pas seulement des gardiens de l'information, mais aussi des conservateurs de confiance, des évaluateurs critiques et des guides qui aident les enseignants et les étudiants à naviguer dans le nouveau monde de la recherche axée sur l'IA¹⁷⁸. » Cette publication est à resituer dans son contexte : il s'agit d'une communication d'une entreprise qui vend une technologie. Cela explique, notamment, un biais sous-jacent du propos concernant le fait que le monde de la recherche se réorganise vers l'IA, et que cet outil doit être adopté. C'est, justement, le rôle des professionnels des bibliothèques que d'interroger de tels présupposés, de questionner les discours qui présentent ces évolutions comme des évidences, et de défendre des pratiques qui garantissent la circulation et la qualité du savoir.

UNE PROFESSION QUI S'ADAPTE POUR RESTER PERTINENTE

Une variété de positionnements face à la technologie

Les entretiens menés dans le cadre de ce mémoire ont été l'occasion d'interroger un panel de professionnels des bibliothèques sur leur rapport à l'IA générative et sur l'utilisation de ces technologies. Si l'échantillon reste restreint, il permet néanmoins de dégager quelques grandes tendances et de formuler de premiers constats. Le premier de ces constats est la diversité des positionnements individuels face à ces outils. Cependant, toutes les personnes interrogées témoignent d'un regard critique sur les possibilités et les limitations des IA génératives, avec une variété allant de la méfiance assumée à l'utilisation précautionneuse.

Certains professionnels déclarent ne pas utiliser les IA génératives. Parmi ceux-ci, deux cas distincts émergent. Certaines personnes s'en abstiennent pour des raisons éthiques et environnementales, considérant que le coût environnemental de ces outils est trop élevé et méconnu, et que les questions quant au respect du droit d'auteur et à la production de la connaissance sont encore insuffisamment résolues aujourd'hui. Sonja Bottger l'exprime clairement : « Je ne les utilise pas personnellement et c'est un choix conscient, pour des raisons éthiques et

¹⁷⁸ « The Modern Librarian's Role in Supporting AI-Driven Research », sur *SciSpace Resources*, 4 avril 2025 (en ligne : <https://scispace.com/resources/university-librarians/the-modern-librarians-role-in-supporting-ai-driven-research/> ; consulté le 10 février 2026)

environnementales. [...] Je vois ça un peu comme un piège pour l'esprit critique¹⁷⁹. » D'autres professionnels déclarent ne pas utiliser les IA génératives car ils n'en ont pas trouvé d'usage particulier dans leur exercice professionnel — c'est d'ailleurs la réponse la plus fréquemment donnée par les personnes interrogées. Émilie Liard reconnaît ainsi n'avoir « pas trouvé d'usage pour [s]on travail qui apporterait une plus-value¹⁸⁰ ». Cela peut être lié, d'après elle, soit à une limitation des outils eux-mêmes soit à un déficit de compétences en matière de prompting ou à un manque d'imagination quant aux usages possibles. Bernard Jacquemin analyse plus largement cette posture comme liée à un manque d'acculturation : le « manque de connaissances et de cultures dans l'univers des bibliothèques relatifs aux outils d'intelligence artificielle agit à la fois comme un repoussoir et comme une boîte noire¹⁸¹ ».

On trouve cependant parmi les personnes interrogées un certain nombre de professionnels déclarant avoir testé des outils pour en explorer les possibilités et évaluer ce qu'ils permettaient de générer. Cette approche exploratoire, guidée par la curiosité, est fréquemment mentionnée lors des entretiens, même si elle n'a pas toujours débouché sur des adoptions des outils dans les pratiques. Elle est parfois motivée par des interactions avec des usagers ou des membres plus jeunes de la famille qui utilisent ces outils de manière enthousiaste. Elle témoigne d'une curiosité prudente et d'un souci des professionnels de comprendre les outils qui se diffusent parmi leurs usagers.

Parmi les usages déclarés au cours des entretiens, le principal est la traduction. L'IA générative est en effet considérée comme très utile pour améliorer des textes rédigés en anglais ou pour faciliter la compréhension de certains documents qui ne sont pas dans la langue natale de l'utilisateur. Certaines personnes interrogées mentionnent par ailleurs l'utiliser pour produire des synthèses de documents, et évoquent parfois une aide à la rédaction ou à la reformulation, bien que celle-ci reste encadrée, et jugée assez peu efficace. Isabelle Filliâtre s'en sert ainsi « pour des reformulations ou pour [s]e donner des idées de plan », en précisant : « je m'en inspire, je ne m'y conforme pas¹⁸² ». Enfin, un dernier usage concerne l'aide à la programmation : Damien Belvèze indique consacrer « 90 % du temps » de son utilisation à « corriger du code source » et « 10 % pour améliorer [s]on anglais¹⁸³ ».

On note par ailleurs que parmi les répondants, beaucoup déclarent privilégier soit les modèles français, à l'instar de Mistral, soit les modèles locaux, lorsque ceux-ci sont mis à disposition par des établissements universitaires – c'est notamment le cas de Damien Belvèze, dont l'université de Rennes donne accès à un outil développé en interne. Ce petit échantillon témoigne donc d'une variété d'usages et d'une absence d'unité dans la manière dont la profession s'empare de cette question.

Comment se former ?

Si les pratiques individuelles varient, il semble important que la profession s'adapte à ces technologies, dans la mesure où elles sont de plus en plus largement adoptées par les usagers – chercheurs et étudiants – et où elles influencent l'ensemble des maillons de l'écosystème de la publication scientifique. La question de la prise en main de ces outils, néanmoins, n'est pas aisée.

¹⁷⁹S. BÖTTGER, « Entretien », 2 octobre 2025

¹⁸⁰E. LIARD, « Entretien », 20 octobre 2025

¹⁸¹B. JACQUEMIN, « Entretien », *op. cit.*

¹⁸²I. FILLIÂTRE, « Entretien », *op. cit.*

¹⁸³D. BELVÈZE, « Entretien », 20 2025

Michel Fraysse insiste : « Je pense que nous devons absolument nous tenir au courant de ces questions, évoluer et être capables de répondre à des demandes¹⁸⁴. » Les bibliothèques ne peuvent ignorer ces mutations au nom de grands principes. Pour autant, la question des voies à emprunter pour se former et pour prendre en main ces outils nouveaux n'est pas simple, dans la mesure où il s'agit d'un terrain neuf, qui touche à des fonctionnements technologiques qui dépassent parfois les compétences des bibliothécaires, et qui évoluent très rapidement. Bernard Jacquemin identifie ainsi comme compétence métier clé à développer « la connaissance, la compréhension de la manière dont fonctionne l'intelligence artificielle¹⁸⁵ ». Raluca Pierrot, dans une posture similaire, précise néanmoins les limites de cette compréhension : les professionnels des bibliothèques doivent être en mesure de comprendre les enjeux informatiques, sans nécessairement développer une expertise technique sur la question : « Je pense qu'il vaut mieux avoir un bibliothécaire qui comprend les enjeux, qui a une culture informatique, qui ne va pas mettre les mains dans le cambouis, mais qui sache discuter avec des informaticiens.¹⁸⁶ » Sébastien Perrin partage cette opinion : « il va falloir monter en compétence technique sur ces outils sans devenir technicien nous-mêmes¹⁸⁷. »

Se pose alors la question des voies à emprunter pour assurer la montée en compétence des bibliothécaires eux-mêmes. Une conservatrice, en entretien identifie deux sujets de légitimité accessibles : « les mentions d'usage et la recherche d'information avec IA », mais reconnaît que « il ne suffit pas de dire qu'on est expert, il faut encore avoir actualisé nos compétences¹⁸⁸ ». Valentine Favel-Kapoian et Susan Kovacs décrivent un travail d'acculturation par autoformation et veille¹⁸⁹. François Magnan, qui forme lui-même à ces technologies, décrit un processus de prise en main collectif, par expérimentation et tests avec les outils : « On se forme par l'expérience : en testant, en observant ce qui fonctionne ou non. Quand l'IA générative est arrivée en 2022, c'est comme ça qu'on a procédé, en s'inspirant aussi les uns des autres¹⁹⁰. » Il identifie, par ailleurs, les bibliothécaires formateurs comme une ressource précieuse pour la formation.

La profession semble aujourd'hui dans une phase d'exploration et de tâtonnement : les initiatives d'acculturation institutionnelle sur ces questions existent mais prennent des formes variées. Les nombreuses journées professionnelles organisées sur le thème de l'IA témoignent d'une conscience collective de la nécessité de réfléchir à l'impact de ces technologies sur les missions des bibliothèques. De telles journées existent à des échelles nationales comme internationales : la conférence annuelle de LIBER en 2025, intitulée *Engage, Collaborate, Innovate: Libraries Working to Address Global Challenges*, traitait en partie de la question de l'intelligence artificielle ; la conférence d'EBLIDA la même année portait sur *Power and Resilience: Books and Reading in the Age of AI* ; l'une des trois journées du Printemps Couperin 2025 était consacrée à la négociation à l'heure du *Big Data*. Les associations professionnelles et consortium, on le voit au travers de ces exemples, s'emparent de la question.

¹⁸⁴M. FRAYSSE, « Entretien », 3 octobre 2025

¹⁸⁵B. JACQUEMIN, « Entretien », *op. cit.*

¹⁸⁶R. PIERROT, « Entretien », *op. cit.*

¹⁸⁷S. PERRIN, « Entretien », *op. cit.*

¹⁸⁸Entretien anonymisé

¹⁸⁹V. FAVEL-KAPOIAN et S. KOVACS, « Accompagner les usagers vers l'IA », *op. cit.*, p. 231

¹⁹⁰F. MAGNAN, « Entretien », 26 juin 2025

À l'échelle des établissements, les initiatives prennent des formes variées. Plusieurs modèles émergent, qui correspondent à des degrés différents de formalisation et d'ambition. Le premier est celui de la cellule IA, un groupe de travail se concentrant spécifiquement sur les questions liées à la profession et aux missions de l'institution. Lucie Albaret décrit le dispositif de l'Université Grenoble Alpes : « Elle s'est montée avec des représentants de chaque direction ou service de la BU. Il y a par exemple un aspect formation interne : on va tous passer une formation obligatoire sur l'IA et les bibliothèques. On sait bien que ça doit être pris en compte dans la formation qu'on donne aux étudiants, donc on commence par nous former. [...] Il y a des sous-groupes de travail à l'intérieur de la cellule, et les sous-groupes rendent compte au sein de la cellule qui se réunit tous les trimestres puis à la direction.¹⁹¹ »

Le deuxième modèle est celui du café IA, à la fois temps d'échange et moment de formation. La BnF en a organisé un premier en décembre 2025. Jean-Philippe Moreux décrit ces cafés comme « des actions de vulgarisation pour parler de manière très concrète de ce qui est en train de se passer, pour que tout le monde soit au courant de ce qu'on fait autour de l'IA et faire remonter les besoins aussi¹⁹² ». Des modules de formation sont par ailleurs intégrés au catalogue de formation de l'établissement.

D'autres établissements privilégient des formats plus souples, comme les formations ad hoc et webinaires. Renaud Delemontez Sage évoque des webinaires organisés par l'université Dauphine pour les personnels administratifs, animés par un doctorant en informatique¹⁹³. Dans un environnement technologique très mouvant, la question de la stabilité de ces formations se pose néanmoins : Isabelle Filiâtre note à propos du séminaire de Dauphine que « si je le regardais aujourd'hui, il y aurait pas mal de choses qui seraient obsolètes¹⁹⁴ ». Le maintien d'une veille, en complément de ces formations, est ainsi essentiel.

Le dernier modèle identifié est celui du cercle d'échange, plus ouvert et plus souple. À Lille, où la question de l'IA a été inscrite dans la stratégie 2030 du SCD, Madeleine Geroudet a lancé un cercle réunissant une trentaine de personnes. « Le principe, c'est que c'est plus large qu'un groupe de travail : on n'a pas fermé la participation. [...] clairement, ça a un impact sur tous les services et donc il faut être le plus ouvert possible. » Le cercle a commencé par des séances de formation suivies d'ateliers : le premier portait sur les règles d'usage internes, le second sur l'IA dans les relations aux usagers. Pour son lancement, un formateur extérieur a formé le comité de direction pendant une demi-journée, suivi d'un atelier de définition des enjeux¹⁹⁵. Ce format illustre une approche qui articule montée en compétence sur les questions technologiques et réflexion collective sur le positionnement du service.

Ces différentes modalités témoignent de mécanismes de réflexion pour faire émerger une culture professionnelle commune face à ces questions. Cette culture est d'autant plus importante que ces technologies ont un impact sur plusieurs des missions fondamentales des bibliothèques.

¹⁹¹L. ALBARET, « Entretien », 4 septembre 2025

¹⁹²J.-P. MOREUX, « Entretien », *op. cit.*

¹⁹³R. DELEMONTÉZ SAGE, « Entretien », 4 octobre 2025

¹⁹⁴I. FILLIÂTRE, « Entretien », *op. cit.*

¹⁹⁵M. GERODET, « Entretien », 22 août 2025

LES MISSIONS DES PROFESSIONNELS DES BIBLIOTHÈQUES FACE À L'IA GÉNÉRATIVE

Pour finir ce travail, nous exposons ci-dessous quelques missions des professionnels des bibliothèques qui semblent particulièrement touchées par ces nouvelles technologies. Le panorama suivant n'est pas complètement stabilisé, l'évolution de certaines missions étant dépendant des progrès technologiques et de l'adaptation du cadre normatif.

Interactions avec les éditeurs

Les interactions avec les éditeurs scientifiques ont été parmi les premières touchées par la question des IA génératives, et ce à deux titres : tout d'abord, certains éditeurs ont tenté d'introduire des clauses restreignant l'usage des IA génératives pour les usagers des bibliothèques, notamment les chercheurs. Par ailleurs, les éditeurs développent progressivement des nouveaux services qui intègrent les technologies d'IA génératives, ce qui modifie leur offre. Face à ces doubles évolutions, les professionnels des bibliothèques doivent monter en compétence sur les questions juridiques, et sur l'évaluation de ces outils.

Les clauses de restrictions d'usages ont été, dans plusieurs entretiens, mentionnées comme la première prise de conscience que les bibliothèques étaient touchées par les évolutions technologiques. Elles ont nécessité une adaptation rapide, qui s'est faite par la mobilisation d'un ensemble d'acteurs collaborant, grâce notamment à la position clé du consortium Couperin sur ces questions. Adeline Regé décrit le processus : « L'objectif était de définir une doctrine commune et une sorte de clause type avec les arguments à présenter aux éditeurs pour distinguer intelligence artificielle, intelligence artificielle générative, et TDM, les usages de recherche, les usages de formation. » Le travail, mené conjointement par des professionnels travaillant en BU et à l'ABES, a consisté à collecter des cas d'usage, des expériences étrangères, et à monter en compétence « sur le plan juridique et technologique¹⁹⁶ ».

En juin 2024, une clause type provisoire a été fournie, pour aider les négociateurs Couperin faisant face à cette question. Elle stipule que « l'usage de l'intelligence artificielle [...] n'est autorisé qu'à des fins exclusives de recherche », que cette utilisation « doit s'effectuer localement ou dans une application hébergée sécurisée », et qu'il est interdit de « donner accès à une société commerciale » pour les contenus¹⁹⁷. Cette réponse rapide de la profession témoigne d'une capacité à monter en compétence et à s'adapter face à l'émergence de nouvelles questions. Cette force devra, probablement, être maintenue à l'avenir. En effet, s'il existe aujourd'hui un cadre réglementaire autour de ces technologies, il est possible que des évolutions futures recomposent les obligations. Les professionnels des bibliothèques doivent donc mener une veille et s'adapter aux évolutions du cadre légal.

Le second type d'interaction avec les éditeurs concerne l'évaluation des services proposés, ceux-ci intégrant de plus en plus, d'options mobilisant des LLM

¹⁹⁶A. REGÉ, « Entretien », *op. cit.*

¹⁹⁷J. BAUDRY, « Enjeux juridiques de l'IA à l'Université : comment la loi encadre les usages scientifiques et documentaires », *op. cit.* La clause figure en annexe.

ou des IA génératives. Aujourd'hui, la profession porte un regard circonspect sur ces nouvelles offres. Sonja Bottger déclare : « Tous les éditeurs s'y mettent et l'utilisent comme stratégie de marketing¹⁹⁸. » Adeline Regé souligne que les éditeurs utilisent l'intégration de l'IA comme argument tarifaire, invoquant les investissements nécessaires, notamment, pour garantir l'intégrité de la publication¹⁹⁹. Or, la question budgétaire est, dans les contextes actuels, très forte. Nombre d'interlocuteurs, lors des entretiens, ont évoqué le décalage entre l'augmentation des coûts et des enveloppes budgétaires fixes pour la documentation électronique. Dans un tel contexte, l'évaluation de la plus-value effective des outils est indispensable. Le groupe de travail IA (GTIA) du consortium Couperin semble prendre ces questions à-bras-le-corps. En effet, dans sa feuille de route 2025-2026, le GTIA pose comme objectif l'évaluation des outils d'IA pour des usages de recherche, qu'il s'agisse d'outils génériques ou d'outils spécifiquement développés à destination de la recherche, afin d'évaluer leur fonctionnement. Le GTIA se donne également comme objectif de réfléchir aux modèles de financements²⁰⁰. L'un des risques liés au développement de modèles payants, identifié par Isabelle Filliâtre, est le développement d'« une recherche à deux vitesses, parce qu'il y a des universités qui auront plus de budget pour acheter ces options IA²⁰¹ ». Le GTIA de Couperin semble prendre la mesure de ces enjeux stratégiques et de la manière dont ils doivent informer les prises de décision et les négociations des professionnels des bibliothèques.

Déterminer le périmètre de la politique documentaire

Outre les négociations avec les éditeurs concernant les droits et les coûts des offres, les outils d'IA générative questionnent, par ailleurs, le périmètre de la politique documentaire pour les outils numériques. Comme le rappelle Bertrand Calenge, l'objectif premier de la politique documentaire réside dans les publics à servir : « La politique documentaire veut fournir à des publics la meilleure ressource utile, au regard du contexte et en fonction des priorités affirmées par la collectivité²⁰². » Cette approche impose une double exigence : s'adapter aux besoins des publics tout en portant un regard critique sur l'écosystème. Néanmoins, la notion de « meilleure ressource utile » n'est aujourd'hui pas nette. La question des services qui ont la responsabilité d'acquérir et de fournir de tels outils au sein des universités se pose, les outils pouvant aussi relever du périmètre soit des VP numériques, soit, au niveau technique, des directions des systèmes informatiques, soit des SCD. L'exemple du développement de RAGaRenn, service expérimental de l'université de Rennes, en témoigne : destiné à un usage pour l'ensemble du personnel, l'outil a été développé hors du SCD. Il s'agit donc d'un outil professionnel, qui dépasse le cadre des ressources à fournir par les bibliothèques.

La nature polymorphe des outils d'IA générative peut cependant entraîner des hésitations. Emilie Liard, lors d'un entretien, mentionne notamment l'exemple de demandes d'abonnement à des outils professionnels d'IA générative par le SCD, en particulier dans le domaine juridique. Dans le cadre de telles requêtes, avoir une ligne claire sur la politique documentaire et sur le champ d'outils visés par le SCD s'impose pour arbitrer entre les demandes. Dans l'exemple susmentionné, Emilie Liard expose sa réflexion à propos des bases juridiques : « nous avons vocation à fournir du matériau

¹⁹⁸S. BOTTGER, « Entretien », *op. cit.*

¹⁹⁹A. REGÉ, « Entretien », *op. cit.*

²⁰⁰« GTIA Feuille de route 2025-2026 », couperin.org, s. d.

²⁰¹I. FILLIÂTRE, « Entretien », *op. cit.*

²⁰²B. CALENGE, « Bertrand Calenge », *op. cit.*

brut qui permet aux étudiants d'apprendre le droit. On ne s'abonne pas aux outils professionnels, destinés par exemple, aux cabinets d'avocats, et l'on considère que les étudiants vont apprendre à les connaître en milieu professionnel. » Elle note cependant, que cette ligne directrice n'est pas, à l'heure actuelle, explicite dans les documents formalisant la politique documentaire : « dans notre charte des collections, on n'aborde pas précisément le cas de figure d'outils fonctionnant avec de l'IA. Peut-être qu'il faudrait qu'on reprecise ce qu'on considère être de la documentation et ce qu'on considère être la limite du service rendu par le SCD²⁰³. » La redéfinition et la précision des documents de politique documentaire afin d'intégrer des lignes directrices et des orientations stratégiques concernant les abonnements à des outils d'IA générative est un chantier qu'il faudra mener dans l'ensemble des SCD dans le futur.

Accompagner les chercheurs

En parallèle d'une précision de la ligne d'acquisition des ressources, les bibliothèques doivent continuer leur travail d'accompagnement aux chercheurs, et l'adapter aux nouveaux besoins suscités par l'émergence des IA génératives. Valentine Favel-Kapoian et Susan Kovacs qualifient cette position d'« agir pédagogique²⁰⁴ ». Les bibliothécaires sont particulièrement bien outillés pour l'investir : la fonction de formation des SCD est établie et reconnue, notamment pour les doctorants. Bernard Jacquemin souligne l'urgence : « il y a quand même beaucoup de bibliothécaires qui disent qu'ils se retrouvent face à un public qui ne comprend pas vraiment les outils. [...] le grand classique maintenant, c'est le lecteur qui vient en salle de lecture pour retrouver une référence qui a été hallucinée par une IA. » Le travail d'accompagnement se décline sur plusieurs questions autour de l'IA générative : le fonctionnement des outils, afin de comprendre leurs possibilités et limitations, les questions de droit d'auteur, et d'intégrité scientifique.

Delhaye et Perrin identifient deux enjeux importants dans le positionnement des bibliothèques face à l'IA générative : le « bon usage » de ces technologies et les « garanties de l'accès au savoir ». Ils soulignent que, du fait de la contractualisation liée aux ressources de documentation électronique, les bibliothèques sont partie prenante des discussions à plusieurs titres : « institutionnellement, elles garantissent la conformité et la bonne mise en œuvre des clauses d'exploitation des licences signées ; pédagogiquement, elles prennent en charge l'explication de ces termes au public, en particulier aux étudiants et chercheurs ; lors de l'acquisition de la ressource, elles évaluent les fonctionnalités de recherche des plateformes, qui intègrent de plus en plus d'outils génératifs mis en place par les fournisseurs.²⁰⁵ » Ces dimensions dessinent un positionnement qui articule étroitement deux pans de l'activité des bibliothécaires : la formation des usagers au bon usage des outils d'IA d'une part, et l'évaluation des fonctionnalités proposées par les fournisseurs d'autre part. Ces dimensions existent de manière fortement interdépendante. C'est parce que les bibliothécaires disposent des compétences nécessaires pour évaluer les outils qu'ils sont légitimes pour former les usagers à leur utilisation. En retour, c'est parce qu'ils sont au contact quotidien des publics qu'ils identifient les interrogations concrètes des chercheurs et des

²⁰³E. LIARD, « Entretien », *op. cit.*

²⁰⁴V. FAVEL-KAPOIAN et S. KOVACS, « Accompagner les usagers vers l'IA », *op. cit.*, p. 230

²⁰⁵M. DELHAYE et S. PERRIN, « Bibliothèque de recherche et IA », *op. cit.*

étudiants, et qu'ils sont sollicités, en tant que professionnels de l'information, sur l'usage de ces technologies dans la diffusion des savoirs et la production scientifique. Cette boucle vertueuse entre évaluation et médiation constitue l'une des forces structurelles de la profession dans le contexte actuel. Par ailleurs, la capacité à mener une veille et à suivre de près l'évolution de ces technologies est précieuse dans un contexte mouvant. La labilité des technologies d'IA génératives et les hésitations sur leur encadrement entraîneront probablement des recompositions fréquentes des pratiques et des cadres. Resnik et Hosseini préconisent que les directives qui encadrent les usages de ces outils soient « révisés périodiquement à mesure que l'IA générative est adoptée pour la recherche²⁰⁶ ». Dans ce processus, les professionnels des bibliothèques ont un double rôle : la veille informationnelle sur les bonnes pratiques et le conseil sur les potentiels et limites de ces technologies, d'une part ; le soutien à la recherche par la formation, d'autre part. Ce travail nécessite une veille et une collaboration étroite avec les responsables de déontologie et les référents d'intégrité scientifique au sein des universités.

Défendre l'accès ouvert

Concernant le besoin de conseil pour les choix de licence et de droit d'auteur, Adeline Regé donne un exemple éclairant. Elle observe que « des chercheurs sont scandalisés que leur éditeur ait vendu leur contenu en gros à une boîte d'IA sans leur demander leur avis. Certes, mais l'éditeur a tous les droits ». Plus révélateur encore, des chercheurs hésitent à publier en accès ouvert, craignant que leurs contenus soient utilisés pour entraîner des modèles. D'autres préfèrent une licence CC BY-NC, pensant se protéger. « Or, ça ne marche pas », explique-t-elle : quand on publie avec une licence CC BY-NC, on fait très souvent une cession de droits exclusive à l'éditeur pour les aspects commerciaux. « Mes collègues de l'université de Californie disent qu'au contraire, la seule façon de se protéger, c'est de publier en open access avec une licence CC BY, parce que là il ne peut pas y avoir de monétisation exclusive par l'éditeur. » La compréhension des implications légales et techniques des outils d'IA génératives est ainsi essentielle pour permettre aux professionnels des bibliothèques de mener à bien leurs missions de conseil auprès des chercheurs. Cette fonction de médiation prend une importance particulière dans le contexte actuel, où les chercheurs font face à des injonctions parfois contradictoires : publier beaucoup et rapidement tout en respectant les principes de la science ouverte ; valoriser leurs travaux dans des revues à fort impact tout en contestant les indicateurs bibliométriques ; accéder aux meilleures ressources tout en participant à la contestation des modèles économiques des grands éditeurs.

Le bon positionnement des bibliothécaires sur ces questions est celui du conseil, non de la prescription. Raluca Pierrot le formule avec netteté : « Il ne faut surtout pas que ce soit une bataille des bibliothécaires qui font la leçon sur l'intégrité scientifique. Il faut que ça vienne des communautés scientifiques, sinon ça ne peut pas marcher²⁰⁷. » Adeline Regé précise : « Notre rôle est plutôt dans le conseil et la sensibilisation, notamment des jeunes chercheurs et des doctorants, pour expliquer ce qui se passe du côté des éditeurs, comment ils travaillent, expliquer aussi ce qu'est un paper mill²⁰⁸. » Madeleine Geroudet synthétise ce positionnement : « ça nous positionne plutôt sur un rôle de mise à disposition de clés dans un débat, pour une décision qui relève

²⁰⁶D. B. RESNIK et M. HOSSEINI, « The ethics of using artificial intelligence in scientific research », *op. cit.*, p. 1515

²⁰⁷R. PIERROT, « Entretien », *op. cit.*

²⁰⁸A. REGÉ, « Entretien », *op. cit.*

effectivement de la recherche derrière²⁰⁹. » Ce rôle de conseil s'exerce avec une acuité particulière sur les questions de science ouverte, que l'émergence de l'IA générative vient à la fois conforter et compliquer.

La vision que les bibliothécaires portent sur la qualité de l'écosystème informationnel peut leur permettre de défendre l'importance du maintien de pratiques de science ouverte, y compris face aux interrogations soulevées par l'IA générative. Cette défense s'exerce simultanément à plusieurs échelles. Les professionnels travaillant à la science ouverte et à la gestion des archives institutionnelles peuvent avoir un rôle de conseil individuel auprès des chercheurs. Julien Baudry formule ainsi le lien entre science ouverte et IA : « même si de prime abord le réflexe serait de se dire qu'on ne veut pas que les travaux des chercheurs soient moissonnés par l'IA parce qu'ils vont être exploités commercialement, etc., personnellement je préfère que ce soient ces publications qui soient le plus visibles et le plus mises en avant et présentes dans les données d'entraînement plutôt que des publications non scientifiques, pour apporter de la fiabilité à des outils de plus en plus utilisés²¹⁰. » Cette mission de défense reste néanmoins dépendante d'orientations stratégiques qui dépassent les seules considérations des bibliothèques. Damien Belvèze rappelle qu'outre l'aspect éthique et de qualité informationnelle, issu du mouvement de l'open source, l'accès ouvert est aussi un levier de croissance économique : « c'est d'ailleurs pour cette deuxième raison que l'*open access*, après avoir eu des difficultés, a finalement pu décoller. C'est grâce à l'argent que la Commission européenne y a mis, et la Commission européenne en attend de la croissance pour tous les pays qui lui sont rattachés. » Les considérations qui président aux politiques d'accès ouvert ne sont donc pas seulement documentaires : elles sont aussi économiques. Les bibliothécaires, in fine, sont des exécutants qui suivent des préconisations adoptées par des instances supérieures. Ce constat, toutefois, ne diminue pas la portée de leur rôle, il la précise. C'est parce que les bibliothécaires n'ont pas de motivations financières ni d'enjeux de croissance qu'ils sont en mesure de porter un regard détaché de ces considérations et de défendre des lignes d'action concentrées sur la qualité de l'information et du savoir mis en circulation. Les associations et regroupements professionnels peuvent porter ces principes à un niveau plus politique et réglementaire. Le récent communiqué de l'ADBU sur la question de la science ouverte à l'heure de l'intelligence artificielle²¹¹ atteste du fait que les instances professionnelles s'emparent de cette question et jouent pleinement leur rôle sur ces questions stratégiques.

²⁰⁹M. GEROUDET, « Entretien », *op. cit.*

²¹⁰J. BAUDRY, « Entretien », 3 novembre 2025

²¹¹ADBU, « Ouvrir la science à l'heure de l'intelligence artificielle : pourquoi et comment ? », *op. cit.*

CONCLUSION

Ce travail a cherché à dresser un panorama des bouleversements que les technologies d'intelligence artificielle générative entraînent dans la manière dont les professionnels des bibliothèques appréhendent la documentation électronique. Pour ce faire, nous avons cherché à identifier la manière dont l'écosystème de la publication scientifique était modifié par ces questions.

Le premier constat qui s'impose est celui de l'ampleur et de la complexité des recompositions en cours. L'écosystème de la publication scientifique n'était pas stabilisé avant l'arrivée de l'IA générative et plusieurs tensions structurelles le traversaient, notamment la concentration du marché éditorial, l'inflation des publications, la fragilité des mécanismes d'évaluation par les pairs, les coûts croissants de la documentation électronique. Loin de constituer une rupture isolée, les technologies d'IA générative agissent comme un catalyseur qui accélère et intensifie des dynamiques préexistantes. Elles créent également de nouvelles questions, le paradoxe central mis en lumière résidant dans l'interdépendance entre l'IA génératives et l'écosystème de la science ouverte : ces technologies se nourrissent des données en libre accès pour fonctionner, mais fragilisent en retour les fondements mêmes du contrat social sur lequel repose l'accès ouvert. La captation des ressources publiques par des acteurs privés à des fins commerciales, la sursollicitation des infrastructures ouvertes et la désaffection des usagers humains pour les sources primaires constituent autant de menaces pour la viabilité d'un écosystème informationnel dont la qualité conditionne pourtant celle des modèles d'IA eux-mêmes.

Le deuxième apport de ce travail concerne l'analyse des solutions technologiques et de leurs implications. L'examen des fondements techniques des LLM, la compréhension de leur nature probabiliste et de leurs limitations mettent en évidence un décalage entre les capacités apparentes de ces outils et leurs limites réelles. Ce décalage est d'autant plus préoccupant que la communication commerciale qui entoure ces technologies tend à minimiser leurs biais et leurs insuffisances. L'existence de solutions plus maîtrisées, comme les systèmes RAG, ouvre néanmoins des perspectives pour des usages documentaires plus fiables et plus transparents. Les initiatives publiques examinées témoignent d'une appropriation prudente mais réelle de ces technologies, mise au service des missions préexistantes des institutions plutôt qu'en substitution de celles-ci.

Le troisième axe de ce travail a mis en évidence l'état encore inachevé du cadre normatif. Si des grandes tendances émergent la prolifération de chartes, de guides et de préconisations, émanant d'acteurs aux légitimités et aux intérêts variables, crée un paysage fragmenté. Les éditeurs privés, du fait de leur réactivité et de leurs moyens, se trouvent en position de prescripteurs de fait des normes d'usage, ce qui soulève des questions quant à la neutralité de ces cadres. La normalisation reste, par ailleurs, compliquée par la rapidité d'évolution des technologies et par la nécessité d'adapter les règles aux spécificités disciplinaires. Des questions de fond demeurent ouvertes, en particulier celle de la réforme des systèmes d'évaluation de la recherche, dont les indicateurs quantitatifs, déjà décriés auparavant, risquent de perdre toute pertinence discriminante dans un contexte où la production de contenu est facilitée par l'IA générative.

Enfin, la dernière partie se concentre sur les compétences et missions des professionnels des bibliothèques. Elle révèle une profession consciente des enjeux, mais dont la légitimité sur ces questions reste fragile et insuffisamment reconnue. Les compétences des professionnels des bibliothèques sont pourtant mobilisables face aux défis posés par ces technologies : évaluation critique des outils de diffusion de l'information, maîtrise des données et des métadonnées, connaissance fine de l'écosystème éditorial, capacité de formation aux compétences informationnelles. La position de neutralité des bibliothèques, exemptes d'intérêts commerciaux dans le traitement des savoirs, constitue un atout structurel dans un paysage dominé par des logiques marchandes. Néanmoins, cette pertinence objective ne suffit pas : elle doit s'accompagner d'une affirmation de la profession. Les entretiens ont révélé un ethos de modestie qui, s'il témoigne d'une rigueur intellectuelle appréciable, peut aussi conduire à une auto-exclusion des espaces de décision.

Plusieurs pistes prospectives se dégagent de ce panorama. La première concerne le maintien et le renforcement de l'accès ouvert, non par un réflexe de fermeture face aux captations, mais par le développement de nouveaux modèles de réciprocité qui permettent de concilier ouverture des données et soutenabilité de l'écosystème. Les bibliothèques, en tant qu'institutions défendant la libre circulation des savoirs, ont un rôle à jouer dans la promotion et l'expérimentation des modèles. La deuxième piste touche à la montée en compétence de la profession. L'acculturation aux fonctionnements techniques de l'IA, sans nécessairement viser une expertise informatique, apparaît comme une condition nécessaire pour que les bibliothécaires puissent évaluer les outils, conseiller les chercheurs et participer aux instances de décision. Les modalités de formation identifiées sont encore tâtonnantes mais réelles. Elles constituent des dispositifs prometteurs qu'il conviendrait de généraliser et de pérenniser. La troisième piste concerne le positionnement stratégique de la profession. Les bibliothécaires ont vocation à être non seulement des utilisateurs et des évaluateurs de ces technologies, mais aussi des contributeurs à leur développement, en tant que fournisseurs de données de qualité et « architectes de confiance documentaire ». Ce repositionnement suppose une reconnaissance institutionnelle qui reste à construire.

Ce travail, on l'a dit en introduction, n'est qu'un instantané. Les technologies d'IA générative continuent d'évoluer à un rythme soutenu, et les recompositions de l'écosystème de la publication scientifique sont loin d'être achevées. De nouvelles questions émergeront probablement. L'impact de ces technologies sur les pensées humaines, et de la manière dont se développeront les habitudes informationnelles des générations qui auront grandi de manière native avec ces outils sont, notamment, des questions qui modifieront en profondeur l'évolution des technologies sur le long terme. Néanmoins, les enjeux fondamentaux identifiés dans ce travail, à savoir la qualité de l'information scientifique, l'équité dans l'accès aux savoirs, la transparence des processus de production et de diffusion des connaissances, sont des préoccupations durables, auxquelles les professionnels des bibliothèques sont outillés pour répondre. Il leur appartient désormais de s'en saisir pleinement.

DÉCLARATION D'USAGE

Plusieurs outils d'IA générative ont été utilisés dans le cadre de ce travail.

- Copilot, et ZoomAI companion ont été utilisés pour la transcription des entretiens menés dans le cadre de cette enquête.

- Claude Opus 4.6 et Claude Sonnet 3.5 ont été utilisés comme aide à la rédaction du texte, en particulier des corrections orthographiques et stylistiques. Aucune idée et aucune structure argumentative n'ont été générées à l'aide d'outils d'IA générative.

BIBLIOGRAPHIE

IA GÉNÉRATIVE, GÉNÉRALITÉS

AGULHON, Sophie, 2023. Numérique ? Algorithme ? Big data ? Comprendre les technologies autour de l'Intelligence Artificielle. In : GUÉNOT, Françoise (éd.), *L'IA éducative. L'intelligence artificielle dans l'enseignement supérieur* [en ligne]. Paris : Bréal. [Consulté le 26 mai 2025]. Disponible à l'adresse : <https://hal.science/hal-04871046>

BENDER, Emily M., GEBRU, Timnit, MCMILLAN-MAJOR, Angelina et SHMITCHELL, Shmargaret, 2021. On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? . In : *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency* [en ligne]. Virtual Event Canada : ACM. 3 mars 2021. pp. 610-623. [Consulté le 18 février 2026]. ISBN 978-1-4503-8309-7. Disponible à l'adresse : <https://dl.acm.org/doi/10.1145/3442188.3445922>

BERGSTROM, Tracy et RUEDIGER, Dylan, 2024. *A Third Transformation? Generative AI and Scholarly Publishing* [en ligne]. Ithaka S+R. [Consulté le 6 janvier 2026]. Disponible à l'adresse : <http://sr.ithaka.org/?p=321519>

BERMÈS, Emmanuelle et LECLAIRE, Céline, 2022. Vers l'intelligence artificielle et au-delà ! Une feuille de route pour la BnF. *Arabesques*. 14 octobre 2022. N° 107, pp. 8-9. DOI [10.35562/arabesques.3082](https://doi.org/10.35562/arabesques.3082).

BIBLIOTHÈQUE NATIONALE ET UNIVERSITAIRE (BNU), 2024. *Le projet Gallica Images. L'intelligence artificielle pour l'accès aux grands corpus d'images* [en ligne]. 18 juin 2024. [Consulté le 14 février 2026]. Disponible à l'adresse : <https://www.youtube.com/watch?v=IVMfpR96jrk>

BOMMASANI, Rishi, KLYMAN, Kevin, KAPOOR, Sayash, LONGPRE, Shayne, XIONG, Betty, MASLEJ, Nestor et LIANG, Percy, 2025. *The 2024 Foundation Model Transparency Index* [en ligne]. 4 mars 2025. arXiv. arXiv:2407.12929. [Consulté le 18 juin 2025]. Disponible à l'adresse : <http://arxiv.org/abs/2407.12929>

BRANDY, Grégor, 2026. Aux Etats-Unis, des bibliothécaires désemparés face aux demandes de livres inventés par l'IA. *Le Monde* [en ligne]. 24 janvier 2026. [Consulté le 6 février 2026]. Disponible à l'adresse : https://www.lemonde.fr/culture/article/2026/01/24/aux-etats-unis-des-bibliothecaires-desempares-face-aux-demandes-de-livres-inventes-par-l-ia_6663922_3246.html

BRITTAİN, Blake, 2025. Anthropic wins key US ruling on AI training in authors' copyright lawsuit. *Reuters* [en ligne]. 24 juin 2025. [Consulté le 26 juin 2025]. Disponible à l'adresse : <https://www.reuters.com/legal/litigation/anthropic-wins-key-ruling-ai-authors-copyright-lawsuit-2025-06-24/>

CESSAC, Marjorie, 2025. Les agents IA, vos « futurs collègues ». *Le Monde*. 19 juin 2025. p. 19.

CHAGUÉ, Alix et ROMARY, Laurent, 2022. L'intelligence artificielle, une ouverture du champ des possibles. *Arabesques*. 7 octobre 2022. N° 107, pp. 4-5. DOI . [10.35562/arabesques.3043](https://doi.org/10.35562/arabesques.3043)

COHEN-BOULAKIA, Sarah, 2025. Vers des IA génératives intègres, 14 avril 2025, Ofis [en ligne]. [Consulté le 15 décembre 2025]. Disponible à l'adresse : <https://www.ofis-france.fr/wp-content/uploads/2025/04/06-Session1-part22025-Sarah-CohenPresentation-IAGen-DATAIA-Ofis.pdf>

CUXAC, Pascal, 2022. L'IA et la fouille de textes à l'INIST : l'IA à portée de tous ? *Arabesques*. 12 octobre 2022. N° 107, pp. 18-19. DOI [10.35562/arabesques.3098](https://doi.org/10.35562/arabesques.3098)

Équipements et usages du numérique, 2026. *Arcep* [en ligne]. [Consulté le 1 mars 2026]. Disponible à l'adresse : <https://www.arcep.fr/actualites/actualites-et-communiques/detail/n/equipements-et-usages-du-numerique-090226.html>

FARRELL, Henry, GOPNIK, Alison, SHALIZI, Cosma et EVANS, James, 2025. Large AI models are cultural and social technologies. *Science* [en ligne]. 14 mars 2025. [Consulté le 19 janvier 2026]. Disponible à l'adresse : <https://www.science.org/doi/10.1126/science.adt9819>

Generative AI Licensing Agreement Tracker, [sans date]. *Ithaka S+R* [en ligne]. [Consulté le 4 septembre 2025]. Disponible à l'adresse : <https://sr.ithaka.org/our-work/generative-ai-licensing-agreement-tracker/>

GEOFFROY, Géraldine, 2022. Pour une approche décomplexée de l'IA. *Arabesques*. 12 octobre 2022. N° 107, pp. 22-23. DOI [10.35562/arabesques.3095](https://doi.org/10.35562/arabesques.3095).

Géopolitique de l'intelligence artificielle : Quo vadis Europa ?, [sans date]. . HALSHS : archive ouverte en Sciences de l'Homme et de la Société. ISBN 978-84-10448-45-2.

HENNO, Jacques, 2017. 1956 : et l'intelligence artificielle devint une science. *Les Echos* [en ligne]. 21 août 2017. [Consulté le 9 janvier 2026]. Disponible à l'adresse : <https://www.lesechos.fr/2017/08/1956-et-lintelligence-artificielle-devint-une-science-181042>

ILAAS, 2025. Communiqué de presse. [en ligne]. 18 juin 2025. [Consulté le 3 février 2026]. Disponible à l'adresse : <https://www.ilaas.fr/wp-content/uploads/2025/07/20250605.CommuniqueDePresse.pdf>

LIESENFELD, Andreas et DINGEMANSE, Mark, 2024. Rethinking open source generative AI: open-washing and the EU AI Act. In : *Proceedings of the 2024 ACM Conference on Fairness, Accountability, and Transparency* [en ligne]. New York, NY, USA : Association for Computing Machinery. 5 juin 2024. pp. 1774-1787. [Consulté le 12 février 2026]. FAccT '24. ISBN 979-8-4007-0450-5. Disponible à l'adresse : <https://dl.acm.org/doi/10.1145/3630106.3659005>

LIGOZAT, Anne-Laure, LEFEVRE, Julien, BUGEAU, Aurélie et COMBAZ, Jacques, 2022. Unraveling the Hidden Environmental Impacts of AI Solutions for Environment Life Cycle Assessment of AI Solutions. *Sustainability* [en ligne]. 25 avril 2022. Vol. 14, n° 9. [Consulté le 6 février 2026]. Disponible à l'adresse : <https://www.mdpi.com/2071-1050/14/9/5172>

MOSHA, Neema Florence, 2025. The role of artificial intelligence tools in enhancing accessibility and usability of electronic resources in academic libraries. *Library Management*. 28 janvier 2025. Vol. 46, n° 1/2, pp. 132-157. DOI 10.1108/LM-08-2024-0088.

Partnership with Axel Springer to deepen beneficial use of AI in journalism, 2024. [en ligne]. [Consulté le 22 janvier 2026]. Disponible à l'adresse : <https://openai.com/index/axel-springer-partnership/>

PÔLE SOCIÉTÉ DU CREDOC, 2026. *Baromètre du numérique* [en ligne]. [Consulté le 1 mars 2026]. Disponible à l'adresse : https://www.arcep.fr/uploads/tx_gspublication/barometre-du-numerique-edition-2026_RAPPORT.pdf

Quelles bonnes IA génératives pour faire de la recherche ? Qu'est-ce qu'un modèle ouvert ?, 2025. [en ligne]. [Consulté le 27 juin 2025]. Disponible à l'adresse : <https://www.youtube.com/watch?v=w0ejerBv6Mc>

SCHONFELD, Roger C., 2024. Tracking the Licensing of Scholarly Content to LLMs. *The Scholarly Kitchen* [en ligne]. 15 octobre 2024. [Consulté le 4 septembre 2025]. Disponible à l'adresse : <https://scholarlykitchen.sspnet.org/2024/10/15/licensing-scholarly-content-llms/>

SHUMAILOV, Ilia, SHUMAYLOV, Zakhar, ZHAO, Yiren, GAL, Y., PAPERNOT, Nicolas et ANDERSON, Ross, 2023. The Curse of Recursion: Training on Generated Data Makes Models Forget. *ArXiv*. 27 mai 2023.

SOMARAJAN, Sajin, [sans date]. Impacts of Data Poisoning in AI. *Infosys* [en ligne]. [Consulté le 13 juin 2025]. Disponible à l'adresse : <https://blogs.infosys.com/digital-experience/emerging-technologies/impacts-of-data-poisoning-in-ai.html>

Sorbonne Université rejoint une alliance inédite sur l'IA avec Mistral AI. *Sorbonne Université* [en ligne]. [Consulté le 6 mai 2025]. Disponible à l'adresse : <https://www.sorbonne-universite.fr/actualites/sorbonne-universite-rejoint-une-alliance-inedite-sur-lia-avec-mistral-ai>

TÉTART, Florent, 2025. L'intelligence artificielle appliquée aux bibliothèques : une opportunité plutôt qu'une menace. *I2D - Information, données & documents*. 3 septembre 2025. N° 1, pp. 78-133. DOI 10.3917/i2d.251.0078.

TOULET, Anne, MICHEL, Franck et TCHECHMEDJIEV, Andon, 2022. Le projet ISSA : l'intelligence artificielle au service de la recherche bibliographique. *Arabesques*. 7 octobre 2022. N° 107, pp. 6-7. DOI 10.35562/arabesques.3046.

SERVICES D'IA GÉNÉRATIVES

Comment utiliser SophIA ?, [sans date]. [en ligne]. [Consulté le 5 février 2026]. Disponible à l'adresse : <https://aide.cairn.info/article/378-comment-utiliser-sophia>

Consensus: AI for Research, [sans date]. [en ligne]. [Consulté le 10 février 2026]. Disponible à l'adresse : <https://consensus.app/search/>

DeepL AI Platform: Translation, Voice & Agentic Workflows, [sans date]. [en ligne]. [Consulté le 10 février 2026]. Disponible à l'adresse : <https://www.deepl.com>

Elicit: AI for scientific research, [sans date]. [en ligne]. [Consulté le 10 février 2026]. Disponible à l'adresse : <https://elicit.com/welcome>

Great writing made easy with ProWritingAid, [sans date]. [en ligne]. [Consulté le 6 février 2026]. Disponible à l'adresse : <https://prowritingaid.com/>

Our Features | Grammarly, [sans date]. [en ligne]. [Consulté le 6 février 2026]. Disponible à l'adresse : <https://www.grammarly.com/features>

Prism, [sans date]. *Openai.com* [en ligne]. [Consulté le 4 février 2026]. Disponible à l'adresse : <https://openai.com/fr-FR/prism/>

SciSpace AI Research Agent | 150+ Tools, 280 M Papers, [sans date]. *SciSpace* [en ligne]. [Consulté le 10 février 2026]. Disponible à l'adresse : <https://scispace.com>

Scopus AI | Elsevier, [sans date]. *www.elsevier.com* [en ligne]. [Consulté le 4 février 2026]. Disponible à l'adresse : <https://www.elsevier.com/products/scopus/scopus-ai>

Trinka: AI Writing and Grammar Checker Tool, [sans date]. [en ligne]. [Consulté le 6 février 2026]. Disponible à l'adresse : <https://www.trinka.ai>

AUTOUR DES QUESTIONS D'OPEN ACCESS

ACION, Laura, RAJNGEWERC, Mariela, RANDALL, Gregory et ETCHEVERRY, Lorena, 2023. Generative AI poses ethical challenges for open science. *Nature Human Behaviour*. novembre 2023. Vol. 7, n° 11, pp. 1800-1801. DOI 10.1038/s41562-023-01740-4.

AMUE, 2022. Le numérique universitaire des BU. . 2022.

Baromètre français de la Science Ouverte, [sans date]. [en ligne]. [Consulté le 26 janvier 2026]. Disponible à l'adresse : <https://barometredelascienceouverte.esr.gouv.fr/>

Budapest Open Access Initiative, [sans date]. [en ligne]. [Consulté le 10 juin 2025]. Disponible à l'adresse : <https://www.budapestopenaccessinitiative.org/read/>

DELHAYE, Marlène, 2017. Portrait métier # 1 : bibliothécaire open access. *Bulletin des bibliothèques de France*. 1 janvier 2017. N° 13, pp. 20-21.

EFFODUH, Okechukwu, 2025. Scholarly Publishing in the Era of Open Access and Generative Artificial Intelligence. *Journal of Scholarly Publishing*. 1 juillet 2025. Vol. 56, n° 3, pp. 529-553. DOI 10.3138/jsp-2024-0102.

EXTANCE, Andy, 2018. How AI technology can tame the scientific literature. *Nature*. 10 septembre 2018. Vol. 561, n° 7722, pp. 273-274. DOI 10.1038/d41586-018-06617-5.

FARGIER, Nathalie et BARBORINI, Yannick, 2024. AP2024 | HAL et l'IA. [en ligne]. 21 novembre 2024. [Consulté le 26 mai 2025]. Disponible à l'adresse : <https://www.canal-u.tv/161272>

GEORGE, Sigy, 2024. Impacts of Open Access and Open Repositories on Acquisitions and Collection Development. *International Journal of Social Science Humanity & Management Research* [en ligne]. 23 août 2024. Vol. 3, n° 08. [Consulté le 20 mai 2025]. DOI 10.58806/ijsshmr.2024.v3i8n15. Disponible à l'adresse : <https://ijsshmr.com/v3i8/15.php>

GOËTA, Samuel, 2025. Ne jetez pas le bébé de l'open data avec l'eau du bain de l'IAG ! *La Gazette des communes*. 16 juin 2025. Vol. 2771, n° 2771, pp. 27.

HOSSEINI, Mohammad, HORBACH, Serge P. J. M., HOLMES, Kristi et ROSS-HELLAUER, Tony, 2025. Open Science at the generative AI turn: An exploratory analysis of challenges and opportunities. *Quantitative Science Studies*. 27 janvier 2025. Vol. 6, pp. 22-45. DOI 10.1162/qss_a_00337.

JAILLANT, Lise et CAPUTO, Annalina, 2022. Unlocking digital archives: cross-disciplinary perspectives on AI and born-digital data. *AI & SOCIETY*. 1 septembre 2022. Vol. 37, n° 3, pp. 823-835. DOI 10.1007/s00146-021-01367-x.

KELLER, Paul, 2025. Beyond AI and copyright. Funding a sustainable information ecosystem. . juin 2025.

KELLER, Paul et TARKOWSKI, Alek, 2021. *The Paradox of Open* [en ligne]. 21 septembre 2021. Open Future. [Consulté le 19 janvier 2026]. Disponible à l'adresse : <https://paradox.openfuture.eu/>

KOSMOPOULOS, Christine et PUMAIN, Denise, 2025. Un grand pas de plus vers la science ouverte. *Cybergeog: European Journal of Geography* [en ligne]. 8 décembre 2025. [Consulté le 14 janvier 2026]. Disponible à l'adresse : <https://journals.openedition.org/cybergeog/42519#ftn2>

LO, Leo S., 2025. Generative AI and Open Access Publishing: A New Economic Paradigm. *Library Trends*. février 2025. Vol. 73, n° 3, pp. 160-176. DOI 10.1353/lib.2025.a961190.

L'offre de services de la BnF pour les acteurs de l'intelligence artificielle (IA), [sans date]. *BnF - Site institutionnel* [en ligne]. [Consulté le 2 février 2026]. Disponible à l'adresse : <https://www.bnf.fr/fr/loffre-de-services-de-la-bnf-pour-les-acteurs-de-lintelligence-artificielle-ia>

PASTEUR, Ceris-Institut, 2025. Bases de données commerciales : après s'être désabonné de Scopus, le CNRS coupe l'accès au Web of Science. *Open science : évolutions, enjeux et pratiques* [en ligne]. 18 décembre 2025. [Consulté le 14 janvier 2026]. Disponible à l'adresse : <https://openscience.pasteur.fr/2025/12/18/bases-de-donnees-commerciales-apres-setre-desabonne-de-scopus-le-cnrs-coupe-laces-au-web-of-science/>

PERRET, Camille, [sans date]. LibGuides: Science ouverte et données de la recherche en bibliothèque: Comprendre la publication en open access. [en ligne]. [Consulté le 4 janvier 2026]. Disponible à l'adresse : <https://enssib.libguides.com/c.php?g=705396&p=5083597>

POUYLLAU, Stéphane, 2024. La place de « l'IA » dans ISIDORE. [en ligne]. 2024. [Consulté le 20 juin 2025]. Disponible à l'adresse : <https://isore.hypotheses.org/files/2025/04/ISIDORE-2030-SPOUYLLAU-2.pdf>

PRIDE, David, 2024. Retour d'expérience d'usage de l'IA : CORE. [en ligne]. 21 novembre 2024. [Consulté le 26 mai 2025]. Disponible à l'adresse : <https://www.canal-u.tv/161265>

SCHAULI, Mickaël, 2025. *Wikipédia et l'intelligence artificielle générative* [en ligne]. janvier 2025. Université de Strasbourg. [Consulté le 26 mai 2025]. Disponible à l'adresse : <https://hal.science/hal-05026643>

SHEARER, Kathleen et WALK, Paul, 2025. *The impact of AI bots and crawlers on open repositories: Results of a COAR survey, April 2025* [en ligne]. Confederation of Open Access Repositories. [Consulté le 22 janvier 2026]. Disponible à l'adresse : <https://coar-repositories.org/wp-content/uploads/2025/06/Report-of-the-COAR-Survey-on-AI-Bots-June-2025-1.pdf>

SUBER, Peter, 2016. 1. Une définition. In : *Qu'est-ce que l'accès ouvert ?* [en ligne]. Marseille : OpenEdition Press. pp. 23-44. Encyclopédie numérique. [Consulté le 9 janvier 2026]. ISBN 978-2-8218-6980-6. Disponible à l'adresse : <https://books.openedition.org/oep/1605>

TAY, Aaron, 2025. The reproducibility and interpretability of academic Ai search engines like Primo Research Assistant, Web of Science Research Assistant, Scopus Ai and more. *Musing about librarianship* [en ligne]. 14 avril 2025. [Consulté le 31 juillet 2025]. Disponible à l'adresse : <https://musingsaboutlibrarianship.blogspot.com/2025/04/the-reproducibility-and.html>

VAN DER GRAAF, Maurits, 2023. *Open science services by research libraries. Organisational perspectives* [en ligne]. LIBER et ADBU. [Consulté le 6 mai 2025]. Disponible à l'adresse : <https://www.enssib.fr/bibliotheque-numerique/documents/71319-open-science-services-by-research-libraries-organisational-perspectives.pdf>

WHITE, Molly, 2025. "Wait, not like that": Free and open access in the age of generative AI. *Citation Needed* [en ligne]. 14 mars 2025. [Consulté le 22 janvier 2026]. Disponible à l'adresse : <https://www.citationneeded.news/free-and-open-access-in-the-age-of-generative-ai/>

ÉCOSYSTÈME DE LA PUBLICATION SCIENTIFIQUE

Academ-AI Suspected AI papers, [sans date]. *Academ-AI* [en ligne]. [Consulté le 6 juin 2025]. Disponible à l'adresse : <https://www.academ-ai.info/>

Access to Science & Scholarship 2024 - Summary Report, 2024. [en ligne]. MIT Press Workshop. [Consulté le 22 janvier 2026]. Disponible à l'adresse : <https://assets.pubpub.org/oiucyi8r/MIT%20Press%20Workshop%20-%20Access%20to%20Science%20&%20Scholarship%202024%20-%20Summary%20Report-71731528809632.pdf>

AI Policies in Academic Publishing 2025: Guide & Checklist - Thesify - Ethical AI Tools That Improve Your Academic Writing, [sans date]. [en ligne].

[Consulté le 28 janvier 2026]. Disponible à l'adresse : <https://www.thesify.ai/blog/ai-policies-academic-publishing-2025>

arXiv announces new policy on ChatGPT and similar tools, [sans date]. *arXiv blog* [en ligne]. [Consulté le 26 mai 2025]. Disponible à l'adresse : <https://blog.arxiv.org/2023/01/31/arxiv-announces-new-policy-on-chatgpt-and-similar-tools/>

Authorship and AI tools, 2023. *COPE: Committee on Publication Ethics* [en ligne]. [Consulté le 2 octobre 2025]. Disponible à l'adresse : <https://publicationethics.org/guidance/cope-position/authorship-and-ai-tools>

CC Signals: A New Social Contract for the Age of AI, [sans date]. *Creative Commons* [en ligne]. [Consulté le 28 août 2025]. Disponible à l'adresse : <https://creativecommons.org/ai-and-the-commons/cc-signals/>

Colloque 2025 - Comment l'IA générative transforme les pratiques de recherche : nouveaux enjeux d'intégrité scientifique, [sans date]. *OFIS* [en ligne]. [Consulté le 6 juin 2025]. Disponible à l'adresse : <https://www.ofis-france.fr/colloque/colloque-2025/>

CROTTY, David, 2025. 2025 Update: Quantifying Consolidation in the Scholarly Journals Market. *The Scholarly Kitchen* [en ligne]. 20 août 2025. [Consulté le 26 janvier 2026]. Disponible à l'adresse : <https://scholarlykitchen.sspnet.org/2025/08/20/2025-update-quantifying-consolidation-in-the-scholarly-journals-market/>

Directory of Open Access Journals – DOAJ, 2026. [en ligne]. [Consulté le 26 janvier 2026]. Disponible à l'adresse : <https://doaj.org/>

DREYFUS, Louis et FENOGLIO, Jérôme, 2024. Intelligence artificielle : un accord de partenariat entre « Le Monde » et OpenAI. *Le Monde* [en ligne]. 13 novembre 2024. [Consulté le 27 janvier 2026]. Disponible à l'adresse : https://www.lemonde.fr/le-monde-et-vous/article/2024/03/13/intelligence-artificielle-un-accord-de-partenariat-entre-le-monde-et-openai_6221836_6065879.html

Edition scientifique et intelligence artificielle (IA) : Les rencontres François Gèze, 2025. [en ligne]. [Consulté le 11 juin 2025]. Disponible à l'adresse : <https://www.youtube.com/watch?v=HTFC7q0A0IA>

Enjeux d'intégrité scientifique à diverses étapes de la recherche, 2025. [en ligne]. [Consulté le 27 juin 2025]. Disponible à l'adresse : <https://www.youtube.com/watch?v=RJ1WNBbWZOY>

EUROPEAN COMMISSION. DIRECTORATE GENERAL FOR RESEARCH AND INNOVATION, 2024. *Study on scientific publishing in Europe: development, diversity, and transparency of costs* [en ligne]. 2024. Publications Office. [Consulté le 23 janvier 2026]. Disponible à l'adresse : <https://data.europa.eu/doi/10.2777/89349>

EVANS, Ian, 2025. How researchers can use GenAI to get the info they need quicker than ever. *www.elsevier.com* [en ligne]. 15 février 2025. [Consulté le 4 février 2026]. Disponible à l'adresse : <https://www.elsevier.com/connect/how-researchers-can-use-genai-to-get-the-info-they-need-quicker-than-ever>

FRAISIER-VANNIER, Ophélie, 2024. Mésusages des LLM dans la publication scientifique. . 30 juillet 2024.

Frontiers introduces FAIR² Data Management, 2025. *Research Information* [en ligne]. [Consulté le 4 avril 2025]. Disponible à l'adresse : <https://www.researchinformation.info/news/frontiers-introduces-fair2-data-management/>

Generative AI, scholarly and cultural language models, and the return of content, 2023. *LorcanDempsey.net* [en ligne]. [Consulté le 25 mai 2025]. Disponible à l'adresse : <https://www.lorcandempsey.net/generative-ai-a-note-about-content/>

HANSEN, Dave, 2025. AI Licensing: An Interview with Ben Denne of Cambridge University Press. *Authors Alliance* [en ligne]. 17 mars 2025. [Consulté le 2 octobre 2025]. Disponible à l'adresse : <https://www.authorsalliance.org/2025/03/17/ai-licensing-an-interview-with-ben-denno-of-cambridge-university-press/>

HANSON, Mark A., BARREIRO, Pablo Gómez, CROSETTO, Paolo et BROCKINGTON, Dan, 2024. The strain on scientific publishing. *Quantitative Science Studies*. 1 novembre 2024. Vol. 5, n° 4, pp. 823-843. DOI 10.1162/qss_a_00327.

HARDINGES, Jack, PEARSON, Sarah et ROSS, Rebecca, 2025. *From Human Content to Machine Data* [en ligne]. juin 2025. Disponible à l'adresse : https://creativecommons.org/wp-content/uploads/2025/06/Human-Content-to-Machine-Data_Final.pdf

HAUSTEIN, Stefanie, SCHARES, Eric, ALPERIN, Juan Pablo, HARE, Madelaine, BUTLER, Leigh-Ann et SCHÖNFELDER, Nina, 2024. *Estimating global article processing charges paid to six publishers for open access between 2019 and 2023* [en ligne]. 23 juillet 2024. arXiv. arXiv:2407.16551. [Consulté le 17 juillet 2025]. Disponible à l'adresse : <http://arxiv.org/abs/2407.16551>

Intelligence artificielle, [sans date]. *Syndicat national de l'édition* [en ligne]. [Consulté le 12 février 2026]. Disponible à l'adresse : <https://www.sne.fr/numerique-2/intelligence-artificielle/>

KAGAN-KANS, Dan, 2025. Baby Shoggoth Is Listening. *The American Scholar* [en ligne]. 29 octobre 2025. [Consulté le 13 février 2026]. Disponible à l'adresse : <https://theamericanscholar.org/baby-shoggoth-is-listening/>

KOBAK, Dmitry, GONZÁLEZ-MÁRQUEZ, Rita, HORVÁT, Emőke-Ágnes et LAUSE, Jan, 2025. Delving into LLM-assisted writing in biomedical publications through excess vocabulary. *Science Advances* [en ligne]. 4 juillet 2025. Vol. 11, n° 27. [Consulté le 20 janvier 2026]. DOI 10.1126/sciadv.adt3813. Disponible à l'adresse : <http://arxiv.org/abs/2406.07016>

KOSMOPOULOS, Christine, 2025. Entre prédation et désincarnation, l'édition scientifique à l'épreuve de l'IA. *Cybergeog: European Journal of Geography* [en ligne]. 13 février 2025. [Consulté le 10 juin 2025]. Disponible à l'adresse : <https://journals.openedition.org/cybergeog/41699>

KWON, Diana, 2024. AI is complicating plagiarism. How should scientists respond? *Nature* [en ligne]. 30 juillet 2024. [Consulté le 6 juin 2025].

DOI 10.1038/d41586-024-02371-z. Disponible à l'adresse : <https://www.nature.com/articles/d41586-024-02371-z>

LAROUSSE, David, 2025a. Comment l'IA bouscule les publications scientifiques. [en ligne]. 7 juillet 2025. [Consulté le 17 juillet 2025]. Disponible à l'adresse : https://www.lemonde.fr/sciences/article/2025/07/07/comment-l-ia-bouscule-les-publications-scientifiques_6619655_1650684.html

LAROUSSE, David, 2025b. Le monde des revues scientifiques au bord de l'asphyxie. e [en ligne]. 7 juillet 2025. [Consulté le 17 juillet 2025]. Disponible à l'adresse : https://www.lemonde.fr/sciences/article/2025/07/07/le-monde-des-revues-scientifiques-au-bord-de-l-asphyxie_6619660_1650684.html

LEMOIGNE, Victoire, 2025. «On n'aura bientôt plus besoin d'écrire»: comment ChatGPT reprogramme notre rapport au langage. *Le Figaro (site web)* [en ligne]. 22 juin 2025. Disponible à l'adresse : https://nouveau.europepress.com/Link/ENSSIBT_1/news%C2%B720250622%C2%B7LFF%C2%B71f04c2e4-98cf-6b94-9dd2-8713057bdca4

Les paper mills, review mills et citation mills, 2024. *Open science : évolutions, enjeux et pratiques* [en ligne]. [Consulté le 6 mai 2025]. Disponible à l'adresse : <https://www.enssib.fr/bibliotheque-numerique/documents/72200-les-paper-mills-review-mills-et-citation-mills.pdf>

Les repères de l'intégrité académique – Responsable Academia, [sans date]. [en ligne]. [Consulté le 11 septembre 2025]. Disponible à l'adresse : <https://irafpa.org/a-propos/regles-dintegrite-academique/>

LIANG, Weixin, IZZO, Zachary, ZHANG, Yaohui, LEPP, Haley, CAO, Hancheng, ZHAO, Xuandong, CHEN, Lingjiao, YE, Haotian, LIU, Sheng, HUANG, Zhi, MCFARLAND, Daniel A. et ZOU, James Y., 2024. Monitoring AI-Modified Content at Scale: A Case Study on the Impact of ChatGPT on AI Conference Peer Reviews. *ICML'24: Proceedings of the 41st International Conference on Machine Learning*. 15 juin 2024. pp. 29575-29620.

Papers and peer reviews with evidence of ChatGPT writing, 2024. *Retraction Watch* [en ligne]. [Consulté le 6 juin 2025]. Disponible à l'adresse : <https://retractionwatch.com/papers-and-peer-reviews-with-evidence-of-chatgpt-writing/>

PERKINS, Mike et ROE, Jasper, 2024. Academic publisher guidelines on AI usage: A ChatGPT supported thematic analysis. *F1000Research*. 16 janvier 2024. Vol. 12, pp. 1398. DOI 10.12688/f1000research.142411.2.

Policy for authors' use of generative AI language tools, [sans date]. *arXiv info* [en ligne]. [Consulté le 26 mai 2025]. Disponible à l'adresse : <https://info.arxiv.org/help/moderation/index.html#policy-for-authors-use-of-generative-ai-language-tools>

PUEHRINGER, Stephan, RATH, Johanna et GRIESEBNER, Teresa, 2021. The political economy of academic publishing: On the commodification of a public good. *PLOS ONE* [en ligne]. 17 juin 2021. Vol. 16, n° 6. [Consulté le 26 janvier 2026]. DOI 10.1371/journal.pone.0253226. Disponible à l'adresse : <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0253226>

Quelles bonnes IA génératives pour faire de la recherche ? Qu'est-ce qu'un modèle ouvert ?, 2025. [en ligne]. [Consulté le 27 juin 2025]. Disponible à l'adresse : <https://www.youtube.com/watch?v=w0ejerBv6Mc>

Rapport de l'UNESCO sur la science 2021, 2021. [en ligne]. [Consulté le 26 janvier 2026]. Disponible à l'adresse : <https://www.unesco.org/reports/science/2021/fr/dataviz/researchers-million-habitants>

Rencontres François Gèze sur l'édition scientifique et l'intelligence artificielle, [sans date]. *Syndicat national de l'édition* [en ligne]. [Consulté le 12 février 2026]. Disponible à l'adresse : https://www.sne.fr/evenement_sne/rencontres-francois-geze-sur-ledition-scientifique-et-lintelligence-artificielle/

RESNIK, David B. et HOSSEINI, Mohammad, 2025. The ethics of using artificial intelligence in scientific research: new guidance needed for a new tool. *AI and Ethics*. 1 avril 2025. Vol. 5, n° 2, pp. 1499-1521. DOI 10.1007/s43681-024-00493-8.

RO, Christine et LEEMING, Jack, 2025. Authorship for sale: Nature investigates how paper mills work. *Nature*. 9 juin 2025. Vol. 642, n° 8068, pp. 823-826. DOI 10.1038/d41586-025-01824-3.

ROUSSEAU-HANS, Françoise, BARTHET, Emilie, WEIL-MIKO, Christine, REGE, Adeline, PERRIN, Sébastien, DELHAYE, Marlène et COLCANAP, Grégory, 2023. *La négociation d'accords transformants* [en ligne]. [Consulté le 23 janvier 2026]. Disponible à l'adresse : <https://hal.science/hal-04009964>

ROYAL SOCIETY, [sans date]. *Science in the age of AI* [en ligne]. [Consulté le 6 mai 2025]. Disponible à l'adresse : <https://royalsociety.org/-/media/policy/projects/science-in-the-age-of-ai/science-in-the-age-of-ai-report.pdf>

RUSSO, Federica, SCHLIESSER, Eric et WAGEMANS, Jean, 2024. Connecting ethics and epistemology of AI. *AI & SOCIETY*. août 2024. Vol. 39, n° 4, pp. 1585-1603. DOI 10.1007/s00146-022-01617-6.

San Francisco Declaration on Research Assessment, 2012. *DORA* [en ligne]. [Consulté le 16 février 2026]. Disponible à l'adresse : <https://sfdora.org/read/>

SINGH CHAWLA, Dalmeet, 2025. Fake AI images will cause headaches for journals. *Nature* [en ligne]. 27 mai 2025. [Consulté le 6 juin 2025]. DOI 10.1038/d41586-025-01488-z. Disponible à l'adresse : <https://www.nature.com/articles/d41586-025-01488-z>

STOKEL-WALKER, Chris, 2023. ChatGPT listed as author on research papers: many scientists disapprove. *Nature*. 18 janvier 2023. Vol. 613, n° 7945, pp. 620-621. DOI 10.1038/d41586-023-00107-z.

Subscribe to Open Community of Practice, [sans date]. *Subscribe to Open* [en ligne]. [Consulté le 17 juillet 2025]. Disponible à l'adresse : <https://subscribetoopencommunity.org/>

UNITED NATIONS EDUCATIONAL, SCIENTIFIC AND CULTURAL ORGANIZATION, 2021. *UNESCO Science Report 2021: The Race Against Time for Smarter Development* [en ligne]. United Nations. [Consulté le 26 janvier 2026]. Disponible à l'adresse : <https://www.un-ilibrary.org/content/books/9789210058575>

VAN NOORDEN, Richard, 2023. More than 10,000 research papers were retracted in 2023 — a new record. *Nature*. 12 décembre 2023. Vol. 624, n° 7992, pp. 479-481. DOI 10.1038/d41586-023-03974-8.

VIAL, Stéphane, 2025. La mort de l'auteur n'aura pas lieu à l'ère de l'IA. *Le Devoir* [en ligne]. 25 août 2025. [Consulté le 31 août 2025]. Disponible à l'adresse : <https://www.ledevoir.com/opinion/idees/911919/mort-auteur-aura-pas-lieu-ere-ia>

YOON, JungWon, KU, Halin et CHUNG, EunKyung, 2024. The road to sustainability: Examining key drivers in open access diamond journal publishing. *Learned Publishing*. 2024. Vol. 37, n° 3, pp. e1611. DOI 10.1002/leap.1611.

ZHOU, Hong et SOULLIÈRE, Marie, 2025. From Detection to Disclosure — Key Takeaways on AI Ethics from COPE's Forum. *The Scholarly Kitchen* [en ligne]. 25 août 2025. [Consulté le 2 octobre 2025]. Disponible à l'adresse : <https://scholarlykitchen.sspnet.org/2025/08/25/from-detection-to-disclosure-key-takeaways-on-ai-ethics-from-copes-forum/>

MISSIONS DES BIBLIOTHÈQUES

2024 : les structures documentaires, piliers de la communauté universitaire et de recherche, 2025. ESGBU.

ADBU, [sans date]. Agir pour l'intégrité scientifique : quels rôles pour les bibliothèques et structures documentaires de l'enseignement supérieur et de la recherche ? [en ligne]. [Consulté le 6 mai 2025]. Disponible à l'adresse : <https://adbu.fr/actualites/agir-pour-lintegrite-scientifique-quels-roles-pour-les-bibliotheques-et-structures-documentaires-de-lenseignement-superieur-et-de-la-recherche>

ADBU, [sans date]. *Ouvrir la science à l'heure de l'intelligence artificielle : pourquoi et comment ?* [en ligne]. [Consulté le 13 janvier 2026 b]. Disponible à l'adresse : <https://adbu.fr/actualites/ouvrir-la-science-a-lheure-de-lintelligence-artificielle-pourquoi-et-comment>

ADBU, 2014. *Evolution des budgets d'acquisition des bibliothèques universitaires* [en ligne]. 2014. [Consulté le 9 janvier 2026]. Disponible à l'adresse : https://adbu.fr/wp-content/uploads/2014/03/Enque%CC%82te_ADBU_2014.pdf

ALI, Muhammad Yousuf et RICHARDSON, Joanna, 2025. AI literacy guidelines and policies for academic libraries: A scoping review. *IFLA Journal*. 13 mars 2025. pp. 03400352251321192. DOI 10.1177/03400352251321192.

ASSOCIATION OF RESEARCH LIBRARIES, 2024. *Research Libraries Guiding Principles for Artificial Intelligence* [en ligne]. Association of Research Libraries. [Consulté le 30 avril 2025]. Disponible à l'adresse : <https://www.arl.org/resources/research-libraries-guiding-principles-for-artificial-intelligence/>

BALNAVES, Edmund, BULTRINI, Leda, COX, Andrew et UZWYSHYN, Raymond (éd.), 2024. *New Horizons in Artificial Intelligence in Libraries* [en ligne]. De Gruyter. [Consulté le 6 juin 2025]. ISBN 978-3-11-133643-5. Disponible à l'adresse : <https://www.degruyterbrill.com/document/doi/10.1515/9783111336435/html>

BAUDRY, Julien, 2025. Enjeux juridiques de l'IA à l'Université : comment la loi encadre les usages scientifiques et documentaires. *Les RDV de l'AMUE* [en ligne]. 15 octobre 2025. [Consulté le 22 janvier 2026]. Disponible à l'adresse : https://www.amue.fr/fileadmin/amue/RDV_Amue/2025/J_BAUDRY_Presentation_IA_-_AMUE_Aspects_juridiques_v2.pdf

BELLIER, Luc, CLAVAUD, Florence et COURTIN, Antoine, 2022. « Notre objectif est de rassembler les professionnels francophones autour de l'ia ». *Arabesques*. 14 octobre 2022. N° 107, pp. 21. DOI 10.35562/arabesques.3093.

BENSAMOUN, Alexandra et FEIRRERA, Lionel, 2024. *Rapport de mission relative à la mise en œuvre du règlement européen établissant des règles harmonisées sur l'intelligence artificielle*. Conseil supérieur de la propriété littéraire et artistique.

BIZOS, Isabelle, 2020. Big deals et open access : quelle stratégie numérique pour les bibliothèques universitaires ? 2020.

BLITZ-FRAYRET, Céline, JALIGOT, Estelle et ORSINI, Colline, 2024. *Note de synthèse - Usages des outils d'Intelligence Artificielle générative dans le domaine de la recherche* [en ligne]. [Consulté le 4 avril 2025]. Disponible à l'adresse : <https://agritrop.cirad.fr/609884/1/240226%20-%20Note%20de%20synthese%20IA%20generatives%20v1.pdf>

BOUCHARD, Aline, 2025. *#WorkInProgress : IA générative et outils de recherche de littérature académique* [en ligne]. février 2025. URFISTinfo. [Consulté le 20 mai 2025]. Disponible à l'adresse : <https://hal.science/hal-04960003>

BROUDOUX, Évelyne, CHARTRON, Ghislaine et EPRON, Benoît, 2025. *Information et intelligence artificielle: opportunités et risques*. Louvain-la-Neuve : De Boeck supérieur. Information & stratégie.

CALENGE, Bertrand, 2015. Permanence bibliothécaire. *Carnet de notes* [en ligne]. 21 décembre 2015. [Consulté le 23 janvier 2026]. Disponible à l'adresse : <https://bccn.wordpress.com/>

COMMISSION RECHERCHE ET DOCUMENTATION ADBU, [sans date]. *L'IA dès aujourd'hui : Des idées pour l'adaptation rapide de nos services aux communautés de recherche – ADBU* [en ligne]. [Consulté le 10 juin 2025]. Disponible à l'adresse : <https://adb.fr/actualites/lia-des-aujourd'hui-des-idees-pour-ladaptation-rapide-de-nos-services-aux-communautes-de-recherche>

COUPERIN, 2024. *L'activité de Couperin en faveur de la transparence des coûts de l'édition scientifique*. 2024. LU : Publications Office.

COX, Andrew, 2023. How artificial intelligence might change academic library work: Applying the competencies literature and the theory of the professions. *Journal of the Association for Information Science and Technology*. mars 2023. Vol. 74, n° 3, pp. 367-380. DOI 10.1002/asi.24635.

COX, Andrew, 2024. Developing a library strategic response to Artificial Intelligence. *eLucidate* [en ligne]. 19 septembre 2024. Vol. 16, n° 4. [Consulté le 5 juin 2025]. DOI 10.29173/elucidate847. Disponible à l'adresse : <https://elucidate-ukeig.org.uk/index.php/elucidate/article/view/847>

COX, Andrew M. et MAZUMDAR, Suvodeep, 2024. Defining artificial intelligence for librarians. *Journal of Librarianship and Information Science*. 1 juin 2024. Vol. 56, n° 2, pp. 330-340. DOI 10.1177/09610006221142029.

DELHAYE, Marc et PERRIN, Sébastien, 2025. Bibliothèque de recherche et IA : vers de nouveaux espaces documentaires ? [en ligne]. 2 juillet 2025. [Consulté le 17 juillet 2025]. Disponible à l'adresse : <https://bbf.enssib.fr/consulter/bbf-2025-00-0000-013>

FAVEL-KAPOIAN, Valentine et KOVACS, Susan, 2025. Accompagner les usagers vers l'IA : défis et motivations des bibliothécaires-“explorateurs” de l'enseignement supérieur. In : *Information et intelligence artificielle* [en ligne]. De Boeck Supérieur. pp. 225-237. [Consulté le 9 février 2026]. Disponible à l'adresse : <https://shs.cairn.info/information-et-intelligence-artificielle--9782807369870-page-225>

FERET, Romain, 2025. Accompagner le développement des services d'appui à la recherche dans les SCD : l'exemple du CRFCB Média Normandie. *Bulletin des Bibliothèques de France* [en ligne]. mars 2025. [Consulté le 26 mai 2025]. Disponible à l'adresse : <https://hal.science/hal-04935704>

GARNIER, Mathilde, CUSSAC, Mathilde, DELAUNE, Camille, ESPY, Camille et PAPON-VIDAL, Laure, 2024. The Pivotal Role of Libraries in Sustainable AI Development. *The Canadian Journal of Information and Library Science*. 1 octobre 2024. Vol. 47, n° 2, pp. 137-149. DOI 10.5206/cjils-rcsib.v47i2.17699.

GEOFFROY, Géraldine, 2025. Une nouvelle interface entre les bibliothèques et l'IA : le Model Context Protocol (MCP). *IA & Bibliothèques* [en ligne]. 11 juillet 2025. [Consulté le 1 août 2025]. Disponible à l'adresse : <https://iaetbibliotheques.fr/2025/07/une-nouvelle-interface-entre-les-bibliotheques-et-lia-le-model-context-protocol-mcp>

GTIA *Feuille de route 2025-2026*, [sans date]. [en ligne]. [couperin.org](https://www.couperin.org). [Consulté le 23 février 2026]. Disponible à l'adresse : https://www.couperin.org/wp-content/uploads/2026/02/GTIA_Couperin_Feuille_de_route_2025-2026.pdf

IFLA FAIFE (COMMITTEE ON FREEDOM OF ACCESS TO INFORMATION AND FREEDOM OF EXPRESSION), 2020. *IFLA statement on Libraries and Artificial intelligence* [en ligne]. [Consulté le 13 juin 2025]. Disponible à l'adresse : https://www.ifla.org/wp-content/uploads/2019/05/assets/faife/ifla_statement_on_libraries_and_artificial_intelligence.pdf

JACOB, Steve, SOUISSI, Seima et MARTINEAU, Charlie, [sans date]. *IA et métiers de la gestion documentaire* [en ligne]. [Consulté le 6 mai 2025]. Disponible à l'adresse : https://bibliopiaf.ebsi.umontreal.ca/bibliographie/8WWN5HAY/download/SDIBUXKJ/ia_et_metiers_de_la_gestion_documentaire.pdf

JACQUEMIN, Bernard et CHAUDIRON, Stéphane, 2024. Les métiers de l'information de demain : Cinq défis à relever. *I2D - Information, données & documents*. 16 juillet 2024. Vol. 1, n° 1, pp. 107-110. DOI 10.3917/i2d.241.0107.

JEANGIRARD, Éric, 2022. L'utilisation de l'apprentissage automatique dans le Baromètre de la science ouverte : une façon de réconcilier bibliométrie et science ouverte ? *Arabesques*. 14 octobre 2022. N° 107, pp. 10-11. DOI 10.35562/arabesques.3084.

KALBANDE, Dattatraya, HEMKE, Digambar et MOTEWAR, Nandkishor, 2025. Artificial intelligence and the five laws: a new vision for library science. *Library Hi Tech News*. 6 mai 2025. Vol. 42, n° 4, pp. 1-3. DOI 10.1108/LHTN-01-2025-0005.

Le Plan national pour la science ouverte 2021-2024 : vers une généralisation de la science ouverte en France, [sans date]. *enseignementsup-recherche.gouv.fr* [en ligne]. [Consulté le 9 janvier 2026]. Disponible à l'adresse : <https://www.enseignementsup-recherche.gouv.fr/fr/le-plan-national-pour-la-science-ouverte-2021-2024-vers-une-generalisation-de-la-science-ouverte-en-48525>

MANGE, Valérie-Anne et TOUSSAINT, Claire, 2024. L'intelligence artificielle : quelles opportunités pour les bibliothèques ? [en ligne]. 12 septembre 2024. [Consulté le 4 avril 2025]. Disponible à l'adresse : https://bbf.enssib.fr/tour-d-horizon/l-intelligence-artificielle-queelles-opportunités-pour-les-bibliothèques_72609

MONTREUIL, Vincent et VOLKWEIN, Alice, [sans date]. 24-25 016B mai 2025 : *L'intelligence artificielle dans les établissements scolaires, sur le plan administratif et pédagogique*. Inspection générale de l'éducation, du sport et de la recherche.

MOUFFLET, Jean-François, 2022. L'intelligence artificielle au service du traitement des archives. *Arabesques*. 14 octobre 2022. N° 107, pp. 14-15. DOI 10.35562/arabesques.3088.

NICOLAS, Yann, 2022. Acclimatation de l'IA à l'Abes : la période des semis. *Arabesques*. 14 octobre 2022. N° 107, pp. 16-17. DOI 10.35562/arabesques.3087.

NOGUEIRA, Leticia Antunes, MOLTUBAKK, Stine Thordarson, FAGERVIK, Andreas et LANGFELDT, Inga Buset, 2025. Cutting through the noise: Assessing tools that employ artificial intelligence. *IFLA Journal*. 1 octobre 2025. Vol. 51, n° 3, pp. 735-749. DOI 10.1177/03400352241304121.

OECD, 2025. *Intellectual property issues in artificial intelligence trained on scraped data* [en ligne]. OECD Artificial Intelligence Papers. 33. [Consulté le 30 octobre 2025]. OECD Artificial Intelligence Papers. Disponible à l'adresse : https://www.oecd.org/en/publications/intellectual-property-issues-in-artificial-intelligence-trained-on-scraped-data_d5241a23-en.html

PASCAL, Frédéric, TADDEI, François, DE FALCO, Marc et GALLIÉ, Emilie-Pauline, 2025. *IA et Enseignement Supérieur : Formation, Structuration et Appropriation par la Société*. Ministère chargé de l'enseignement supérieur et de la recherche.

Printemps Couperin - Sciencesconf.org, [sans date]. [en ligne]. [Consulté le 11 février 2026]. Disponible à l'adresse : <https://printempscoup.sciencesconf.org/resource/page/id/1>

RABOT, Cécile, 2015. Bibliothécaire, un « métier modeste » dans une institution marginalisée. In : QUIJOUX, Maxime (éd.), *Bourdieu et le travail* [en ligne]. Rennes : Presses universitaires de Rennes. pp. 211-227. Le sens social.

[Consulté le 21 février 2026]. ISBN 978-2-7535-6327-8. Disponible à l'adresse : <https://books.openedition.org/pur/69743>

TARDIF, Laurent, CHALLAH, Rana, SERREAU, Matthieu, MESSINA, Virginie, LÉBOUCHER, Caroline, LAFONTAINE, Antoine, KNEFATI, Anas, BOUILLARD, Noé, DELALANDE, Pascaline, MELAYERS, Lucie, CHUSSEAU, Elsa, CAROFF, Charleyne, WONG HEE KAM, Olivier et GAILLAT, Thomas, 2022. Les technologies numériques et la transformation pédagogique. In : *Transformations pédagogique et numérique dans l'enseignement supérieur. Quatre années pour changer les pratiques*. [en ligne]. Presses des mines. [Consulté le 3 juin 2025]. Disponible à l'adresse : <https://hal.science/hal-03670138>

TAY, Aaron, 2025. Testing AI Academic Search Engines - What to find out and how to test (2). *Musings about librarianship* [en ligne]. mai 2025. [Consulté le 31 juillet 2025]. Disponible à l'adresse : <https://musingsaboutlibrarianship.blogspot.com/2025/05/testing-ai-academic-search-engines-what.html>

The Modern Librarian's Role in Supporting AI-Driven Research, 2025. *SciSpace Resources* [en ligne]. [Consulté le 10 février 2026]. Disponible à l'adresse : <https://scispace.com/resources/university-librarians/the-modern-librarians-role-in-supporting-ai-driven-research/>

ZELLER, Arnaud et CHEVRY PÉBAYLE, Emmanuelle, 2024. Usage des Intelligences artificielles génératives à l'université : regards croisés entre usagers et professionnels des bibliothèques universitaires. *Revue COSSI : communication, organisation, société du savoir et information* [en ligne]. 2024. N° 13. [Consulté le 26 mai 2025]. DOI 10.34745/numerev_1962. Disponible à l'adresse : <https://hal.science/hal-04953684>

CHARTES ET GUIDES

AMI DEMOES, 2025. *IA générative & ESR : propositions de charte & guides* [en ligne]. 23 juin 2025. [Consulté le 31 août 2025]. Disponible à l'adresse : <https://www.demoes-pro3.fr/wp-content/uploads/2025/06/Charte-et-bonnes-pratiques-IA-23-juin-2025.pdf>

EUROPEAN COMMISSION, 2025. *Living guidelines on the responsible use of generative AI in Research* [en ligne]. ERA Forum Stakeholder's document. avril 2025. [Consulté le 6 janvier 2026]. Disponible à l'adresse : https://research-and-innovation.ec.europa.eu/document/download/2b6cf7e5-36ac-41cb-aab5-0d32050143dc_en?filename=ec_rtd_ai-guidelines.pdf

Guidance for generative AI in education and research, 2023. [en ligne]. UNESCO. [Consulté le 14 février 2026]. ISBN 978-92-3-100612-8. Disponible à l'adresse : <https://unesdoc.unesco.org/ark:/48223/pf0000386693?locale=en>

ISM-IAE DE VERSAILLES ST-QUENTIN-EN-YVELINES, 2025. *Charte de l'usage de l'intelligence artificielle dans les travaux universitaires*. juin 2025.

UNIVERSITÉ DE LILLE, [sans date]. *Charte des bonnes pratiques d'utilisation des intelligences artificielles génératives dans les travaux pédagogiques*.

UNIVERSITÉ DE RENNES, [sans date]. *Charte pour la science ouverte*. [en ligne]. [Consulté le 6 mai 2025]. Disponible à l'adresse : <https://ged.univ-rennes1.fr/nuxeo/site/esupversions/d2c2533b-365d-49c4-b6fc-58f34ee4bd53?inline>

UNIVERSITÉ DE TOULOUSE, 2023. *Charte du bon usage des IA génératives à l'UT* [en ligne]. 1 mai 2023. [Consulté le 13 février 2026]. Disponible à l'adresse : <https://edintegrity.biomedcentral.com/articles/10.1007/s40979-023-00133-4>

UNIVERSITÉ D'ORLEANS, 2024. *Charte sur l'utilisation de l'Intelligence Artificielle* [en ligne]. 18 octobre 2024. [Consulté le 23 juillet 2025]. Disponible à l'adresse : <https://www.enssib.fr/bibliotheque-numerique/documents/73003-charte-sur-l-utilisation-de-l-intelligence-artificielle.pdf>

LOIS ET RÈGLEMENTS

ALLEA, 2023. *Code de conduite européen pour l'intégrité en recherche* [en ligne]. juin 2023. [Consulté le 13 février 2026]. Disponible à l'adresse : <https://allea.org/wp-content/uploads/2024/12/Code-de-conduite-europeen-pour-lintegrite-en-recherche.pdf>

Ordonnance n° 2021-1518 du 24 novembre 2021 complétant la transposition de la directive 2019/790 du Parlement européen et du Conseil du 17 avril 2019 sur le droit d'auteur et les droits voisins dans le marché unique numérique et modifiant les directives 96/9/CE et 2001/29/CE - Légifrance, [sans date]. [en ligne]. [Consulté le 28 juillet 2025]. Disponible à l'adresse : <https://www.legifrance.gouv.fr/jorf/id/JORFTEXT000044362034>

Règlement (UE) 2024/1689 du Parlement européen et du Conseil du 13 juin 2024 établissant des règles harmonisées concernant l'intelligence artificielle et modifiant les règlements (CE) n° 300/2008, (UE) n° 167/2013, (UE) n° 168/2013, (UE) 2018/858, (UE) 2018/1139 et (UE) 2019/2144 et les directives 2014/90/UE, (UE) 2016/797 et (UE) 2020/1828 (règlement sur l'intelligence artificielle) (Texte présentant de l'intérêt pour l'EEE), 2024. [en ligne]. [Consulté le 28 juillet 2025]. Disponible à l'adresse : <http://data.europa.eu/eli/reg/2024/1689/oj/fra>

ANNEXES

Table des annexes

LISTE DES PERSONNES INTERVIEWÉES DANS LE CADRE D'ENTRETIENS QUALITATIFS.....	96
TRAME DE QUESTIONS – ENTRETIENS SEMI-DIRECTIFS.....	98
EXEMPLE DE CLAUSE TYPE PROVISOIRE FOURNIE AUX.....	100
NÉGOCIATEURS COUPERIN, JUIN 2024.....	100
MODÈLE DE RÉSUMÉ DU CONTENU UTILISÉ POUR ENTRAÎNER LE MODÈLE D'IA À USAGE GÉNÉRAL.....	101

LISTE DES PERSONNES INTERVIEWÉES DANS LE CADRE D'ENTRETIENS QUALITATIFS

Lucie Albaret, Cellule Data – Université Grenoble Alpes

Julien Baudry, Responsable du pôle Collections remarquables et patrimoine scientifique – SCD Université Bordeaux Montaigne

Damien Belvèze, Coordination des formations doctorales – SCD ; Membre du groupe ARDEL et de l'Atelier de la Donnée ARDOISE ; Accompagnement des chercheurs dans la gestion de leurs données de recherche et de leurs codes source – Université de Rennes

Sonja Böttger, ITRF, Chargée du portail HAL et des ressources électroniques – SCD Université de Perpignan

Aurélie Clédat, Chargée de mission Science ouverte & données de la recherche – SCD Université de Perpignan

Nathalie Clément, Sales & Subscriptions Manager – EDP Sciences

Renaud Delemontez Sage, Responsable du pôle Appui à la recherche et à la science ouverte – SCD Université Paris Dauphine-PSL

Isabelle Filliâtre, Administratrice fonctionnelle et technique des systèmes de découverte, d'accès et de signalement des ressources numériques – Pôle Système d'information documentaire, Direction du numérique – Université Paris Dauphine-PSL

Anne-Laure Fischer, Responsable du Département Science, Culture et Numérique – SCD Université de Montpellier

Michel Fraysse, Directeur du Département des services à la recherche – Bibliothèques de l'Université Toulouse Capitole

Madeleine Geroudet, Responsable du Département Services à la recherche et aux chercheurs – SCD Université de Lille

Bernard Jacquemin, Maître de conférences en sciences de l'information – Enssib

Selina La Barbera, Publishing Director Journals – EDP Sciences

Arnaud Laborderie, Chef de projet Gallica, chargé de l'exploitation des données pour la recherche – BnF

Emilie Liard, Co-responsable du pôle Collections – Université de Rennes

François Magnan, Documentaliste et formateur

Jean-Philippe Moreux, Chargé de mission IA – BnF

Thomas Parisot, Directeur général adjoint – Cairn.info

Sébastien Perrin, Directeur des bibliothèques – Mines Paris-PSL

Raluca Pierrot, Directrice projet réinformatisation – ABES

Adeline Regé, Responsable du Pôle Appui à la diffusion de la recherche, Université de Strasbourg ; Chargée de mission Open Access auprès du VP Recherche ; Co-responsable du Département des Négociations documentaires – Consortium Couperin

TRAME DE QUESTIONS – ENTRETIENS SEMI-DIRECTIFS

ÉVOLUTIONS MÉTIER DU BIBLIOTHÉCAIRE

- Quelles sont vos principales missions aujourd'hui ?
- Utilisez-vous des IA génératives dans le cadre de votre travail ? Si oui, lesquelles ?
- Depuis l'émergence des IA génératives, avez-vous observé des changements dans certaines de vos pratiques professionnelles ? Lesquelles et comment ?
- Identifiez-vous de nouvelles compétences métier qu'il faudrait développer, notamment en ce qui concerne la documentation électronique ?
- Quels blocages ou interrogations rencontrez-vous dans l'évolution des missions touchant à la documentation électronique ?

IA GÉNÉRATIVE ET EXPERTISE DES BIBLIOTHÉCAIRES

- L'IA générative change-t-elle la manière de travailler avec les collègues de l'université ?
- Le bibliothécaire joue-t-il un rôle particulier dans le développement des projets IA générative ? Si oui, en quoi ses compétences métier sont-elles utiles ?
- Y a-t-il des domaines dans lesquels l'expertise des bibliothécaires peut être particulièrement intéressante ? (Exemples : traitement des métadonnées, respect des droits d'auteur, open access)

FORMATION ET ACCOMPAGNEMENT DES USAGERS

- Avez-vous adapté vos services d'accompagnement ou de formation des usagers à l'arrivée de ces nouvelles technologies ? Comment ?
- Observez-vous des modifications dans les habitudes de recherche ou de consultation des usagers ?
- Quel regard portez-vous sur les outils de recherche utilisant l'IA, tels que les outils de cartographie bibliographique (Research Rabbit) ou de recension bibliographique (SciSpace, Consensus, Scite) ?
- Observez-vous des différences dans les besoins de formation entre étudiants et chercheurs ? En fonction des disciplines ? Quelles compétences deviennent nécessaires pour un bibliothécaire, un doctorant, un chercheur ?
- Y a-t-il des différences selon les disciplines (STM, SHS, LLH) ?

IMPACT SUR LES BASES DE DONNÉES ET RESSOURCES DOCUMENTAIRES

- Avez-vous connaissance d'expérimentations sur l'indexation avec IA, la recherche dans des textes bruts avec IA, ou le text mining dans vos outils de documentation électronique ?
- Avez-vous participé à de tels projets ? Quelle est votre opinion sur leur qualité et la fiabilité des algorithmes utilisés ?
- Quels autres projets ou perspectives pourraient impacter les missions des bibliothécaires ?

- Quelle est, actuellement, la place de l'IA dans les négociations avec les fournisseurs et éditeurs ? Les dispositions légales pour le TDM (text and data mining) vous semblent-elles suffisantes aujourd'hui ?

IMPACT DE L'IA GÉNÉRATIVE SUR LA PUBLICATION ET LA DOCUMENTATION SCIENTIFIQUES

- Observez-vous des changements dans les tendances et habitudes de publication des chercheurs du fait de l'IA générative ? Par exemple dans les choix de droits d'auteur et d'autorisation de diffusion ? Des spécificités disciplinaires ?
- Les modèles de diffusion en open access sont-ils, selon vous, remis en question par l'IA ?
- L'IA pose-t-elle des questions en termes de fiabilité des sources et d'intégrité scientifique ?
- En quoi le travail du bibliothécaire s'articule-t-il avec ces questionnements autour de l'intégrité scientifique ? Le droit d'auteur est-il suffisamment protégé ? L'utilisation de l'IA pour la documentation scientifique contribue-t-elle à l'ouverture de la science ?

PRÉCONISATIONS ET ÉVOLUTIONS INSTITUTIONNELLES

- Votre établissement a-t-il pris des positions officielles concernant l'usage des IA ? Lesquelles ? (Charte, règlement des études, autre)
- Quel impact les aspects réglementaires (RGPD, propriété intellectuelle, AI Act...) ont-ils sur vos missions et votre vision des outils d'IA générative ?

PROSPECTIVE

- Si vous deviez conseiller un jeune collègue bibliothécaire travaillant en documentation électronique, que lui diriez-vous ? Quelles compétences lui suggèreriez-vous de développer en priorité ?
- Quelles évolutions souhaiteriez-vous voir dans la formation des bibliothécaires ?
- Y a-t-il des aspects importants que nous n'avons pas abordés et que vous souhaiteriez mentionner ?

EXEMPLE DE CLAUSE TYPE PROVISOIRE FOURNIE AUX

NÉGOCIATEURS COUPERIN, JUIN 2024

L'usage de l'intelligence artificielle sur les données détenues par XXX ou pour lesquelles elle détient les droits de distribution n'est autorisé qu'à des fins exclusives de recherche.

Par Intelligence artificielle, on entend les techniques et outils visant à imiter les facultés cognitives humaines telles que l'apprentissage, le raisonnement, la résolution de problèmes, la compréhension du langage naturel, la reconnaissance de motifs, et la prise de décisions.

L'utilisation de l'intelligence artificielle à des fins de recherche doit s'effectuer localement ou dans une application hébergée sécurisée.

Il est interdit dans tous les cas de :

- Reproduire et divulguer les contenus générés
- Avoir un usage pouvant entraver la bonne exploitation par XXX
- Donner accès à une société commerciale au contenu de XXX

MODÈLE DE RÉSUMÉ DU CONTENU UTILISÉ POUR ENTRAÎNER LE MODÈLE D'IA À USAGE GÉNÉRAL

Extrait de BENSAMOUN Alexandra et Lionel FEIRRERA, *Rapport de mission relative à la mise en œuvre du règlement européen établissant des règles harmonisées sur l'intelligence artificielle*, Conseil supérieur de la propriété littéraire et artistique, 2024.

IV. Modèle de résumé

Point de contact pour toute demande ou procédure.

S'agit-il de la mise à jour d'un résumé après évolution du modèle ? o / n

Si oui, fournir le résumé du modèle initial.

L'entraînement du modèle repose-t-il en tout ou partie sur un modèle d'IA déjà existant ? o / n

Si oui, lien vers le résumé de ce modèle.

Certains des contenus utilisés font-ils l'objet d'accords ? o / n

Si oui, dans quel secteur ?

Si le partenaire est un organisme public, le mentionner.

Contenus libres de droit*					
Moissonnés depuis internet				Récupérés auprès de tiers	
Jeu de données	Méthode de moissonnage	Nom de domaine moissonné	Taille et type de données (image, son, multimodal, etc.)	Jeu de données	Le cas échéant, lien vers le jeu de données
(liste jeux de données)	(description) (1)	(pour chaque jeu de données, liste des domaines moissonnés)	(pour chaque jeu de données, taille et type de données)	(liste des jeux de données)	(pour chaque jeu de données, lien si applicable)

* uniquement pour les jeux de données constitués exclusivement de données libres de droit, c'est-à-dire relevant du domaine public à l'issue de la protection au titre du droit d'auteur et des droits voisins ou contenus dont l'utilisation a été expressément autorisée par l'ayant droit en licence libre. La mission met en garde sur le périmètre de la licence libre qui peut néanmoins interdire ce type d'usage. A défaut ou en cas de doute sur le licence, considérer que le jeu de données relève de la catégorie « Autres contenus ».

Autres contenus (non livres de droits)											
Moissonnés directement ou par un tiers mandaté depuis internet					Jeux de données récupérés auprès de tiers			Prompts		Données synthétiques générées à partir de données humaines	
Jeu de données	Méthode de moissonnage	Méthodologie pour assurer la conformité avec le droit de l'Union, (notamment clause de réserve de droits, exclusion des sites pirates...)	URLs moissonnées + date de moissonnage	Taille et type de données (image, son, multimodal, etc.)	Jeu de données	Méthodologie pour assurer la conformité avec le droit de l'Union (ex. : clause de réserve de droits...)	Description	Description	Méthodologie pour assurer la conformité avec le droit de l'Union	Jeu de données	Méthodologie pour assurer la conformité avec le droit de l'Union
(liste jeux de données datés)	(description) (1)	(description) (2)	(pour chaque jeu de données, liste des URLs avec date de collecte)	(pour chaque jeu de données, taille et type de données...) (3)	(liste des jeux de données datés)	(description, notamment fondement légal de la collecte de données) (4)	(5)	(préciser si les prompts utilisateurs peuvent être enregistrés et utilisés pour l'entraînement du modèle)	(description) (6)	(liste des jeux de données datés, modèles utilisés)	(description, notamment méthodes permettant d'atténuer le risque de contrefaçon...)

A titre non exhaustif, type de questions qui pourraient être intégrées au modèle de résumé par le Bureau de l'IA :

- (1) Identification des moissonneurs utilisés ? Base légale du moissonnage ? Dates de moissonnage ?
- (2) Fondement légal de la collecte ? Méthode utilisée pour identifier les métadonnées pertinentes ? Méthodes pour éviter la suppression de ces métadonnées au cours de l'entraînement ou la génération de contenu ? Standards et protocoles reconnus ? Méthodes mises en œuvre pour respecter les instructions lisibles par la machine ? Autres méthodes mises en œuvre pour garantir un traitement licite ? Méthodes mises en œuvre pour assurer la mise en

conformité a posteriori si du contenu protégé a été utilisé sans autorisation ? Certification ou contrôle externe mobilisés pour assurer la licéité de l'opération ? Méthodes pour neutraliser les données dont les droits ont été réservés ? Identification des standards reconnus (type ISBN, etc.) ?

- (3) Taille des données d'entraînement ? Taille de chaque type de donnée ?

- (4) Par ex., garantie du fournisseur de données ?

- (5) Identification et taille du jeu de données ? Type de données ? Liste des types d'identifiants uniques (ISBN, DOI, ISAN, etc.) inclus ? Lien vers le jeu de données si applicable ou, à défaut, liste des contenus (par les identifiants disponibles) du jeu de données à date avec suffisamment de détail relatif au contenu pour faciliter aux ayants droit l'exercice de leur droit ? Pour chaque type d'identifiant, pourcentage d'item avec cet identifiant ?

- (6) Méthode pour apprécier si le *prompt* reproduit des données protégées ? Méthode pour neutraliser le cas échéant ces données ?

TABLE DES MATIÈRES

SIGLES ET ABRÉVIATIONS.....	9
INTRODUCTION.....	11
L'ÉCOSYSTÈME DE LA PUBLICATION SCIENTIFIQUE BOULEVERSE PAR LES IA GÉNÉRATIVES.....	15
Contexte général.....	15
<i>Vision d'ensemble.....</i>	<i>15</i>
<i>Le mouvement pour la science ouverte et l'accès ouvert.....</i>	<i>17</i>
<i>Inflation des publications scientifiques.....</i>	<i>19</i>
<i>Remise en question du modèle de la revue et de l'article.....</i>	<i>21</i>
Un écosystème changé par l'IA générative.....	22
<i>Un potentiel d'utilisation dans la totalité des étapes de la publication.....</i>	<i>22</i>
<i>La pertinence de ces usages en question.....</i>	<i>23</i>
<i>Des questionnements de natures multiples.....</i>	<i>25</i>
<i>Du refus à l'encadrement.....</i>	<i>25</i>
<i>De nouvelles opportunités commerciales.....</i>	<i>26</i>
Un écosystème permettant aux IA génératives d'exister, mais fragilisé par les IA génératives.....	28
<i>Les données en libre accès comme source d'entraînement pour les IA génératives.....</i>	<i>28</i>
<i>Une remise en cause du contrat social autour de l'open access.....</i>	<i>30</i>
<i>Tentation de la fermeture et ses limites.....</i>	<i>32</i>
<i>Adaptations.....</i>	<i>33</i>
LES ÉVOLUTIONS TECHNOLOGIQUES DANS LA DOCUMENTATION ÉLECTRONIQUE.....	37
Aux fondements technologiques de l'IA générative.....	37
<i>Des outils statistiques ayant des limites structurelles.....</i>	<i>37</i>
<i>La technologie RAG : réduire les biais par la sélection documentaire.....</i>	<i>39</i>
Un continuum de solutions privées.....	40
<i>Outils d'aides à la rédaction et à la soumission.....</i>	<i>40</i>
<i>Outils de découvrabilité.....</i>	<i>41</i>
<i>Des outils holistiques : de la recherche à l'écriture.....</i>	<i>42</i>
<i>Enjeux critiques des solutions privées.....</i>	<i>43</i>
Des initiatives publiques.....	44
<i>HAL : améliorer le traitement des données.....</i>	<i>44</i>
<i>L'ABES.....</i>	<i>44</i>
<i>La BnF.....</i>	<i>45</i>
<i>ILaaS : une infrastructure mutualisée pour l'enseignement supérieur et la recherche.....</i>	<i>45</i>
Quelles expertises des bibliothécaires pour comprendre ces technologies ?	46
<i>Évaluer l'évolution des pratiques de consultation d'information.....</i>	<i>46</i>
<i>Évaluer méthodiquement la qualité des résultats fournis par l'IA générative</i>	<i>47</i>
<i>Être partie prenante des solutions technologiques.....</i>	<i>49</i>
ENCADREMENTS LÉGAUX, NORMATIFS ET PRATIQUES.....	52
Un cadre normatif en cours de constitution.....	52
<i>La loi.....</i>	<i>52</i>
<i>Préconisations d'acteurs variés sans coordination complète.....</i>	<i>54</i>

<i>Cas concrets : les chartes universitaires.....</i>	<i>56</i>
<i>La place des bibliothèques dans ces réflexions.....</i>	<i>57</i>
Questions de fond : la responsabilité scientifique dans l’usage des IA	
généralité.....	59
<i>Tendances concernant le signalement.....</i>	<i>59</i>
<i>Questions latentes.....</i>	<i>61</i>
QUELLES ÉVOLUTIONS POUR LES PROFESSIONNELS DES	
BIBLIOTHÈQUES ?.....	63
Des compétences pertinentes, une légitimité à construire.....	63
<i>Regard critique et capacité de médiation.....</i>	<i>64</i>
<i>Une reconnaissance et une identité professionnelle fragiles sur ces questions</i>	<i>65</i>
Une profession qui s’adapte pour rester pertinente.....	66
<i>Une variété de positionnements face à la technologie.....</i>	<i>66</i>
<i>Comment se former ?.....</i>	<i>67</i>
Les missions des professionnels des bibliothèques face à l’IA générative....	70
<i>Interactions avec les éditeurs.....</i>	<i>70</i>
<i>Déterminer le périmètre de la politique documentaire.....</i>	<i>71</i>
<i>Accompagner les chercheurs.....</i>	<i>72</i>
<i>Défendre l’accès ouvert.....</i>	<i>73</i>
CONCLUSION.....	75
DÉCLARATION D’USAGE.....	77
BIBLIOGRAPHIE.....	79
IA générative, généralités.....	79
Services d’IA génératives.....	81
Autour des questions d’open access.....	82
Ecosystème de la publication scientifique.....	84
Missions des bibliothèques.....	89
Chartes et guides.....	93
Lois et règlements.....	94
ANNEXES.....	95
TABLE DES MATIÈRES.....	103