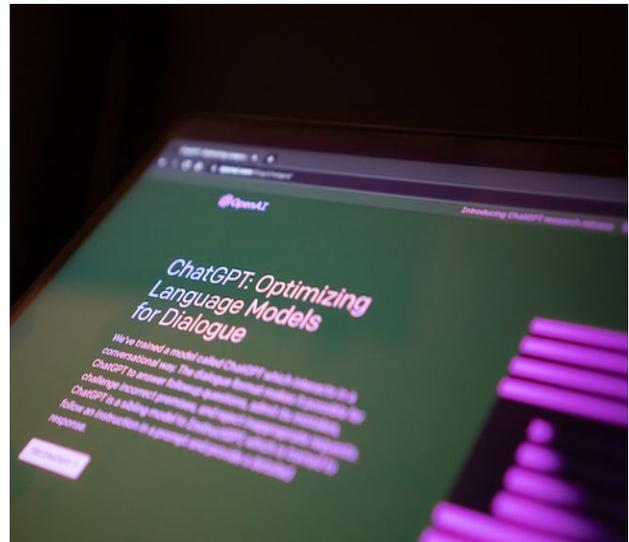


Contexte

L'apparition en accès libre du chatbot ChatGPT et la sortie le 14 mars de GPT4, forcent la communauté des scientifiques à s'interroger : quelles seront les répercussions des outils d'intelligence artificielle conversationnelle sur les pratiques de recherche ? Une question concrète -un programme d'intelligence artificielle peut-il être reconnu comme auteur d'une publication scientifique ?- a déjà conduit les éditeurs à se prononcer.



© Jonathan Kemper

Mi-janvier dernier, le journal *Nature* alertait sur les risques liés à l'utilisation de ChatGPT pour écrire des articles scientifiques en mettant en avant les résultats d'une étude, encore sous forme de *preprint*. Une équipe de l'université *Northwestern* à Chicago y montrait que ChatGPT pouvait non seulement produire des *abstracts* d'articles scientifiques de recherche médicale tout à fait convaincants, mais surtout que ces textes pouvaient duper des chercheurs : cette équipe a en effet fait écrire une cinquantaine d'*abstracts* par ChatGPT-3 à partir d'une sélection d'articles publiés dans *JAMA*, *The New England Journal*, *The BMJ*, *The Lancet* et *Nature Medicine*. Les *abstracts* ont été ensuite soumis à une révision par les pairs et seuls 68 % d'entre eux ont été détectés. De plus, l'outil conversationnel a déjà été crédité comme auteur dans quelques articles publiés.

Prises de position des éditeurs

Des éditeurs scientifiques ou associations et groupes d'éditeurs ont réagi rapidement. Certains d'entre eux se sont positionnés en mettant à jour leur notice de politique éditoriale ou en publiant une déclaration (*Nature*, *Science*, *WAME*, *Jama*, *COPE* etc.). Si les arguments varient un peu, pour l'heure, tous s'accordent globalement à considérer qu'une intelligence artificielle ne peut être considérée comme un auteur d'une publication scientifique et demandent aux chercheurs qui utilisent ces outils d'en expliciter l'usage. Voici un échantillon de premières prises de position :



Nature - Dans un éditorial du 24 janvier 2023, *Tools such as ChatGPT threaten transparent science; here are our ground rules for their use*, annonce la mise à jour de sa notice pour les auteurs. Deux nouvelles règles y figurent et valent pour tous les journaux du groupe *Springer Nature*. La première stipule que ChatGPT comme tout autre modèle de langage de ce type (LLM : *Large Language Model*), ne peut pas être reconnu comme co-auteur d'un article. Parce que, précise la notice, « ils ne satisfont pas actuellement à nos critères d'autorat. Notamment, l'attribution de l'autorat implique la responsabilité, ce qui ne peut pas être appliqué aux LLM ». La deuxième règle exige que l'utilisation d'un outil de ce type soit correctement documentée dans la section « Méthodes du manuscrit » (ou dans une partie similaire appropriée, si cette section n'existe pas).

Science - *Science* est un peu plus radical que *Nature* dans la mise à jour de sa politique éditoriale en matière d'I.A. puisqu'aucun texte généré par un outil d'I.A., comme ChatGPT, pas plus que les figures, images, graphiques produites par ces mêmes outils, ne peut être publié dans *Science*, sans l'autorisation préalable des éditeurs. Le journal va même au-delà. « Toute violation de cette politique constitue une faute scientifique ». « Au même titre que la modification d'images ou le plagiat de travaux existants » renchérit Holden Thorp, son rédacteur en chef dans un éditorial du 27 janvier 2023 *ChatGPT is fun but not an author*. Il précise néanmoins que les nombreuses données générées par des outils d'I.A. dans le cadre du travail de recherche (mais qui ne constituent pas le texte de l'article) ne sont pas visées par cette nouvelle politique.

WAME- *World Association for Medical Editors*. Dans sa déclaration publiée le 20 janvier 2023, *Chatbots, ChatGPT, and Scholarly Manuscripts*, l'association WAME, qui regroupe des éditeurs de revues médicales, insiste sur le fait que les chatbots n'ont pas de personnalité juridique et s'inquiète particulièrement du faible pourcentage de références correctes dans les articles qu'ils produisent. WAME donne 4 lignes directrices :

1. Les chatbots ne peuvent pas être des auteurs.
2. Les auteurs doivent faire preuve de transparence lorsqu'ils utilisent des chatbots et fournir des informations sur la manière dont ils ont été utilisés.
3. Les auteurs sont responsables du travail effectué par un chatbot dans leur article (y compris de l'exactitude de ce qui est présenté et de l'absence de plagiat) et de l'attribution appropriée de toutes les sources (y compris pour le matériel produit par le chatbot).
4. Les éditeurs ont besoin d'outils appropriés pour les aider à détecter le contenu généré ou modifié par l'I.A. et ces outils doivent être accessibles pour tous.

Détecter ChatGPT

Sur ce dernier point, le texte mentionne les développements en cours, notamment ceux de la plateforme *STM Integrity Hub* ou des projets tels que le *Public Knowledge Project*, qui vise à intégrer des outils de détection automatique dans des logiciels d'édition libres, à disposition de tous les éditeurs.

COPE -Committee on Publication Ethics. Outre des éditeurs, COPE regroupe aussi des universités, des organismes de recherche et publie des guides de bonnes pratiques en matière de publication scientifique qui font foi. COPE a emboîté le pas le 23 février 2023, en s'alignant sur les déclarations de *WAME* ou encore de *JAMA, Nonhuman "Authors" and Implications for the Integrity of Scientific Publication and Medical Knowledge*.

Son texte, *Artificial intelligence and authorship*, est donc du même acabit : « Les auteurs qui utilisent des outils d'I.A. pour la rédaction d'un manuscrit, la production d'images ou d'éléments graphiques de l'article, ou la collecte et l'analyse de données, doivent faire preuve de transparence en indiquant dans la section "Matériel et méthodes" (ou une section similaire) de l'article comment l'outil d'I.A. a été utilisé et quel outil a été utilisé. Les auteurs sont entièrement responsables du contenu de leur manuscrit, même des parties produites par un outil d'I.A., et sont donc responsables de tout manquement à l'éthique de la publication. »

Pistes de réflexion

Au-delà de la question du statut d'auteur et de celle de la transparence sur l'utilisation du chatbot qui demandaient d'adopter rapidement des politiques explicites, c'est une réflexion globale qu'appelle de ses vœux une partie de la communauté scientifique. Une réflexion sur les changements de pratiques de recherche que ces outils ne manqueront pas d'entraîner.

Dans leur article « *ChatGPT: five priorities for research* », des chercheurs et chercheuses de l'université d'Amsterdam portent cette voix. « Plutôt que de s'engager dans une course à l'armement futile entre les chatbots et les détecteurs de chatbot, mieux vaut réfléchir à comment utiliser ces chatbots avec intégrité, transparence et honnêteté. Cela défie les définitions binaires de l'autorat, du plagiat ou des données sources. Les politiques devront s'adapter, mais la transparence restera clé ».

Pour sortir de l'ornière actuelle du manque de transparence de ces outils (codes sources, données), ils préconisent donc que les universités, les ONG, les organisations telles que les Nations-Unies, les géants de la technologie, les éditeurs etc. investissent massivement dans un projet en libre accès, indépendant et à but non lucratif. Et pour engager la communauté des chercheurs dans un débat sur l'ensemble des implications de cette technologie, ils proposent 5 axes de réflexions :

1. Maintenir la vérification des résultats par un humain
2. Elaborer de règles de responsabilité
3. Investir dans des technologies (modèles de langage) véritablement ouverts
4. Profiter des avantages de l'intelligence artificielle
5. Elargir le débat

Ils appellent de leurs vœux l'organisation d'un forum international sur le développement et l'utilisation responsable des outils d'intelligence artificielle conversationnelle pour la recherche avec tous les acteurs pour à terme déboucher sur des directives concrètes et pratiques pour les chercheurs et chercheuses.