



05/05/2026

POSITION DU CONSORTIUM COUPERIN sur l'intégration de l'intelligence artificielle dans les ressources documentaires

L'intelligence artificielle générative s'est désormais largement diffusée dans la société, transformant les usages et les attentes des utilisateurs, et posant de nouvelles questions pour le secteur de l'information scientifique et technique. Voici quelques points importants à retenir :

- **Modification profonde des pratiques et dérives des usages :**
 - L'IA générative se présente comme une technologie d'amélioration des moteurs de recherche et bases de données bibliographiques traditionnelles et vient s'ajouter aux outils classiques utilisés dans l'information / documentation ;
 - Les utilisateurs obtiennent désormais des réponses directes (et non plus une liste de sources), ce qui modifie leur rapport à l'information. Cependant, ces grands modèles de langage (LLM) invisibilisent parfois les sources et peuvent favoriser ainsi le plagiat non intentionnel ;
 - La facilitation de la production des travaux scientifiques permis par l'IA impacte l'ensemble de la chaîne : la fraude ou les analyses biaisées peuvent à terme mettre en doute la valeur des corpus fournis par les éditeurs.
- **Nécessité de mise en place de garde-fous dans le contexte académique**
 - Les IA mises à disposition du public académique doivent intégrer des outils ou des moyens de contrôle des défauts bien connus de l'IA générative (biais, réponse aléatoire) ;
 - Les fournisseurs de services utilisant des IA génératives doivent fournir des garanties sur les données récupérées et le pistage. En effet, les IA captent souvent les contenus de sources externes ainsi que les données fournies par les utilisateurs, en particulier via leurs requêtes et le suivi de leur profil, sans véritable garantie concernant les autorisations accordées et l'utilisation qui en est faite.

Dans ce contexte en mutation rapide, de nombreux éditeurs scientifiques ou académiques ont commencé à déployer des fonctionnalités exploitant l'IA générative au sein de leurs plateformes. **Le souci de fournir une IA académique dite de confiance constitue un horizon et un intérêt partagé par tous les producteurs et distributeurs d'information. Il nous apparaît donc nécessaire d'exprimer clairement le positionnement du consortium sur ces évolutions.**

Des modèles économiques préoccupants

Les modèles économiques qui se dessinent soulèvent plusieurs inquiétudes, alors même que les établissements d'enseignement supérieur et de recherche sont confrontés à une tension budgétaire croissante qui limite leur capacité d'investissement.

En effet, certains éditeurs proposent des modules d'IA comme des extensions payantes des abonnements existants, avec des surcoûts pouvant atteindre 30 % du montant initial. Cette stratégie de commercialisation est susceptible de générer une charge financière considérable que nos institutions ne pourront pas absorber, à



plus forte raison alors que les acteurs poussent de plus en plus les fonctionnalités de recherche approfondie ou les agents IA.

Au-delà de l'aspect financier immédiat, ce développement de l'IA générative comme moyen, en même temps, d'accès commode aux ressources documentaires pose question dans la mesure où ces technologies n'ont pas encore fait la preuve de leur plus-value réelle pour la recherche et la construction des savoirs.

Cette logique d'empilement tarifaire entre en contradiction avec notre vision du développement des plateformes de contenus scientifiques dans un contexte intègre et responsable. Si nous reconnaissons volontiers les efforts de R&D consentis par les éditeurs, nous refusons que la communauté académique soit la seule à financer l'innovation. Les améliorations apportées par l'IA constituent souvent des enrichissements fonctionnels plutôt que des ruptures technologiques. À ce titre, elles relèvent de l'amélioration continue des plateformes, et doivent être incluses dans l'évolution tarifaire normale des ressources négociées, et non facturées comme des prestations supplémentaires. Le risque est autrement que l'usage de ressources d'IA génératives des chercheurs et étudiants se déportent massivement sur les acteurs « grand public ».

Par ailleurs, dans un contexte accru de préoccupations environnementales du monde de la recherche, cela encouragerait un usage systématique de l'IA, alors même que cette technologie est extrêmement coûteuse en ressources, notamment énergétiques. L'usage de l'IA dans l'enseignement supérieur doit au contraire être raisonné, ciblé sur les situations où elle apporte une réelle valeur ajoutée, et non déployé de manière automatique sur l'ensemble des fonctionnalités.

Le savoir comme bien partagé de l'ensemble des acteurs de la chaîne de production de l'information

L'arrivée de l'IA générative est une occasion pour les établissements du supérieur d'échanger au sujet de la conception technique des outils proposés par les fournisseurs de solution.

Il ne s'agit pas de se substituer aux acteurs privés, mais de comprendre les outils qui seront proposés aux étudiants et chercheurs. Cette démarche peut associer les nouvelles plateformes d'IA générative tout comme les acteurs traditionnels : il s'agit d'obtenir de ces fournisseurs les moyens d'arriver rapidement à une information validée, par exemple les articles appuyant la génération du texte.

Il s'agit de co-construire, par un dialogue suivi avec les éditeurs et diffuseurs, un diagnostic des attentes et des besoins des communautés et des réponses techniques attendues. D'autant que certains de ces nouveaux services sont largement dépendants de l'écosystème de la science ouverte, dont les membres de Couperin sont parties prenantes : enrichissement des profils des chercheurs, outils de revues de littérature, informations sur les publications et leur impact, etc.

Notre vision des services d'IA « de confiance »

Les fonctionnalités d'intelligence artificielle intégrées aux plateformes éditoriales doivent respecter des principes fondamentaux. Elles doivent faciliter l'accès aux documents primaires et permettre leur meilleure valorisation, et non se substituer à leur consultation par des synthèses décontextualisées. La traçabilité et la citation systématique des sources sont indispensables pour maintenir l'intégrité de la démarche scientifique et éviter la prolifération de contenus synthétiques coupés de leurs fondements documentaires. D'une manière générale, le recours aux IA génératives dans le fonctionnement de l'outil doit toujours être rendu visible dans les interfaces et déclaré dans les réponses fournies.

Une IA académique de confiance repose sur la « transparence ». Cette dernière peut être qualifiée en ces termes :

- **Documenter au maximum les choix qui ont présidé à sa création** : nature et licéité des sources ayant servi lors de l'entraînement des LLM, sélection du corpus et données (full text, métadonnées) utilisés pour la génération de la réponse du RAG (Retrieval-Augmented Generation), critères bibliométriques et/ou statistiques de choix des sources présentées, présentation de la requête booléenne éventuellement créée, citation explicite des sources dans le texte généré, etc. ;
- **Garder la possibilité d'utiliser les services sans recourir à ces IA génératives** (ex. : recherche par mots-clés stricts, requêtes booléennes) afin de répondre aux différents cas d'usage des ressources documentaires. Ce point est particulièrement important pour les cas nécessitant une démarche reproductible ou exhaustive, comme la revue systématique de littérature ;
- **Garantir la confidentialité des données fournies et leur absence d'utilisation** notamment pour l'entraînement des modèles et prendre en compte la protection de la vie privée dès le stade la conception du modèle, et empêcher la divulgation des données personnelles, confidentielles et de recherche ;
- **Justifier leur déploiement par une véritable utilité fonctionnelle** compte tenu de l'empreinte environnementale des technologies d'IA renforcée par la recherche approfondie et l'IA agentique. Un usage raisonné suppose de réserver l'IA aux tâches où elle apporte un bénéfice manifeste, plutôt que de l'appliquer systématiquement à l'ensemble des services.

Enfin, il convient de fournir un accompagnement des utilisateurs et des communautés (aide, FAQ, banque d'exemples variés) afin de favoriser le développement d'une véritable littératie de l'IA dans un contexte académique. Au-delà des données et des technologies, l'enjeu majeur est celui de la capacité des étudiants et des chercheurs à utiliser les outils de manière éclairée.

Vers une intégration transparente et innovante de l'IA dans les ressources documentaires

Des alternatives concrètes existent et méritent d'être mises en lumière : certains éditeurs ont déjà fait le choix d'intégrer les développements liés à l'IA directement dans leurs offres, en répercutant ces investissements par des ajustements tarifaires, mesurés et transparents. Cette démarche nous semble exemplaire et c'est celle que nous soutenons dans les négociations que nous menons.

Au sein des institutions, la mise à disposition des corpus et leur utilisation par l'IA ouvrent la voie à de nombreuses innovations dans le domaine de la recherche et de son pilotage. Il est essentiel que les établissements d'enseignement et de recherche conservent **un droit à la réutilisation et à l'expérimentation sur les corpus négociés**, par exemple par des RAG (Retrieval-Augmented Generation) ou par le protocole MCP (Model Context Protocol) sur des serveurs locaux, dans le respect de la propriété intellectuelle et en application de l'exception TDM pour l'enseignement et la recherche. Il nous semble indispensable de promouvoir des modèles économiques qui reconnaissent pleinement la valeur de l'innovation tout en garantissant l'accessibilité durable des ressources documentaires pour les communautés académiques. Dans ce contexte, il nous semble également important de pouvoir continuer à profiter d'offres différenciées intégrant, ou non, ces nouvelles fonctionnalités IA aux ressources documentaires, mais aussi des indicateurs différenciés des usages.

Ces enjeux appellent une réflexion collective au niveau international. Les échanges en cours avec les industriels de l'information scientifique et technique ont facilité une rapide montée en compétence sur ces outils. **Nous invitons tous les partenaires de Couperin - privés ou publics, étrangers ou français - à se rapprocher du consortium afin de construire ensemble ces approches.**