

Usages en recherche de logiciels (dits d'« intelligence artificielle ») de production de texte, code ou images : questions d'intégrité scientifique – partie 2

64^e congrès de la Société des Anglicistes de l'Enseignement supérieur

Panel 1 – Transitions induites par l'Intelligence Artificielle dans la recherche en anglistique

6 juin 2025

Catherine Tessier
referent-integrite-ethique@onera.fr

Biais induits chez l'utilisateur

- **qualité perçue** du résultat (bien écrit, bien construit, sans fautes...)
→ peut induire la sensation que ce résultat est scientifiquement valide
- **biais d'automatisation** (si l'outil le dit, c'est que c'est vrai)
et **biais de confirmation** (renforcement de ses propres croyances)
→ susceptibles d'être accentués
- risque d'une **illusion de compréhension**
→ suis-je **expert(e)** du domaine ?

Questions (1) : l'usage de l'outil est-il approprié ?

quels arguments pour justifier l'usage ?

- Arguments scientifiques, lesquels ?
- C'est « à la mode », « tout le monde » l'utilise
- L'outil fait « gagner du temps », évaluer le « temps gagné » (paradoxe ?)
- L'outil permet de pallier un déficit de ressources ou de compétences, lesquelles ?
- L'outil est-il susceptible d'être biaisé : biais culturels, idéologiques, déficit de corpus... ?

Questions (2) : recherche bibliographique et production automatisée d'état de l'art [Bouchard24]

- Quel corpus d'entraînement du logiciel ?

risques : comprend revues ou conférences de qualité médiocre, articles corrigés ultérieurement ou rétractés, documents non évalués par les pairs (archives ouvertes, blogs), opinions...

- Représentativité des résultats ?

fraîcheur, pertinence, complétude (ouvrages ? documents payants ?), existence

– ne pas prendre connaissance des références (citer les références sans les avoir lues ni même consultées) ↑

– plagiat d'un état de l'art ou d'un *survey* existant, plagiat de citations

+ privilégier des outils de littérature académique, en utiliser plusieurs, croiser les résultats

+ vérifier les références suggérées et les lire !

[Bouchard 24] Aline Bouchard. *Au-delà de ChatGPT – Recherche d'informations académiques et intelligence artificielle*. URFIST Paris, mai 2024. [lien](#)

Questions (3) : où vont les données des requêtes (*prompts*) ?

Comment sont-elles **utilisées** par les sociétés qui proposent les logiciels ?

+ Vigilance relativement aux données

- sensibles (données personnelles, données de recherche)
 - confidentielles (hypothèses, données pas encore publiées)
 - protégées (propriété intellectuelle)
-
- production (très) facilitée de données créées de toutes pièces (fabrication), de données « arrangées » (falsification)
 - risque de pression temporelle accrue sur la production de résultats scientifiques, « parce que l'IA fait gagner du temps » (ex: traduction)

Questions (4) : usages pour la rédaction ?

- Un logiciel n'est pas « co-auteur » d'un document scientifique (consensus)
- Les auteurs d'un document scientifique qui utilisent un tel logiciel :
 - + doivent le signaler dans le document : nom de l'outil, version, (requêtes)
 - + expliquer les usages
rédaction, production d'images, de graphiques, usages pour le processus de recherche...
 - + s'assurer de la correction de ce qui est produit
 - + vérifier qu'il n'y a pas atteinte à la propriété intellectuelle d'autrui
 - difficile ?
- plagiat délibéré : facilité et de plus en plus difficile à détecter
plusieurs passes dans le logiciel pour reformuler, adapter le style, changer la langue...
- inflation de faux documents scientifiques

Questions 5 : examen par les pairs ?

Par définition (sauf si pré-publication de la version soumise)

l'examen par les pairs porte sur des documents **non publiés**, (articles, thèses, propositions de projets, etc.), donc **à ne pas diffuser**

- divulgation de connaissances originales par téléchargement dans un logiciel externe, ouvert
- multiplication d'évaluations automatisées (qualité ? sens ?)
- + si usage (mise en forme du texte d'évaluation), le signaler

D'autres questions d'éthique

Quel est le **statut**, quel est le **sens** d'un document produit (partiellement ou totalement) de manière automatique ?

Que devient l'**autonomie** (de pensée) du chercheur ?

Qu'est-ce qu'**apprendre** ?

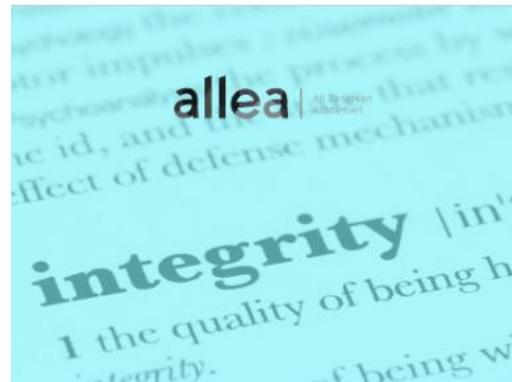
Documents utiles

L'Ofis FAIT LE POINT
OFFICE FRANÇAIS
DE L'INTÉGRITÉ SCIENTIFIQUE

Février 2024

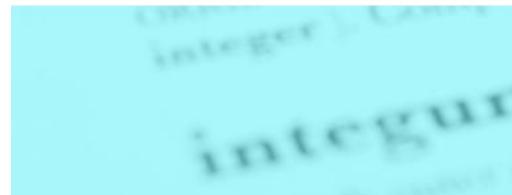
 Systèmes d'intelligence artificielle générative :
quelques points de vigilance

[lien](#)



**The European
Code of Conduct for
Research Integrity**

REVISED EDITION 2023



[lien](#)



[lien](#)