

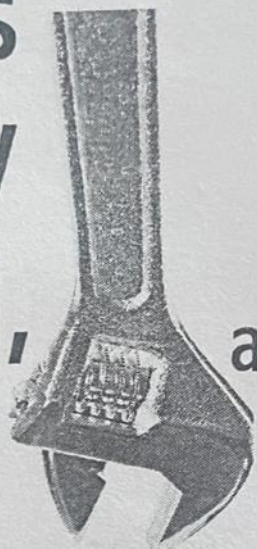
Archives ouvertes et intelligences artificielles

Madeleine Géroudet

Journées CasuHAL 2025

AI is
very
bad, actually

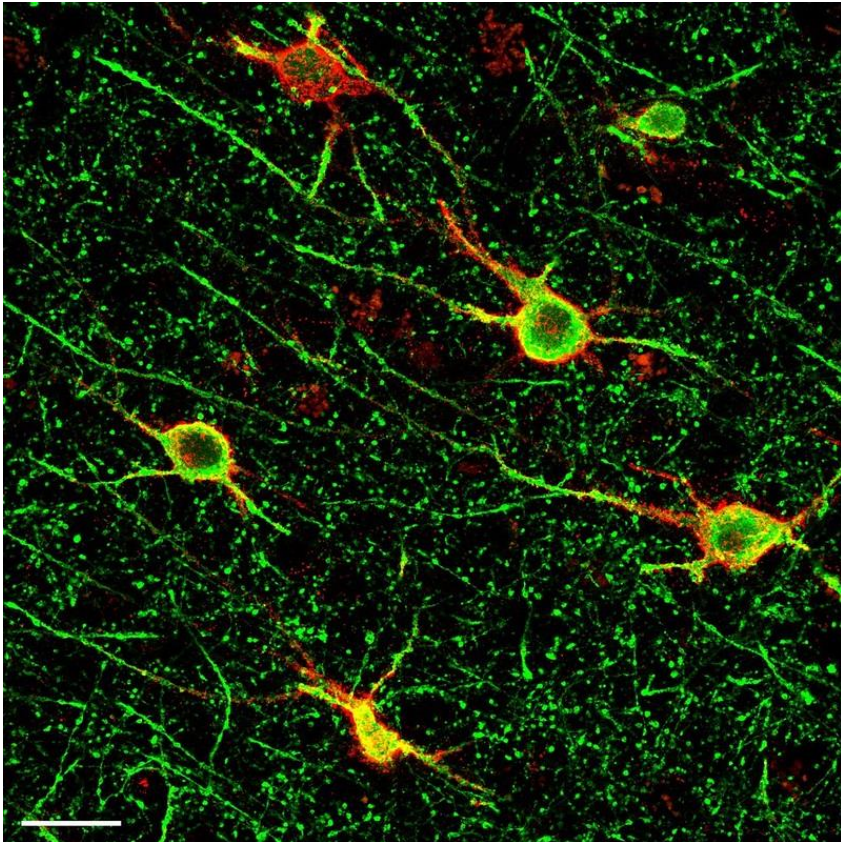
A manifesto



L'IA et moi



Angle d'approche pour aujourd'hui



- **Intelligences artificielles génératives** et ses conséquences pour la communication scientifique et la pratique de la recherche
- Sujet fortement évolutif, avec de multiples inconnues
- Liens et ramifications avec différents enjeux liés à l'écosystème de la recherche

Arnaud Tanti/Inserm.licence CC-BY-NC 4.0
<https://www.flickr.com/photos/inserm/52233529820>

Ramifications



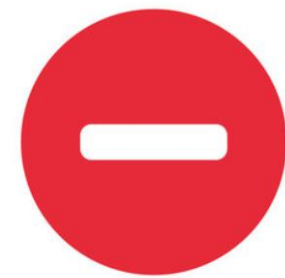
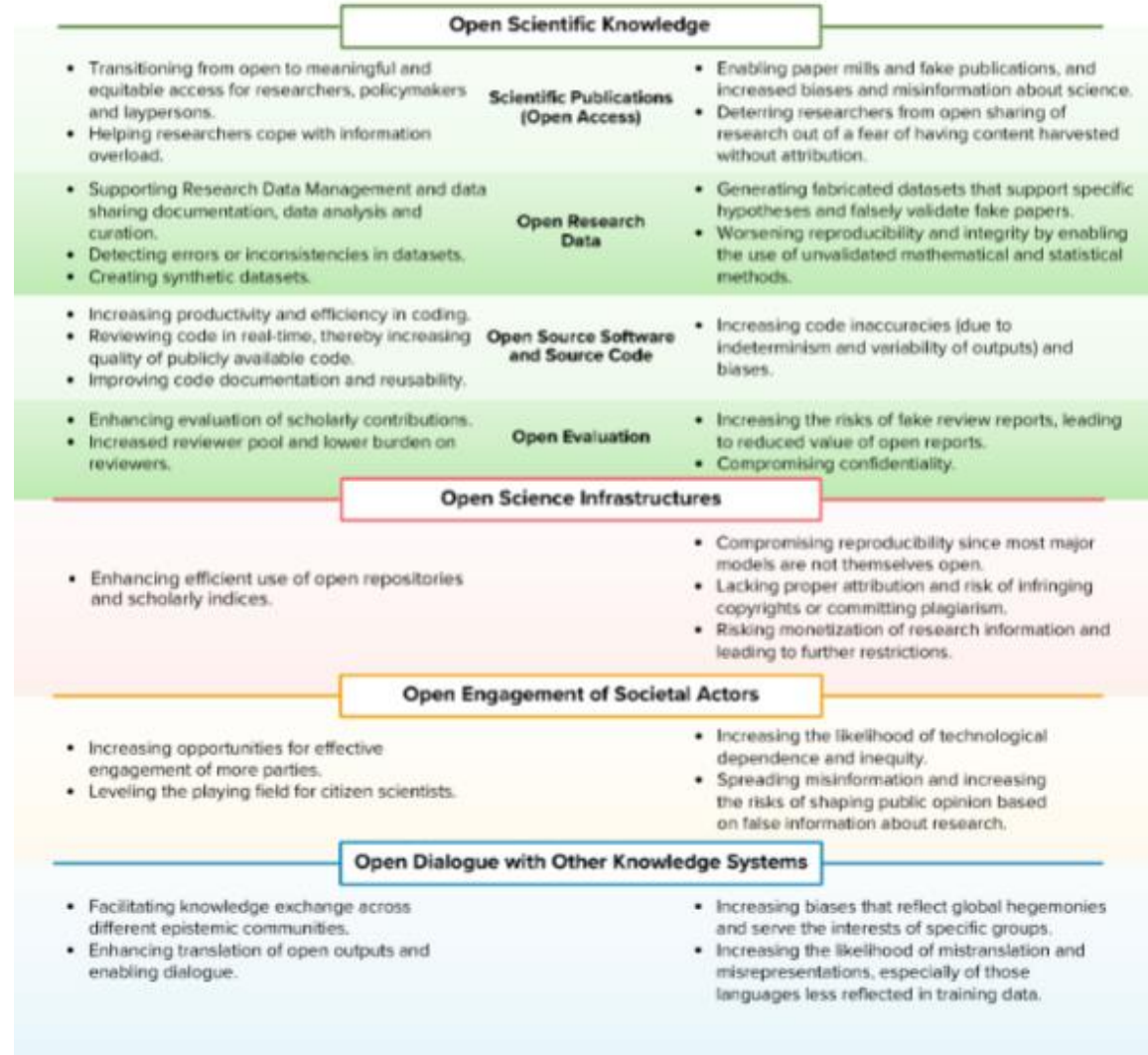
C. Beuret, Aiguillage à Neuchâtel, CC BY-NC-SA 4.0

- Édition et communication scientifique
- Intégrité et éthique
- Respect des droits d'auteurs
- Fouille de texte et de données
- Reproductibilité de la recherche
- Confiance dans la science et esprit critique face à l'information
- Activités de recherche et activités administratives
- Impact sur les métiers
- Sobriété numérique
- ...

Que font les intelligences
artificielles génératives à la
science (ouverte) ?

Possible **Positive** Impacts of GenAI upon Open Science

Possible **Negative** Impacts of GenAI upon Open Science



Trois portes d'entrée



Créé avec Canva

Porte d'entrée n°1 : le droit



Enjeux juridiques



- Règlement relatif à l'intelligence artificielle (RIA ou AI Act) du 13 juin 2024
- Règlement général sur la protection des données (RGPD) du 27 avril 2016
- Directive relative au droit d'auteur et aux droits voisins dans le marché unique numérique (DAMUN) du 17 avril 2019 (dont l'exception TDM pour la recherche)

RIA et droits d'auteurs



Les IA dites générales doivent :

- respecter le droit d'auteur et les droits voisins, et prendre toute mesure visant à s'assurer du respect de ces droits
- Élaborer et rendre disponible un résumé suffisamment détaillé des sources de données utilisées pour entraîner le modèle

« Autrement dit, la liste des ingrédients doit être rendue publique, mais pas la recette. »

Opt-out



POLITIQUE FOUILLE DE TEXTES ET DE DONNEES < TDM-RESERVATION: 1>

[...] s'oppose à toutes opérations de moissonnage et de fouille de textes et de données au sens de l'article L. 122-5-3 du code de la propriété intellectuelle.

Cette opposition couvre l'ensemble du Site et des Contenus auxquels il donne accès.

Toutes opérations de moissonnage et de fouille de textes et de données visant le Site et ses Contenus, y compris par des dispositifs de collecte automatisée de données constituent donc des actes de contrefaçon sauf obtention d'un accord spécifique formellement exprimé de [...].

L'article R. 122-28 du code de la propriété intellectuelle précisant que l'opposition mentionnée au III de l'article L. 122-5-3 peut être exprimée par tout moyen, y compris par le recours à des conditions générales d'utilisation d'un site internet ou d'un service, l'absence de métadonnées associées au Site, répertoires du Site, Contenus du Site est sans incidence sur l'exercice du droit d'opposition exprimé par les présentes conditions générales d'utilisation.

Pour faciliter la lecture de ce droit d'opposition par tout dispositif de collecte automatisée de données, cette opposition est également exprimée ainsi < TDM-RESERVATION: 1>.

<https://www.sne.fr/actu/une-clause-type-pour-sopposer-a-la-fouille-de-textes-et-de-donnees-par-les-intelligences-artificielles/>

« L'exception ou la limitation prévue au paragraphe 1 s'applique à condition que l'utilisation des œuvres et autres objets protégés visés audit paragraphe n'ait pas été expressément réservée par leurs titulaires de droits de manière appropriée, **notamment par des procédés lisibles par machine pour les contenus mis à la disposition du public en ligne.** »

Extrait de l'article 4 (DAMUN, **2019**).

Les IAG questionnent le périmètre même du droit d'auteur

BeauxArts

Agenda Vidéos Expos Insolite À la loupe Reportages Lifestyle **L'ENCYCLO** La Newsletter Conférences Le Magazine La Boutique 🔒 🔍

ACTU

Le succès massif des images façon studio Ghibli générées par ChatGPT relance la question des droits d'auteur face à l'IA

Par **Joséphine Bindé**

Publié le 9 avril 2025 à 16h00, mis à jour le 9 avril 2025 à 16h05

Comment le droit doit-il penser l'usage de données pour l'entraînement de modèles, qui oublie ensuite leurs données sources ?

« Ces données, et notamment les données culturelles, constituent les seuls intrants de la chaîne dont la valeur commerciale, bien qu'évidente, est remise en cause. »



La science ouverte doit-elle être ouverte ou fermée... aux intelligences artificielles génératives ?



Créé avec Canva

- Des craintes exprimées par les communautés de recherche sur le pillage de leurs résultats par les IAG.
- Les IAG ne font que renforcer la logique cumulative de la construction des savoirs scientifiques.
- Est-il préférable de disposer d'IAG avec des contenus scientifiques validés ou avec des contenus de sites web douteux ?

To be in IA or not to be ?

Intelligence artificielle : un accord de partenariat entre « Le Monde » et OpenAI

Cet accord pluriannuel, le premier entre un média français et un acteur majeur de l'IA, permettra à la société de s'appuyer sur le corpus du journal pour établir et fiabiliser les réponses de son outil ChatGPT, moyennant une source significative de revenus supplémentaires.

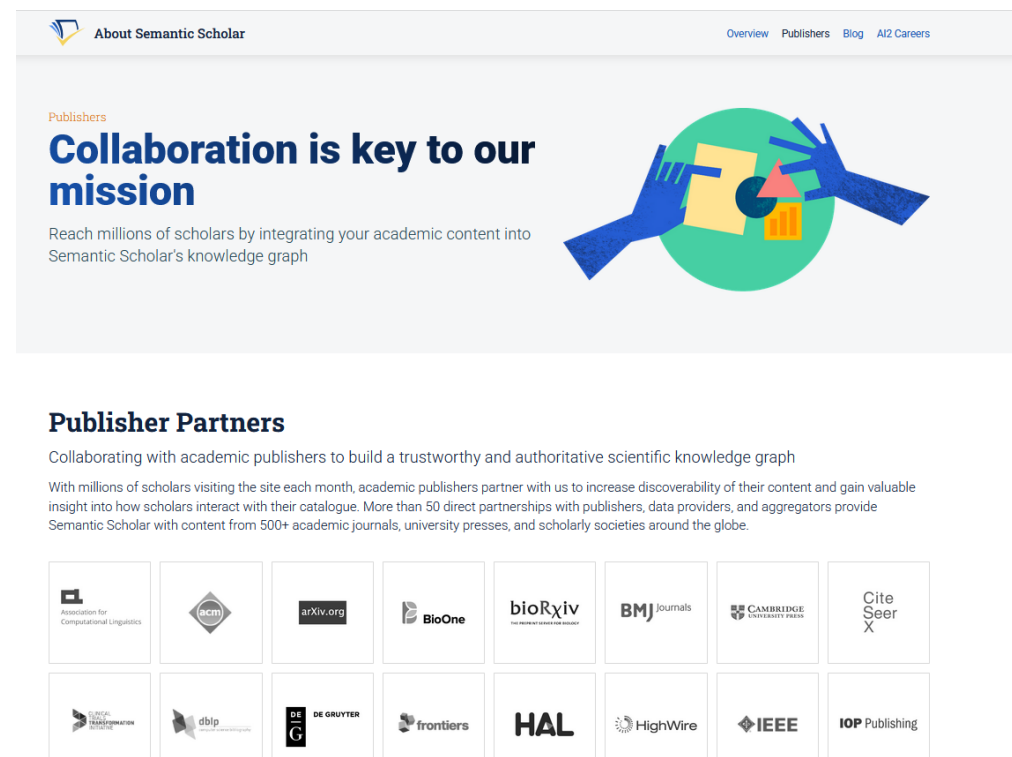
https://www.lemonde.fr/le-monde-et-vous/article/2024/03/13/intelligence-artificielle-un-accord-de-partenariat-entre-le-monde-et-openai_6221836_6065879.html

⚠ Archives ouvertes non intégrées directement (comme HAL) :

Voici quelques autres archives ouvertes reconnues que je ne peux pas interroger automatiquement, mais dont je peux analyser les contenus manuellement si tu m'en donnes les liens :

- HAL (France – multi-disciplinaire) – <https://hal.archives-ouvertes.fr>
- Zenodo (UE/CERN – multi-disciplinaire) – <https://zenodo.org>
- Open Science Framework (OSF) – <https://osf.io>
- E-LIS (sciences de l'information et bibliothéconomie) – <https://eprints.rclis.org>

Scholar GPT, 25 avril 2025



The screenshot shows the Semantic Scholar website. At the top, there's a navigation bar with 'About Semantic Scholar' and links for 'Overview', 'Publishers', 'Blog', and 'AI2 Careers'. The main content area is titled 'Publishers' and features the headline 'Collaboration is key to our mission'. Below this, it states 'Reach millions of scholars by integrating your academic content into Semantic Scholar's knowledge graph'. To the right is an illustration of two hands holding a puzzle piece. Further down, the 'Publisher Partners' section is displayed, listing various academic publishers and their logos in a grid format.

Semantic scholar

« À partir de 2023, la BnF a renforcé sa collaboration avec les pouvoirs publics et l'écosystème IA pour organiser la mise à disposition de données patrimoniales libres de droits dans le cadre de partenariats encadrés. Dès 2025, la BnF ambitionne de proposer une solution d'intérêt général dédiée à la fourniture de jeux de données massifs et spécifiques libres de droits, au service de l'écosystème IA. »

<https://www.bnf.fr/fr/lintelligence-artificielle-un-axe-strategique#bnf-le-tournant-de-lia-g-n-rative-la-valeur-des-collections-comme-donn-es>

Ouvrir aux IAG, mais comment ?



Créé avec Canva

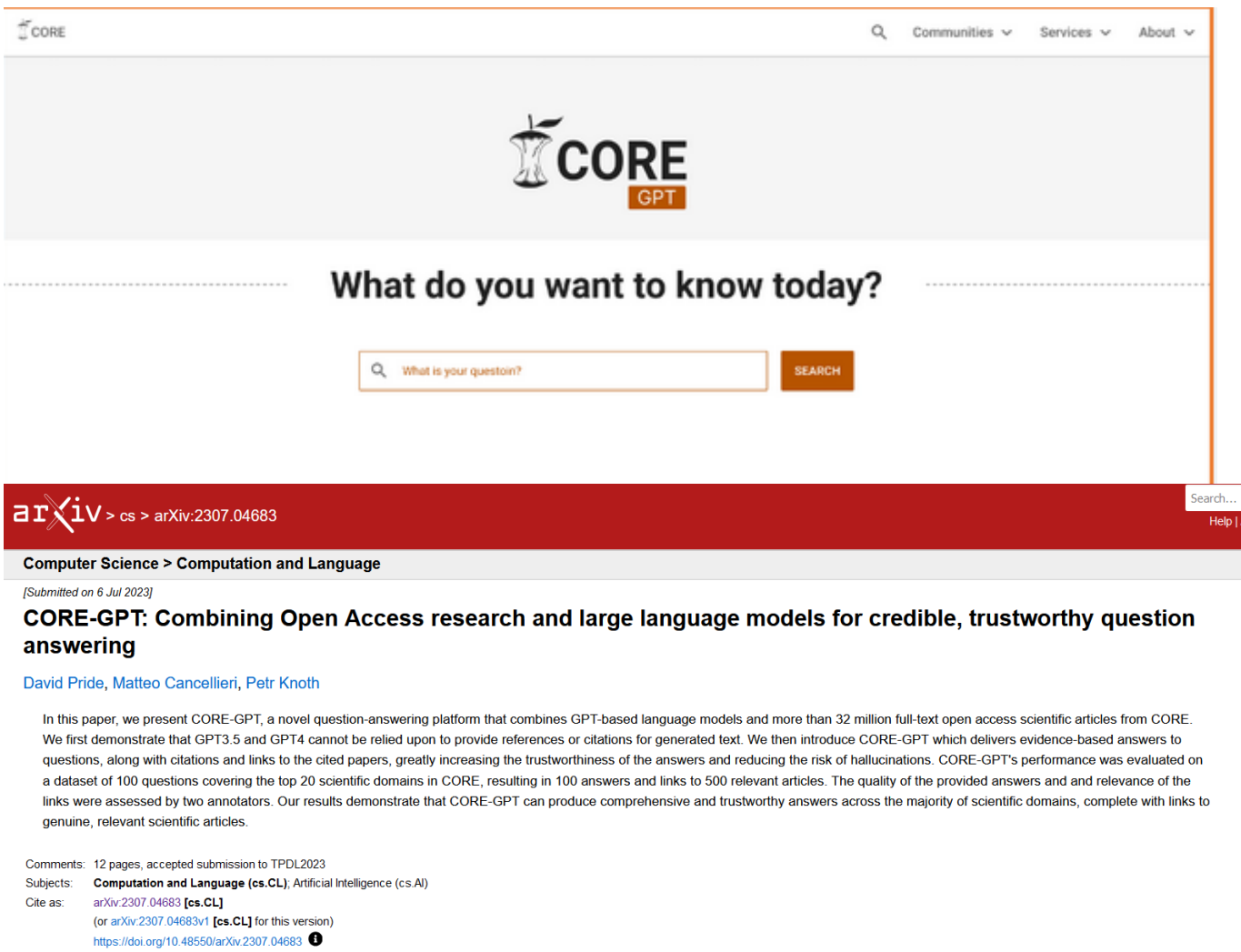
Qui décide d'ouvrir aux IAG ?

Les auteurs ? Les plateformes ? La licence effectivement utilisée ?

Peut-on envisager une ouverture intelligente aux IAG ?

Créer un mécanisme de licence de type Share Alike permettant d'ouvrir uniquement aux IAG ouvertes ?

Des IAG sur des contenus validés



ISTEX

*Le socle de la bibliothèque
scientifique numérique nationale*

AIKO : Programme AI for Publication Knowledge

Inria

Des IAG ouvertes sur des données ouvertes : l'exemple de Pleias

Modèles IA ouverts et moins consommateurs d'énergie

- Un accent mis sur le traitement du PDF
- Travailler sur des données ouvertes

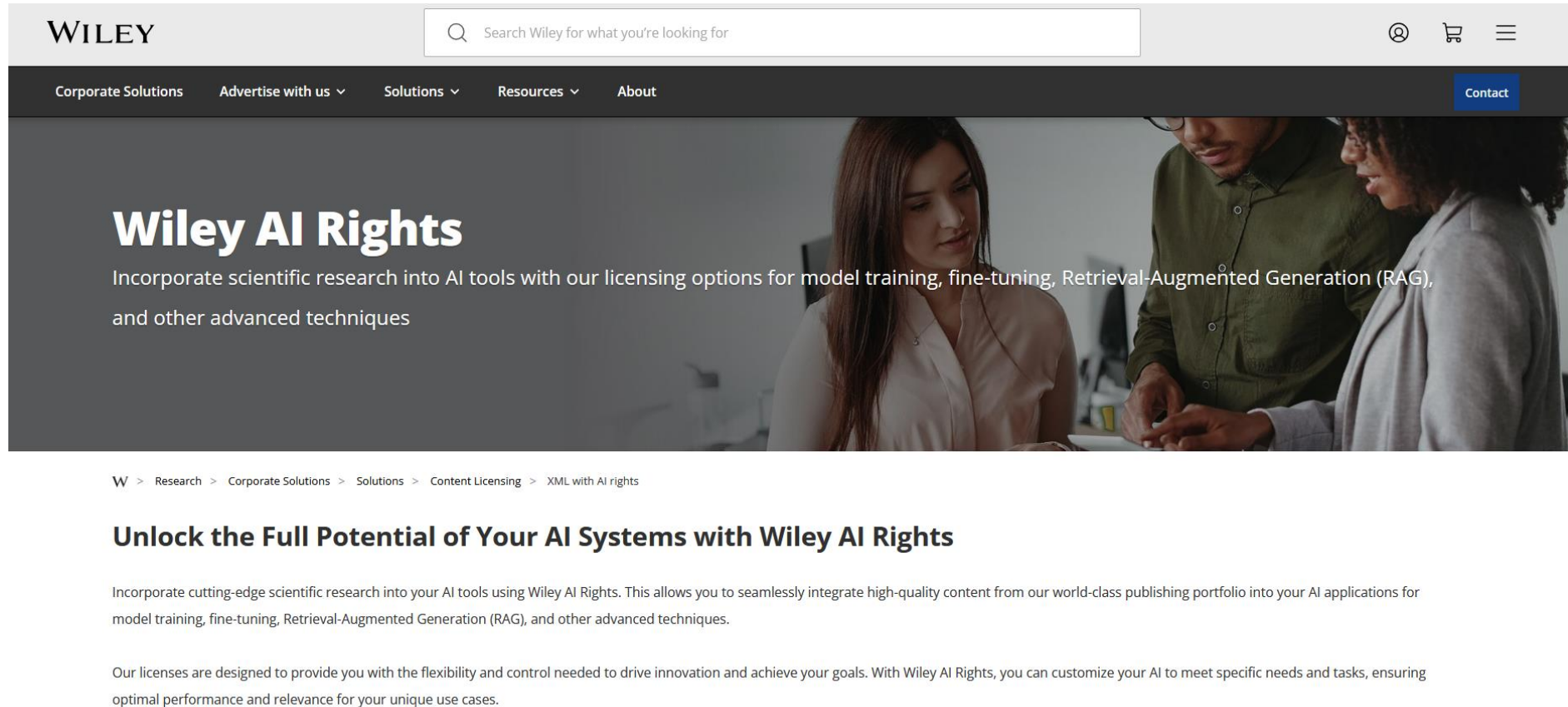
WE DESIGN AND PRETRAIN WORLD'S MOST EFFICIENT
LLMS TO PROVE THAT AI CAN BE SCALED
ON BUDGET - WITH OPEN DATA

About Common Corpus

Common Corpus is made of six carefully curated collections:

- **OpenCulture**: our largest collection at 885,982,490,090 tokens, featuring public domain books, newspapers from cultural heritage repositories and open projects like Wikisource and Gutenberg. We're developing innovative tools of OCR correction based on Pleias Models to correct historical digitization errors, while implementing advanced toxicity filtering to ensure content meets modern ethical standards.
- **OpenGovernment**: 406,581,454,455 tokens of financial and legal documents, including Finance Commons (from sources like SEC and WTO) and Legal Commons (including Europarl and Caselaw Access Project), providing enterprise-grade training data from regulatory bodies and administrative sources.
- **OpenSource**: 283,227,402,898 tokens of high-quality code in open source from GitHub, filtered using ArmoRM to ensure only the top 80% of submissions by quality rating are included.
- **OpenScience**: 281,193,563,789 tokens of academic content from Open Alex and other open science repositories, processed using vision-language models to preserve crucial document structure and formatting.
- **OpenWeb**: 73,217,485,489 tokens from Wikipedia (official releases from the [Wikimedia Foundation](#) on Huggingface), YouTube Commons and Stack-Exchange.
- **Open Semantic**: 67,958,671,827 tokens from Wikidata (official releases from the [Wikimedia Foundation](#) on Huggingface). The data has been reprocessed thanks to support and help of Wikidata and Wikimedia Germany. It includes the transcriptions of all the semantic triplets into natural language statements in over 300 languages.

Ouvrir aux IAG, les communautés de recherche ont-elles vraiment la main ?



The screenshot shows the Wiley AI Rights webpage. At the top is the Wiley logo and a search bar. Below the navigation bar, the main heading is "Wiley AI Rights" with a subtext: "Incorporate scientific research into AI tools with our licensing options for model training, fine-tuning, Retrieval-Augmented Generation (RAG), and other advanced techniques". A background image shows three people in a meeting. Below this is a breadcrumb trail: "W > Research > Corporate Solutions > Solutions > Content Licensing > XML with AI rights". The main heading "Unlock the Full Potential of Your AI Systems with Wiley AI Rights" is followed by two paragraphs of text.

WILEY

Search Wiley for what you're looking for

Corporate Solutions Advertise with us Solutions Resources About Contact

Wiley AI Rights

Incorporate scientific research into AI tools with our licensing options for model training, fine-tuning, Retrieval-Augmented Generation (RAG), and other advanced techniques

W > Research > Corporate Solutions > Solutions > Content Licensing > XML with AI rights

Unlock the Full Potential of Your AI Systems with Wiley AI Rights

Incorporate cutting-edge scientific research into your AI tools using Wiley AI Rights. This allows you to seamlessly integrate high-quality content from our world-class publishing portfolio into your AI applications for model training, fine-tuning, Retrieval-Augmented Generation (RAG), and other advanced techniques.

Our licenses are designed to provide you with the flexibility and control needed to drive innovation and achieve your goals. With Wiley AI Rights, you can customize your AI to meet specific needs and tasks, ensuring optimal performance and relevance for your unique use cases.

<https://www.wiley.com/en-us/business/research/corporate-solutions/solutions/content-licensing/xml-machine-learning>

Les IAG redonnent-elles du poids aux éditeurs scientifiques ?

Scopus AI: Trusted content.
Powered by responsible AI.

Empower your research journey with Scopus AI – your dynamic GenAI-powered research companion.

Navigate through the vast expanse of human knowledge faster with a trusted guide designed to enhance your understanding, enrich your insights, and transform your overall research experience.

<https://www.elsevier.com/products/scopus/scopus-ai>

- Augmentation de la valeur du capital scientifique détenu par les éditeurs
- Valeur ajoutée des services augmentés développés par les éditeurs
- Fragilité des investissements économiques actuellement réalisés par les éditeurs
- Potentiel renforcement d'une recherche à 2 vitesses

Les IAG poussent les éditeurs à la fermeture de leurs contenus



Sciences : « Enrichir Microsoft, Meta ou Google au détriment des éditeurs privés serait une erreur »

Avec l'envol des intelligences artificielles génératives, qui absorbe tous les contenus, la diffusion en « open access » de publications scientifiques, sans droits d'auteur, fragilise l'édition et la recherche, alerte, dans une tribune au « Monde », Vincent Montagne, président du Syndicat national de l'édition.

Publié le 20 avril 2023 à 05h30 | 🕒 Lecture 2 min.

https://www.lemonde.fr/idees/article/2023/04/20/sciences-enrichir-microsoft-meta-ou-google-au-detriment-des-editeurs-privés-serait-une-erreur_6170254_3232.html

- AI clauses will permit the use of AI for any and all legal purposes that support consortia members' core missions of non commercial research, teaching, learning, and equitable access to information.
- Multi-year licenses that include AI clauses will formally establish opportunities for the revision and/or updating of language related to AI during the term of the agreement to reflect legal or technological developments.
- AI clauses will not restrict user actions in ways that are fundamentally unenforceable.
- AI clauses will in no way prevent authorized users from making licensed content fully accessible to other authorized users in any legal manner.
- AI clauses will not be introduced into a negotiation in a way that limits a consortium's opportunity to fully review the clauses in accordance with the consortium's regular review procedure.
- AI clauses will not introduce new liability clauses beyond the scope of what has previously been agreed upon, especially for actions of authorized users.

<https://icolc.net/system/files/ICOLC%20Statement%20on%20AI%20in%20Licensing%203.22.2024.pdf>

Les intelligences artificielles génératives renforcent la valeur des contenus scientifiques

***Une trilogie à suivre sur vos
écrans***

Learning Apocalypse



Jean-Pierre Luminet, CC BY-SA 4.0

https://commons.wikimedia.org/wiki/File:LUMINET_%27S_BLACK_HOLE_DRAWING-1982.jpg

Learning collapse

Hypothèse selon laquelle la présence massive de données synthétiques produites par les intelligences artificielles génératives dans les corpus utilisés pour l'entraînement des modèles pourrait ralentir, voire empêcher l'apprentissage des modèles

<https://www.inria.fr/fr/risque-effondrement-collapse-ia-generatives>

L'attaque des bots

The impact of AI bots and crawlers on open repositories:
Results of a COAR survey, April 2025

June 3, 2025

Kathleen Shearer and Paul Walk



Image from: <https://www.flickr.com/photos/katerha/> CC BY 2.0

- 66 entrepôts répondants
- 90% font face à des bots agressifs
- 80% ont fait face à des ruptures de services, en lien avec ces bots
- Des moyens humains et informatiques importants dédiés à cette difficulté

<https://coar-repositories.org/wp-content/uploads/2025/06/Report-of-the-COAR-Survey-on-AI-Bots-June-2025-1.pdf>

Les AO contre-attaquent

*“A **repository** that indiscriminately **blocks machines** from accessing its content **cannot be FAIR.**”*



AI Bots and
repositories Task Force

- Technical group
- Monitoring group

“The goal should be to create tools and frameworks that allow repositories to remain accessible, discoverable, and resilient even in an environment shaped increasingly by machine users.”

En cours de lancement
par COAR en juin 2025

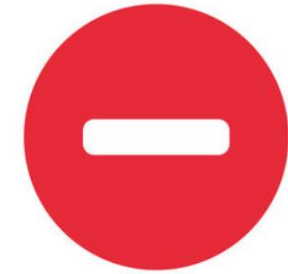
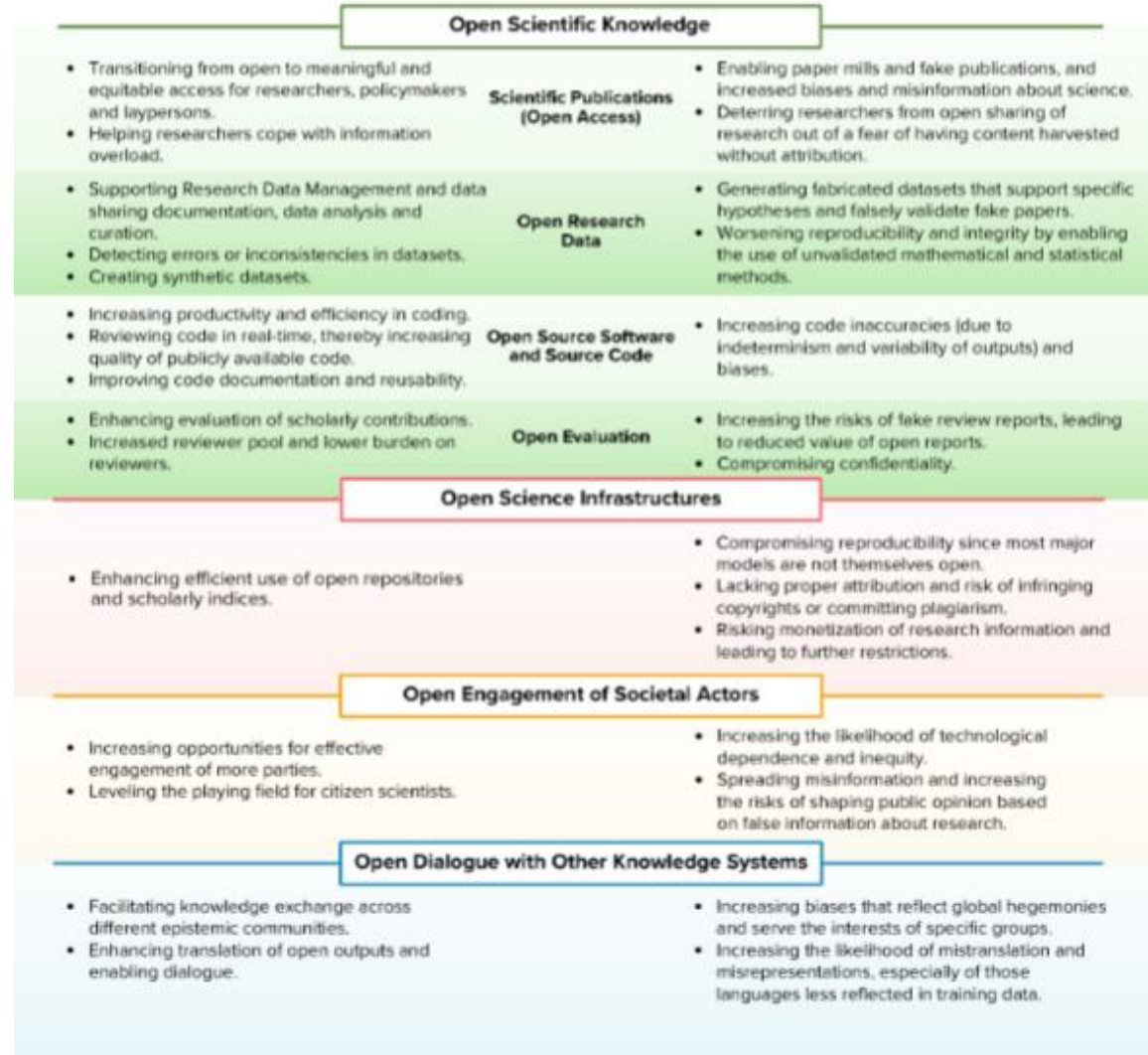
Porte d'entrée n°2 : la confiance





Possible **Positive** Impacts of GenAI upon Open Science

Possible **Negative** Impacts of GenAI upon Open Science



- Approximations
- Biais
- Paper mills
- Faux jeux de données
- Aggravation des problèmes de reproductibilité

Responsabilité, transparence, confiance

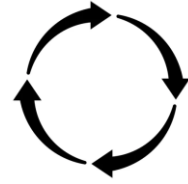


<https://www.ofis-france.fr/les-entretiens-de-lofis-lintegrite-scientifique-vu/>

« Prenons un exemple : Alice écrit de manière tout à fait classique, mais le résumé ne la satisfait pas. Elle demande alors à ChatGPT de réécrire ce résumé de manière plus percutante à partir de cette première version. Puis elle le relit, le modifie au besoin, pour assumer la pleine responsabilité du résultat : si elle déclare l'avoir écrit avec l'assistance d'IA, **ce n'est pas problématique**. En revanche, si elle copie directement un texte de ChatGPT sans le relire et qu'en plus, elle ne le déclare pas, **c'est inadmissible**. »

« Les entretiens de l'Ofis » : Intelligence artificielle et enjeux d'intégrité scientifique, l'analyse d'**Alexei Grinbaum**

Un cercle vicieux



- Les IAG intensifient la pollution de la science : phrases torturées, *regenerate response*...
- Les rétractions d'articles peuvent créer une réaction en chaîne sur les articles citant des articles rétractés
- Cette science polluée vient ensuite directement impacter la qualité des IAG

nature

Explore content ▾

About the journal ▾

Publish with us ▾

Subscribe

[nature](#) > [comment](#) > article

COMMENT | 28 August 2024

Chain retraction: how to stop bad science propagating through the literature

Awareness of when a paper can't be trusted is often too low – but adopting some easy-to-use technological solutions can help researchers, publishers and referees to clean things up.

By [Guillaume Cabanac](#) 

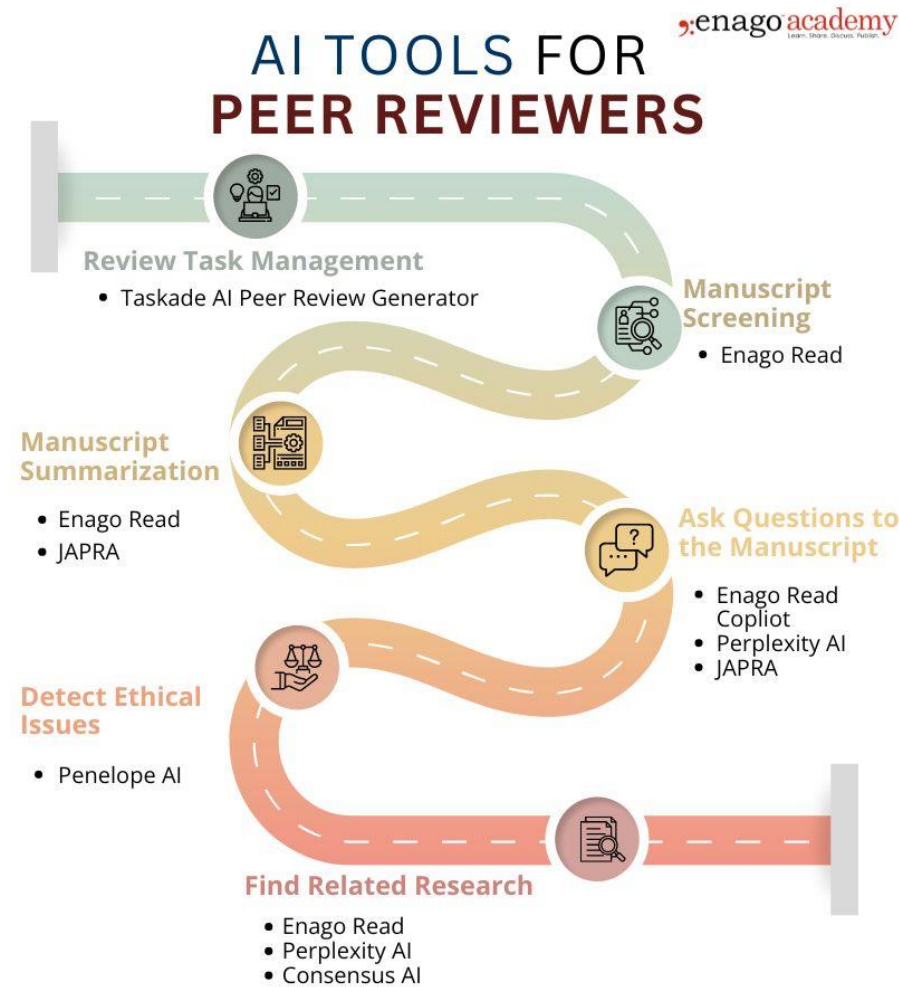
<https://www.nature.com/articles/d41586-024-02747-1>

IA et peer-reviewing

« I submit a manuscript for review in the hope of getting comments from my peers. If this assumption is not met, the entire social contract of peer review is gone. »

Timothée Poisot,
Université de Montréal

Myriam Naddaf. "AI is transforming peer review — and many scientists are worried." *Nature*, 26 mars 2025, <https://doi.org/10.1038/d41586-025-00894-7>

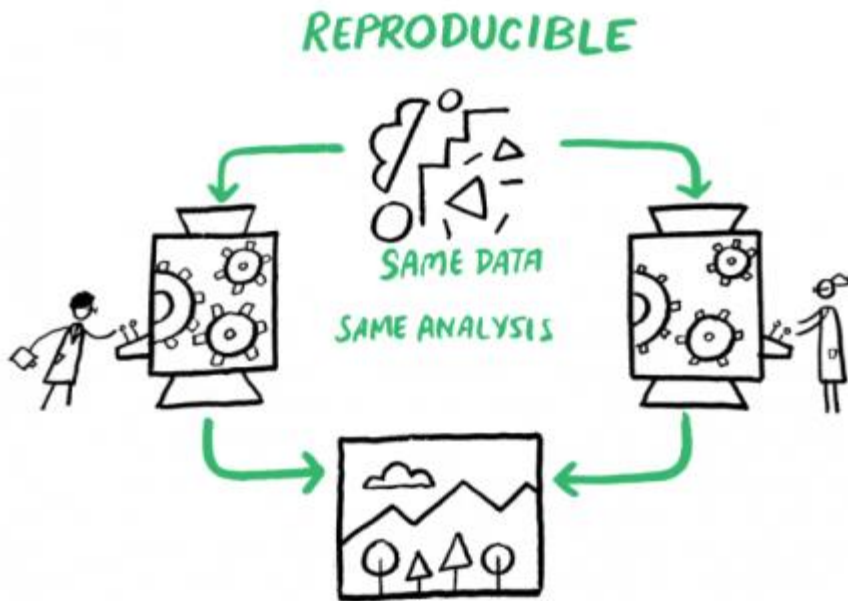


<https://www.enago.com/academy/6-ai-tools-peer-review-process/>

- Une rupture du lien de confiance entre auteurs, peer-reviewers et éditeurs ?
- Une simplification du processus de peer-reviewing ?

IAG et reproductibilité de la recherche

- Les IAG peuvent simplifier la documentation des process et la production de métadonnées associées
- Mais l'usage des IAG actuelles dans un processus de recherche constitue un obstacle fort à la reproductibilité de la recherche



Source : INRAE

Les modèles IA ouvertes, une clé pour la reproductibilité

Open weight vs Open
source



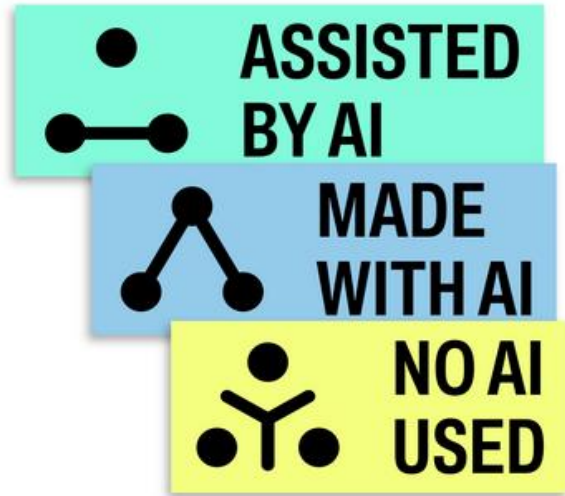
Un modèle IA Open Source :

- Code source
- Poids du modèle
- Données d'entraînement
- Documentation complète

Un modèle Open Weight :

- Ouverture des poids
- Modèle peut être utilisé et ajusté,
mais pas décortiqué
- Contrôle maintenu sur l'origine du
modèle

La transparence par les labels



**Fairly Trained
Certified**

Des labels pour les œuvres d'art
<https://ai-label.org/>

Des labels pour les IA
respectueuses du droit d'auteur
<https://www.fairlytrained.org/>

Poursuivre le rapprochement entre science ouverte et intégrité scientifique



Agir pour l'intégrité scientifique : quels rôles pour les bibliothèques et structures documentaires de l'enseignement supérieur et de la recherche ?

Propositions de l'ADBU

Préambule

Enjeu majeur pour la confiance dans la science, l'intégrité scientifique est définie par le Code de la recherche comme "l'ensemble des règles et des valeurs qui doivent régir les activités de recherche pour en garantir le caractère honnête et rigoureux" (article L.211-2). Elle implique des valeurs de respect, de responsabilité et d'honnêteté tout au long du processus de recherche et de fiabilité dans la conception et l'usage des ressources.

Accompagnement sur les pratiques de science ouverte, formations sur la citation des sources ou sur le droit de l'information, signalement et diffusion des résultats de la recherche : les bibliothèques et services documentaires de l'enseignement supérieur et de la recherche apportent aujourd'hui une contribution importante à l'information de la communauté académique sur les enjeux de l'intégrité scientifique. Leurs actions pour la diffusion de la science et pour le développement des compétences informationnelles en font des acteurs majeurs dans un environnement fragilisé par l'apparition des technologies d'intelligence artificielle.

Dans la continuité de leur accompagnement des évolutions de la production et de l'accès à l'information scientifique, les bibliothèques et services documentaires ont ainsi vocation à renforcer leurs actions pour l'information des communautés et la prise en compte de l'intégrité scientifique dans le signalement de la production scientifique.

Adressées à la communauté académique et aux bibliothèques de nos établissements, les propositions de l'ADBU ont pour objectif de mieux définir le rôle possible des bibliothèques et de faciliter le développement de leur action aux côtés de l'Office Français de l'Intégrité Scientifique, du réseau des référents et référentes Intégrité Scientifique et de l'ensemble des communautés de recherche engagés pour l'intégrité scientifique.

Propositions aux acteurs de l'intégrité scientifique dans l'enseignement supérieur et la recherche

1. Poursuivre le développement d'un réseau national impliquant les professionnels de l'information scientifique aux côtés des référents intégrité scientifique : favoriser le développement des liens entre tous les acteurs engagés au service de l'intégrité scientifique, préciser les rôles respectifs de ces différents acteurs.

« 9. Contribuer à la promotion des pratiques responsables : gestion des références, citation des sources, utilisation raisonnée de l'intelligence artificielle, bonnes pratiques éditoriales, pratiques vertueuses en matière de gestion des données, bonne attribution des fonctions d'auteurs... »

<https://adbu.fr/actualites/agir-pour-lintegrite-scientifique-quels-roles-pour-les-bibliotheques-et-structures-documentaires-de-lenseignement-superieur-et-de-la-recherche>

Porte d'entrée n°3 : les métiers



Une équation à multiples inconnues



Les IAG apprennent progressivement à faire la distinction entre les muffins et les têtes de chihuahua

<https://blog.adatechschooll.fr/intelligence-artificielle/>

- Hypothèses technologiquement fondées ou prédictions marketing ?
- Stabilisation des modèles économiques ?
- Niveaux de fiabilité et impact sur nos choix humains de délégation ?

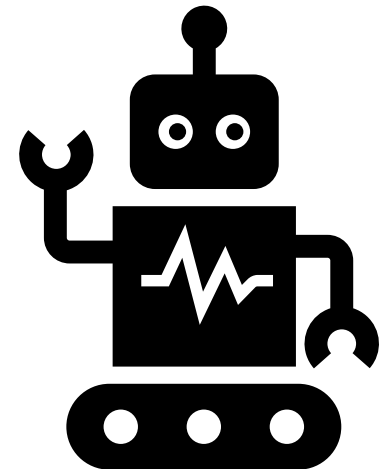
Les multiples fractures du numérique



Francesco Ungaro, Sol sec

- Inégalités d'accès au numérique, du Nord au Sud, des GAFAM à notre quotidien professionnel...
- Du potentiel technique à la réalité des feuilles de route de nos outils.
- Une question de temps ? De choix ? L'arrivée de nouvelles variables dans un monde en profond bouleversement ?

L'impact des IAG est potentiellement profond sur l'écosystème de la communication scientifique et de la science ouverte en général.



En quoi les IAG peuvent changer la donne ?



Maksim Goncharenok

- Automatisation de la création de documents structurés et des métadonnées
- Facilitation de l'usage des entrepôts ouverts et des indicateurs de la recherche
- Gestion de la documentation et du contrôle qualité des jeux de données
- Amélioration de la documentation des codes et de leur potentiel de réutilisation

**Faciliter la mise en pratique de la science ouverte,
38 au niveau du chercheur comme du professionnel de l'information**

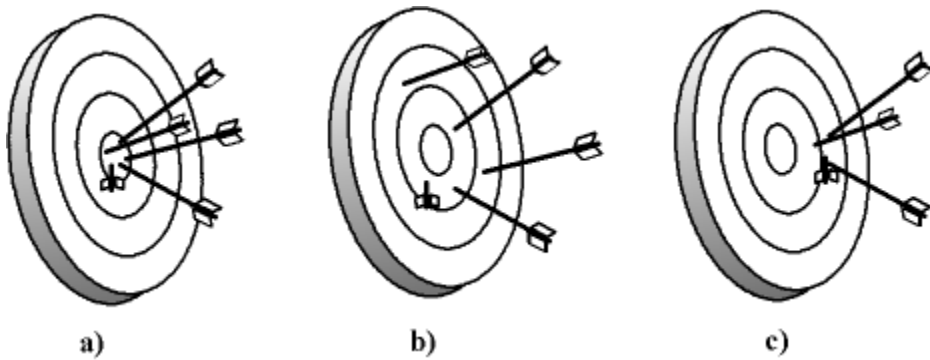
Dans un monde idéal



CC0

- Le dépôt initial par les chercheurs devient l'exception, au profit du développement d'autres approches ouvertes.
- Le travail de curation des métadonnées et de validation technique des fichiers devient semi-automatisé, sous un contrôle humain final.
- Le temps humain dégagé peut être mobilisé :
 - Pour reprendre le contrôle des processus éditoriaux, à l'aide de ces mêmes technologies
 - Pour promouvoir la bibliodiversité et renforcer l'accompagnement et le suivi des pratiques disciplinaires

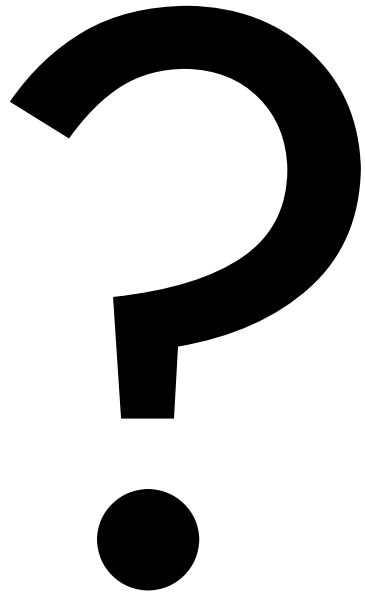
Mais plus probablement...



Christophe Dang Ngoc Chan, CC BY-SA
https://commons.wikimedia.org/wiki/File:Incertitude_illustree_par_fleches_et_cible.png

- La plus-value apportée par le secteur éditorial se déplace(ra) vers d'autres objets.
 - Facilitation des états de l'art
 - Services d'analyse
- Le développement technologique des archives ouvertes sera ralenti par la fracture numérique entre privé et public.
- La perte de confiance dans la science se renforcera et fragilisera les archives ouvertes.

Les IAG viennent questionner l'ensemble de notre production écrite, des métadonnées au texte



Quelle est la plus-value de notre activité ?

L'exemple du plan de gestion des données

« Je vais produire mon plan de gestion de données avec une IA. »
Une chercheuse de l'Université de Lille

- Fonction initiale du PGD vs perception par les chercheurs
- Valeur et sens d'un PGD produit par une IAG ?
- Comment faisons-nous évoluer les modalités de notre accompagnement sur la gestion des données et sur les PGD ?



Copie d'écran de DMP Opidor

Principes de délégation et de réversibilité



Delegation by [Nick Youngson](#) CC BY-SA 3.0 Alpha Stock Images

« Prévoir la réversibilité des services métier au travers d'un mode utilisable sans système d'IA »

« Il est proscrit de mettre en œuvre des technologies d'IA G dans le but de remplacer une action humaine, mais le faire pour compléter, prolonger et démultiplier l'action humaine. »

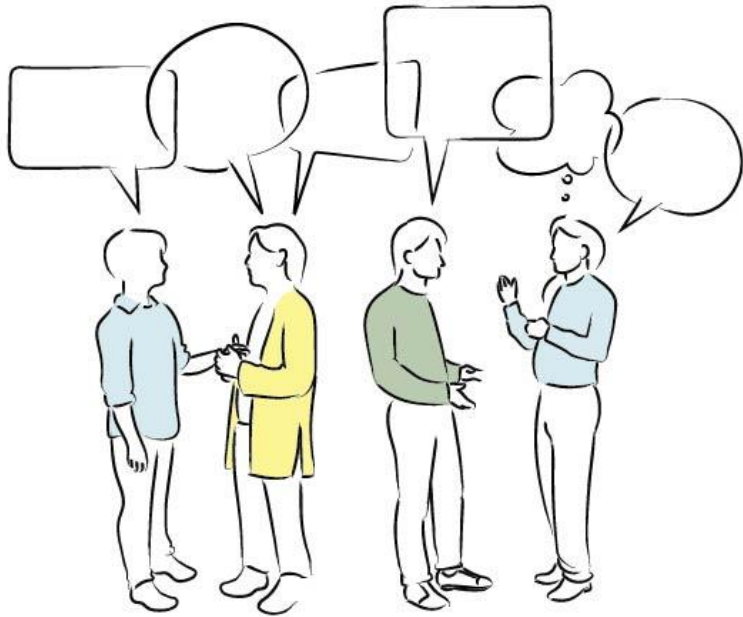
Projet Demoes. IA génératives et ESR : projets de chartes et de guides

<https://demos-pro3.fr/wp-content/uploads/2025/03/Charte-et-bonnes-pratiques-IA-29-mars-2025.pdf>

Mais alors, que faire ?



Ouvrir le débat



Waitala, CC BY-SA 4.0

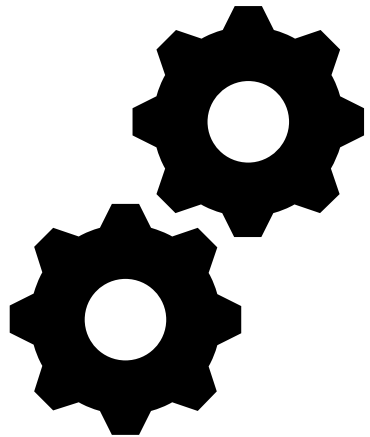
Avec la communauté scientifique

- Communautés de recherche
- Porteuses et porteurs de revues et communauté des éditeurs diamant

Avec les acteurs de l'écosystème des publications

- Le CCSD
- Les éditeurs
- Les acteurs de l'intégrité scientifique

Investir le champ de l'intégrité scientifique



- Penser collectivement l'évolution des archives ouvertes à l'ère de la désinformation et de la *fake science*.
- Sensibiliser et former à la mention des usages de l'IA, à la réutilisation de contenus produits par IA et à l'enjeu des rétractions en chaîne.
- Imaginer des pistes pour la labellisation des contenus scientifiques et pour la transparence dans l'usage des IAG.

Penser dès à présent l'évolution de nos activités et de nos compétences



- Développer une expertise mutualisée sur l'usage de nos collections ouvertes dans les IAG et sur l'usage de nos productions scientifiques par les éditeurs.
- Imaginer de nouvelles méthodes pour le contrôle qualité des métadonnées et des contenus (délégation et réversibilité).
- Déployer de nouvelles manières d'accompagner la science ouverte et la bibliodiversité.