

Risques de mésinformation liés à l'utilisation des IAG pour la recherche d'information scientifique

Contexte :

Recherche d'informations scientifiques à l'aide de moteurs de recherche généralistes (Google), de moteurs de recherche spécialisés (Google Scholar, Semantic Scholar) mais aussi de moteurs de réponse (ChatGPT, Persplexity), aussi appelés systèmes de recherche conversationnels ou moteurs de recherche génératifs (Sharma et al., 2024 ; Zhou & Li, 2024).

Degré de maîtrise faible à bon, chez les étudiants du supérieur, des bases de données et des moteurs de recherche scientifiques (Soung & Dumouchel, 2019).

Défiance des éditeurs (p. ex. presse généraliste) conduisant au blocage des robots d'exploration collectant les données dédiées à l'entraînement des IAG (Viseur et Delcoucq, 2024), ce qui favorise les hallucinations et les biais, et augmente le risque épistémique de « *botshit* » (Hannigan et al., 2024).

Notre question de recherche :

Quel est le risque de mésinformation scientifique associé aux agents conversationnels génératifs type ChatGPT dans un contexte d'accès aux données de qualité handicapé par le conflit entre producteurs d'IAG et éditeurs de contenus ?

Notre méthodologie :

Calcul des taux de blocage (via l'analyse du *robots.txt*) et du biais global (Viseur et Delcoucq, 2024 ; Sun et al., 2007) associés aux revues prédatrices de la liste de Beall et aux revues de niveau 0, 1 et 2 issues de la « *Norwegian Register for Scientific Journals, Series and Publishers* ».

Nos hypothèses :

H1 : les robots des IA génératives sont davantage bloqués que les robots des moteurs de recherche.

H2 : le robot GPTbot est davantage bloqué que les robots d'autres IA génératives.

H3 : les éditeurs scientifiques commerciaux dominants bloquent davantage les robots d'IA génératives que les autres éditeurs.

H4 : les revues prédatrices bloquent moins les robots d'IA génératives que les revues non prédatrices.

H5 : mieux une revue scientifique est classée et plus elle bloque les robots d'IA génératives.

Nos résultats :

Les cinq hypothèses sont vérifiées.

Robot	Citations	Blocages	Taux de blocage (si robot cité)	Taux de blocage (tous les sites)
googlebot	689	9	1,31 %	0,08 %
bingbot	367	26	7,08 %	0,23 %
ccbot	763	666	87,29 %	5,96 %
gptbot	921	834	90,55 %	7,46 %
chatgpt-user	679	603	88,21 %	5,39 %
google-extended	720	657	91,25 %	5,88 %

Robot	Taux de blocage (ccbot)	Taux de blocage (google-extended)	Taux de blocage (gptbot)
Revue prédatrice	0 %	0 %	0,18 %
Revue de niveau 0	1,89 %	1,77 %	3,39 %
Revue de niveau 1	5,36 %	5,33 %	7,04 %
Revue de niveau 2	15,08 %	17,88 %	27,37 %
Top 50	19,35 %	32,26 %	48,39 %

Notre analyse des résultats :

Risque de mésinformation scientifique prenant, soit la forme d'hallucinations (Sun et al., 2024 ; Ye et al., 2023), soit celle de biais (Ferrara, 2023), lié :

- à la qualité réduite des données brutes d'entraînement compte tenu des difficultés d'accéder à des données de qualité (blocage) ainsi que de la facilité d'accéder à des contenus prédateurs ou d'un degré moindre de finition (p. ex. *preprints*) (= notre apport principal) ;
- au délai de production et aux limitations intrinsèques des grands modèles de langage (car pas de compréhension du contenu et fonctionnement statistique ; Hannigan et al., 2024) ;
- aux erreurs et aux biais de confirmation éventuellement induits lors de l'encodage du *prompt* (Sharma et al., 2024).

Recommandations :

Formation des étudiants (BMD) à l'écriture de *prompts* (neutralité, structuration).

Adoption de mesures de contrôle des réponses proportionnées au risque de « *botshit* » (cf. Hannigan et al., 2024, pour une typologie de risques).

Préférence pour des outils dédiés à la recherche d'information scientifique assistée par IA, soit commerciaux (Consensus, Scispace, Elicit), soit promus par les scientifiques (Artirev).

Dr Ir Robert VISEUR <robert.viseur@umons.ac.be>

Service TIC (FWEG)
Université de Mons



Journées MACCA 2025, 9-10 janvier 2025, Évry-Courcouronnes