

Bibliothèque de recherche et IA : vers de nouveaux espaces documentaires ?

Marc Delhaye

Chargé d'enseignement sur l'usage du numérique
à l'Institut d'études politiques IEP d'Aix-en-Provence

Sébastien Perrin

Directeur de la bibliothèque de l'École des Mines de Paris
et membre de la coordination exécutive de Couperin

L'arrivée, inattendue, des outils d'intelligence artificielle générative semble pour l'instant rester extérieure aux bibliothèques. Leur définition même ne fait pas consensus. On se reportera donc à la définition européenne, telle que rédigée dans l'Artificial Intelligence Act du 13 juin 2024 : « système d'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. »¹

La conscience du potentiel transformatif de ces outils dans le secteur de la documentation est réelle. Néanmoins, il convient d'affirmer au plus vite les apports des bibliothèques à ce nouvel écosystème en train de se constituer. Autrement, il est à craindre que d'autres acteurs prennent la place laissée vide, à commencer par les géants du numérique.

L'expertise des bibliothèques doit être positionnée dans ce nouvel écosystème de la recherche documentaire².

À première vue, elles ne sont ni créatrices de contenus (ou marginalement), ni dotées des compétences informatiques en Machine Learning, leur place est à trouver. De plus, le cœur de métier traditionnel des bibliothèques, la production de métadonnées est peu adaptée à l'IA générative³. Mais, en réalité, les bibliothèques sont d'ores et déjà partie prenante des changements en cours. À certains égards, tout se passe comme si la place laissée vacante attendait d'être occupée. L'examen des clauses d'interdiction de l'IA dans les contrats passés avec les fournisseurs de ressources numériques a ainsi initié un important mouvement de réflexion et de prise en compte des enjeux dans certains secteurs de la documentation et de l'enseignement supérieur et de la recherche (ESR).

Loin du souci de défendre une corporation, nous souhaitons démontrer que, pour l'enseignement supérieur, cet espace est celui du « bon usage » de ces technologies et de garantie de l'accès au savoir. À rebours des réseaux massifs détenus par quelques seuls acteurs, il s'agit de donner corps à d'autres « réseaux », complémentaires et comme ajoutés aux réseaux de neurones créés par la technique.

Pourquoi les bibliothèques sont-elles déjà embarquées dans l'IA ?

Du fait de la contractualisation liée aux ressources de documentation électroniques, les bibliothèques sont partie prenante des discussions, à plusieurs titres⁴ :

1 European Parliament et Council of the European Union, « 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) », Pub. L. No. 32024R1689 (2024), en ligne : <https://eur-lex.europa.eu/legal-content/FR/ALL/?uri=CELEX:32024R1689>

2 Leo S. Lo et Cynthia Hudson Vitale, « Quick Poll Results : ARL Member Representatives on Generative AI in Libraries », ARL Views (blog), 9 mai 2023. En ligne : <https://www.arl.org/blog/quick-poll-results-arl-member-representatives-on-generative-ai-in-libraries/>

3 Signalons néanmoins l'excellent blog de Géraldine Geoffroy : *IA et Bibliothèques*, <https://iaetbibliotheques.fr/>

4 International Coalition of Library Consortia (ICOLC), « ICOLC Statement on AI in Licensing », ICOLC, 22 mars 2024, <https://icolc.net/statements/icolc-statement-ai-licensing> et Consortium Couperin, « Printemps Couperin 19 au 21 mars 2025 Paris (France) », *Sciencesconf*, 19 mars 2025, <https://printempscouperin.org/resource/page/id/1>

- 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 et « chatbots » mis en place par les fournisseurs.

L'organisation de l'achat et ses modalités consortiales ont pu faire ignorer cette réalité. Pourtant, à partir de décembre 2023, les consortia européens et américains ont vu arriver de la part de plusieurs fournisseurs (Elsevier, Springer, etc.) la demande d'introduction de clauses d'interdiction pure et simple de l'IA dans leurs licences.

Afin de proposer un argumentaire solide aux négociateurs comportant les lignes rouges du consortium en matière d'utilisation de l'IA, Couperin a mis en place une démarche agile, impliquant l'ensemble des expertises concernées au sein d'un groupe de travail (GT) *ad hoc* et limité dans le temps : juristes, chercheurs en IA, délégué à la protection des données (DPO), experts en propriété littéraire et artistique, et bibliothécaires. Ces spécialistes provenaient de différents établissements (Commissariat à l'énergie atomique et aux énergies alternatives, Institut national de recherche en informatique et en automatique, Saclay, etc.).

Outre la discussion à propos des licences, domaine par essence soumis aux aléas de l'accord des parties, ce GT fut l'opportunité de préparer des éléments de nature générale pour mieux défendre les intérêts de l'ESR. Couperin a en effet toujours défendu l'intérêt des communautés de recherche, en particulier lors des échanges autour de la loi Lemaire et de la mise en place d'une exception pour le Text et Data Mining (TDM).

Le premier souci du GT fut de distinguer, au sein du « *continuum de l'intelligence artificielle* », les techniques historiques de traitement automatique des données des nouveaux outils d'intelligence générative. En effet, par leur nature extensive, les clauses de prohibition des éditeurs pouvaient, tel un marteau trop large, entraîner sinon une interdiction du TDM – en théorie chose difficile, du fait de l'existence d'une exception légale⁵ – du moins contraindre cette dernière.

Un important travail de définition des termes fut donc entrepris et alimenté par les chercheurs experts. Ensuite, sur la base de négociation menée en collaboration avec l'Agence bibliographique de

Outils génératifs éditoriaux et bibliothèque : un mariage de raison ?

Le marché de la documentation électronique promet de connaître une évolution majeure, à trois titres :

- concernant la politique documentaire, les éditeurs partageant leurs données courent le risque que leurs fonds soient rendus disponibles par d'autres acteurs (ChatGPT, Deep Seek, etc.), interrogeant les politiques d'acquisition ;
- le recours à des IA génératives opaques menace de « contaminer » ou de « polluer » les résultats de la recherche, et de la rendre invalide du simple fait de ce soupçon ;
- 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.

Face à ces menaces, les éditeurs proposent de plus en plus des IA spécifiques, reposant sur leurs corpus de publications et données validées et vérifiées, nourries par un outil génératif propre. Réponse pertinente face aux acteurs techniques, cette solution facilite également la création de bases de données de confiance. S'agit-il d'une promesse de création d'un écosystème informationnel propre à la recherche ?

Pour que ces outils répondent à cet objectif, l'expertise utilisatrice des services documentaires est nécessaire. En effet, les choix de conception des outils doivent être endossés par les clients finaux pour être expliqués au public et positionnés dans l'offre documentaire.

Prenons un exemple : un éditeur de bases bibliométriques aura sans doute la tentation de s'appuyer sur ses propres outils d'évaluation des publications pour générer un texte (citation, H-index, Ranking, etc.). On lui reprochera alors à bon droit d'utiliser des outils critiqués. Mais s'il ne les utilise pas, il s'agit alors de l'arbitraire d'un algorithme.

Nous ne prétendons pas trancher le débat. En revanche, il est évident que la transparence sur les protocoles de recherche et les étapes suivies par l'algorithme sont fondamentales, pour la reconnaissance par les institutions de cette IA académique « de confiance ».

Ici, se joue le nœud d'une évaluation à reconstruire. L'approche technique se caractérise par une méthode de test et d'erreur, difficile à mettre en place¹. Mais en réalité, il s'agit bien plus de concevoir et d'implémenter des protocoles d'informations documentaires sérieux et documentés, qui, en bonne part, sont présents dans les techniques de recherche documentaire depuis longtemps.

5 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, *Journal officiel « Lois et décrets »*, 25 novembre 2021, n° 274, en ligne : <https://www.legifrance.gouv.fr/jorf/id/JORFTEXT000044362034>

1. Consortium Couperin, « Printemps Couperin 19 au 21 mars 2025 Paris (France) », *Sciencesconf*, 19 mars 2025, en ligne : <https://printempscoup.sciencesconf.org/resource/page/id/1>

l'enseignement supérieur (Abes) pour Elsevier, des conditions licites d'utilisation facilement acceptables par le plus grand nombre de fournisseurs furent délimitées (exemple : possibilité de recourir à une IA dans un environnement hébergé et sécurisé, refus de toute clause semblant s'étendre au TDM), permettant de refuser les clauses moins-disantes. Dernier point : au-delà de la discussion commerciale, la proposition, reprise des recommandations de l'ICOLC⁶, du GT Couperin d'intégrer des clauses partenariales et des clauses de revoyure sur les usages constatés fut généralement très bien accueillie par les éditeurs – preuve du souhait de l'ensemble des acteurs de se laisser le temps de s'approprier les évolutions en cours⁷.

Ce travail n'est pas terminé : par définition, il s'adapte aux évolutions des demandes des fournisseurs, lesquels opèrent dans un environnement dont la compréhension évolue. On notera par exemple la récente intégration, dans certaines licences, d'un droit d'entraîner des IA ou non selon le Large Language Model (LLM) utilisé, distinguant LLMs « locaux » (développés en interne) et LLMs « Tiers » (ChatGPT par exemple). Le souci de protection des contenus éditoriaux n'y est sans doute pas étranger.

Les nouveaux espaces documentaires : conception de corpus et d'outils

Une autre activité associée étroitement les bibliothèques académiques, ou au moins leurs opérateurs nationaux, à l'IA : l'acquisition de vastes corpus de recherche. Certes, il s'agit de collections numérisées, mais aussi et surtout de collections acquises via les programmes ISTEEX et COLLEX. Ces dernières, grâce à leur contenu académique et aux enrichissements contrôlés de leurs métadonnées constituent une base de confiance pour la mise en place d'outils d'IA. Il s'agit de la même logique de création d'IA caractérisée par un corpus spécifique que celle mise en place par les éditeurs.

Les outils mis à disposition par l'Institut de l'information scientifique et technique (INIST) constituent un ensemble de services autour de la fouille de texte qui devraient être largement promus par les bibliothèques⁸. Il est à souhaiter que des expérimentations concernant l'IA générative soient systématisées sur ce corpus, les outils de TDM étant en soutien, l'ESR disposant grâce à lui d'un riche vivier de sources validées.

L'expérience d'ISTEX démontre que les bibliothèques ont *a minima* une expertise de sélection et de mise en qualité des métadonnées à faire valoir. Surtout, ces infrastructures dessinent un horizon, qui serait celui d'un usage « FAIR » de l'IA⁹.

En effet, sauf pour les IA grand public et dans la vision fantasmagique de certains acteurs exclusivement technologiques, l'IA n'a pas vocation à s'autonomiser de l'ensemble du capital acquis concernant la gestion des données. La réalité de la recherche et des établissements consistera bien plus certainement non dans l'utilisation d'une IA unique, mais dans l'exploitation de multiples IA, génératives ou non, positionnées chacune à un endroit de la recherche¹⁰. Plus au fond, l'IA générative n'a pas vocation à remplacer le continuum d'outils de recherche existant. Dans plusieurs cas d'usage (revue de littérature par exemple), une recherche dans une base bibliographique « classique », permettra d'extraire un corpus, comme une IA, mais surtout les résultats seront explicables et prédictibles.

« Des corpus qui parlent juste »

L'outil n'est rien sans le corpus, et le corpus n'est rien sans un cadre éthique de construction. C'est là que les bibliothèques peuvent et doivent intervenir – non pas comme expertes en algorithmie, ce qu'elles ne sont pas, mais comme « architectes de confiance documentaire », s'appuyant sur des corpus constitués (ISTEX, Gallica, Persée ou HAL par exemple). L'usage du TDM, en permettant d'enrichir à partir des corpus éditeurs, repose une sélection documentaire.

Mais pour que cette potentialité devienne réalité, encore faut-il penser l'IA non comme un substitut, mais comme un instrument au service d'un savoir situé. Cela suppose :

- une transparence sur les corpus utilisés (origine, méthode, licences) ;
- une explicitation des chaînes de transformation des données, afin de rendre explicables les résultats : extraction, sélection, annotation, vectorisation, pondération, restitution ;
- une lisibilité cognitive et sémantique des réponses générées ;
- une explicitation des critères de sélection activés par le modèle : thésaurus mobilisés, poids sémantique, hiérarchie documentaire, exclusions éventuelles ;

9 « TRUST-AI », TrustAI, 2025 2021, <https://trustai.eu/about.html>

10 <https://www.ibm.com/fr-fr/think/insights/artificial-intelligence-future>. Sur les expérimentations dans l'ESR : « RAGaRenn Un service expérimental de l'université de Rennes », RAGaRenn, 2 mai 2025, en ligne : <https://ragarenn.eskemm-numerique.fr/> et « Aristote, une IA au service de l'enseignement supérieur », CentraleSupélec Université Paris Saclay, 26 avril 2024, en ligne : <https://www.centralesupelec.fr/aristote-une-ia-au-service-de-lenseignement-superieur>

6 Consortium Couperin, « Couperin signe la déclaration IA de l'ICOLC », *Couperin.org*, 2 mai 2025, <https://www.couperin.org/le-consortium/actus/couperin-signe-la-declaration-ia-de-licolc/>

7 Sur ce travail, voir le Printemps Couperin : <https://printempscoup.sciencesconf.org/resource/page/id/1>

8 Istex TDM <https://services.istex.fr/>

• un droit de regard sur les modèles sollicités (LLM propriétaires ou locaux), incluant: l'identification claire du modèle utilisé (nom, version, éditeur); une transparence sur les corpus d'entraînement (types de sources, provenance, inclusion ou non de bases académiques); un accès à la documentation technique synthétique (étapes de la recherche documentaire, filtrages, vectorisation); la possibilité de choisir ou restreindre les modèles selon les usages; et « l'interrogeabilité » des réponses générées (justification, rétroliens, critères activés).

L'IA et son interprétation doivent être traçables, situées, disputables. C'est exactement ce que savent faire les bibliothécaires quand ils indexent, éditorialisent, transmettent. Il ne s'agit donc pas de se former à l'IA, mais de reconnaître que l'IA doit se former aux exigences du monde documentaire.

Au cœur de cette analyse se trouve l'idée d'IA « de confiance »¹¹. Cette dernière découle du constat que certes, l'utilisation d'IA grand public est devenue massive, y compris dans la recherche et l'enseignement, mais que ses fragilités (hallucinations, corpus mal maîtrisés, impossibilité de reproduction des résultats, dépendance aux demandes initiales, « pollution » des résultats, etc.) la rendent trop incertaine pour des utilisations nécessitant une fiabilisation absolue des résultats.

Par métaphore, l'action des bibliothèques vis-à-vis de l'IA consiste à l'intégrer dans des corpus de règles et de normes *déjà existantes mais à formaliser et qui constituent un autre réseau que les réseaux techniques de neurones*. Pour l'ESR, ces corpus et règles sont rendus nécessaires en raison de la solidité que doit produire une institution dans ses prises de décision et d'arbitrage, et de la confiance nécessaire dans les résultats de la recherche. Cette *compétence*, car il s'agit bien de cela, permettra de positionner les bibliothèques dans le processus de production du projet de recherche et de son ingénierie documentaire à l'ère de l'IA: conception des corpus, traitement des fichiers, conservation et réutilisation des données produites, sélection et recommandations des outils d'IA.

Réalités des missions des bibliothèques de recherche à l'heure de l'IA

Pour affirmer cette compétence, il s'agit de peser sur les régulations à l'œuvre dans le secteur. Le règlement européen sur l'IA du 13 juin 2024 constitue à cet égard une étape importante. Son article 53, 1, c oblige les fournisseurs de modèles d'IA à prendre

des mesures visant à respecter le droit d'auteur et les droits voisins. Il impose également à ces mêmes acteurs d'élaborer et de rendre publiquement disponible un « *résumé suffisamment détaillé* » des données ayant servi à l'entraînement de leur modèle (article 53, 1, d).

Un récent rapport du Conseil supérieur de la propriété littéraire et artistique (CSPLA) (*IA et Transparence des données d'entraînement*, 11 décembre 2024¹²) a présenté des pistes pour préciser cette obligation et un modèle de « *résumé suffisamment détaillé* ».

« *Dans le respect du secret des affaires, le rapport préconise une "approche par type de contenus, avec un degré de détail croissant" selon que les contenus sont libres de droit ou plus sensibles. Avec pour objectif affiché de permettre l'exercice des droits, le résumé devrait être "complet en termes de contenu", sans révéler les techniques utilisées.* » Le CSPLA indique dans ses réflexions: « *La portée normative du résumé doit être proportionnée à l'objectif poursuivi: aider les intéressés à faire valoir leurs droits.* »

La transparence de la recherche mobilise d'autres enjeux¹³. En effet, sans minimiser le travail du CSPLA, il s'agit en priorité pour les chercheurs de juger des corpus utilisés dans la recherche, non en vue uniquement de défendre leurs droits en tant qu'auteurs, mais afin de satisfaire aux besoins de la communauté des chercheurs et de son bon fonctionnement. En résumé, pour la recherche, « l'explicabilité » est une dimension essentielle de la transparence.

En définitive, il s'agit de rebâtir une expertise autour du cœur de métier des bibliothèques – « donner accès à une information validée » :

- compréhension d'une diversité d'outils et capacité de les évaluer, fondée sur une évaluation des corpus mobilisés ;
- capacité à s'inscrire dans le « faire » des projets de recherche, en s'appropriant pleinement les techniques de fouille de texte et en sélectionnant des outils au sein du continuum de l'IA (à commencer par la gamme de services TDM produite par ISTEEX) ;
- édicition de règles d'usage transposant les principes de la recherche FAIR, en particulier de capacité de réutilisation des données produites.

11 On trouvera ici une définition de l'INRIA : INRIA, « Construire une IA digne de confiance en Europe », 17 juillet 2024, <https://www.inria.fr/fr/ia-confiance-europe>

12 Ministère de la Culture, « IA et Transparence des données d'entraînement : publication du rapport d'Alexandra Bensamoun sur la mise en œuvre du règlement européen établissant... », 11 décembre 2024. En ligne : <https://www.culture.gouv.fr/nous-connaître/organisation-du-ministère/Conseil-supérieur-de-la-propriété-littéraire-et-artistique-CSPLA/travaux-et-publications-du-cspla/missions-du-cspla/ia-et-transparence-des-données-d-entraînement-publication-du-rapport-d-alexandra-bensamoun-sur-la-mise-en-œuvre-du-règlement-européen-etablissant>

13 On notera d'ailleurs l'absence de représentant de l'ESR dans la liste des contributeurs.

Évaluation des outils : le jugement documentaire à l'ère de l'IA

L'évaluation des outils IA est une compétence clé pour faire entendre l'expertise des services documentaires. Pour cela, on pourra distinguer à l'usage plusieurs niveaux d'IA générative « de confiance » :

- l'IA « éthique », acceptant de donner accès à sa gouvernance¹ ;
- une IA « éditoriale », telle que décrite ci-dessus ;
- l'IA « meta-réflexive », c'est-à-dire donnant à voir son raisonnement et en mesure d'expliquer ses choix².

On pourra également s'appuyer sur l'excellente analyse rédigée par Aaron Tay, fournissant une méthode d'évaluation des moteurs de recherche académiques permettant d'objectiver leur capacité de reproduction des résultats sur la base de leurs technologies³.

Plus génériquement, l'évaluation des outils repose sur une expérience commune des utilisateurs raisonnée du numérique :

- Le discernement cognitif : il ne suffit pas que l'outil réponde : il doit rendre lisibles ses mécanismes. Cela signifie pouvoir retracer les corpus sollicités, connaître les critères de réponse activés, et identifier les biais éventuels. Ce discernement inclut la possibilité de pondérer les résultats, d'accéder à des rétroliens, et de comprendre les logiques sémantiques mobilisées : thésaurus, *clustering*, exclusions. En somme, une interface qui soutient la lecture critique.
- L'enracinement éthique : chaque outil mobilise des données : d'où viennent-elles ? Qui les a validées, sélectionnées, labellisées ? Selon quels principes FAIR ou quelles chaînes de responsabilité ? L'outil respecte-t-il les droits d'auteur, les contrats de licence, les politiques d'accessibilité ? La traçabilité des usages – tant dans l'entraînement que dans la réponse générée – est une condition de confiance.

Sans elle, l'outil devient un espace opaque, voire un danger.

- L'utilité située pour l'utilisateur : un bon outil est un allié de l'apprentissage informationnel. Il doit pouvoir s'adapter à des publics variés (étudiants, chercheurs, professionnels), proposer une interaction pédagogique (guidée ou libre), et surtout renforcer l'autonomie critique. Cela suppose de sortir du « prêt-à-penser » algorithmique et de permettre à l'utilisateur de questionner, reformuler, ajuster. L'IA devient alors un partenaire dialogique, et non une boîte noire autoritaire⁴.

1. Unesco, « Recommandation sur l'éthique de l'intelligence artificielle », 2022, en ligne : https://unesdoc.unesco.org/ark:/48223/pf0000381137_fre

2. Pour une première approche passant par une « marketplace », voir l'IA mise en place par le gouvernement : <https://www.comparia.beta.gouv.fr/>

3. Aaron Tay, « The Reproducibility and Interpretability of Academic Ai Search Engines like Primo Research Assistant, Web of Science Research Assistant, Scopus Ai and More », *Aaron Tay's Musings about Librarianship* (blog), 14 avril 2025, <https://musingsaboutlibrarianship.blogspot.com/2025/04/the-reproducibility-and.html>

4. On renverra à la méthodologie proposée à ce lien, qui souligne justement la difficulté d'évaluer l'output des IAG : Aaron Tay, « Testing AI Academic Search Engines - What to Find out and How to Test (2) », *Aaron Tay's Musings about Librarianship* (blog), 1er mai 2025, en ligne : <https://musingsaboutlibrarianship.blogspot.com/2025/05/testing-ai-academic-search-engines-what.html>

Cette approche est nécessaire, alors que des zones de tension se manifestent :

- la gouvernance des corpus : les négociations contractuelles avec les éditeurs deviendront plus âpres, avec la montée en puissance des clauses d'interdiction des usages IA et la volonté de contrôler les corpus d'entraînement. Des questions commerciales vont notamment se poser avec les éditeurs qui auront passé des accords avec les IA « généralistes » ;
- la mutation silencieuse des interfaces : les IA éditoriales deviendront les guichets obligés de consultation. Si les bibliothécaires ne participent pas à leur conception, ils seront relégués au rôle de simples accompagnateurs, voire invisibilisés si les projets de déploiement de chatbots ne permettant pas la consultation des textes viennent à se confirmer. Il faut donc inscrire l'évaluation en amont, dans les protocoles, les interfaces, les algorithmes ;

- la contradiction croissante entre science ouverte et protection des données : face au risque de pillage généralisé des corpus, l'Open Science ne pourra rester univoque. Il faudra assumer des formes de clôtures éthiques, capables de préserver sans trahir.

Mais il convient de ne pas sous-estimer certaines ruptures qui posent des problèmes authentiquement nouveaux. On songe en particulier à la propriété des données générées totalement ou en partie par traitement génératif. Qui peut se réclamer le propriétaire de ces œuvres composites ? L'éditeur, le concepteur du prompt, le possesseur de la base de données ? Comment résoudre la tension entre « explicabilité » et « efficacité » des moteurs de recherche ? Ces territoires restent, eux, vraiment à découvrir. ☺