

« Il devient urgent de réfléchir à ce que l'IA fait à la recherche académique »

Tribune [Antonin Broi Philosophe](#) , [Thibaut Giraud Philosophe](#)

La vitesse à laquelle progressent les capacités de l'intelligence artificielle annonce un
5 bouleversement dans tous les champs du travail intellectuel, estiment, dans une tribune au
« Monde », les philosophes Antonin Broi et Thibaut Giraud, alias Monsieur Phi.

Lundi 10 août, la revue *Philosophy & Public Affairs*, l'une des plus prestigieuses qui soient en
matière de philosophie morale et politique, publiait un article défendant un argument inédit contre
10 l'épistocratie, ce régime où le pouvoir reviendrait aux plus compétents. Cette nouvelle n'aurait dû
intéresser que quelques spécialistes, la philosophie académique internationale publiant des milliers
d'articles chaque année. Sauf que l'essentiel de ce texte a été pensé et rédigé par une intelligence
artificielle (IA).

L'auteur signataire, Simon Goldstein, professeur à l'université de Hongkong [*Chine*], ne s'en cache
15 pas : il a fourni la thèse principale, développé une partie des arguments, corrigé des erreurs et validé
les versions successives. En résumé, ainsi qu'il l'affirme lui-même, il s'est comporté comme un
« directeur de thèse très généreux » envers Claude Opus 4.7 et Fable 5, les modèles d'Anthropic,
qui se sont occupés du reste. La contribution de ces grands modèles de langage (LLM, pour « large
20 langage model ») n'apparaît pourtant que dans les remerciements à la fin du texte, la revue
interdisant qu'une IA puisse avoir le statut d'auteur.

Pour qui suit les progrès récents des LLM, l'événement n'a rien de stupéfiant. Beaucoup de
chercheurs, y compris en philosophie, leur confient des opérations de plus en plus variées. La
parution de cet article marque toutefois un basculement, largement commenté au sein du monde
académique : l'IA ne représente plus seulement un outil d'appoint, elle est devenue le principal
25 rédacteur d'un travail de recherche que l'institution académique a jugé digne de l'une de ses revues
les plus sélectives. Certes, M. Goldstein est resté aux commandes, mais l'IA se taille une part
toujours plus grande dans le partage des tâches.

On ne peut plus balayer ces progrès d'un revers de la main. Loin d'un emballement médiatique
orchestré par les grandes entreprises américaines, ils annoncent des bouleversements majeurs. Les
30 mathématiques semblent avoir atteint un tournant, après quelques années de progression à une
vitesse stupéfiante. En 2023, les meilleurs modèles se trompaient encore couramment sur des
problèmes qu'un élève d'école primaire peut résoudre. En 2025, ils remportaient l'or aux
Olympiades internationales de mathématiques, égalant le niveau des meilleurs étudiants de licence.
En 2026, des dizaines d'avancées mathématiques majeures sur des problèmes de recherche ouverts
35 prestigieux ont été obtenues grâce à des modèles d'OpenAI et d'Anthropic.

En réaction, a paru, le 2 juin, la déclaration de Leiden [*Leyde (Pays-Bas)*], soutenue par l'Union
mathématique internationale, qui a recueilli plus de 1 000 signatures en vingt-quatre heures, dont
celles de l'Australien Terence Tao et de l'Allemand Peter Scholze, tous deux lauréats de la médaille
Fields. Elle expose l'ampleur des défis que pose l'irruption de l'IA dans la recherche en
40 mathématiques. A court terme, une crainte majeure porte sur l'engorgement concernant la
vérification des preuves. Les mathématiciens réclament la transparence sur l'usage des outils et le
maintien de la responsabilité humaine quant aux résultats publiés.

La philosophie académique court également le risque d'un tel engorgement, mais ce sont les
conséquences à plus long terme qui donnent le vertige. Si la philosophie et les mathématiques ont

45 longtemps été perçues comme des domaines qui sollicitent l'intelligence humaine par excellence, voir des IA apporter des contributions décisives dans ces domaines fait vaciller ce statut et questionne plus généralement la nature de l'intelligence, qu'elle soit humaine ou artificielle.

Comment devons-nous, en tant que philosophes, accueillir ces apports de l'IA à nos champs de recherche ? Pour les uns, la philosophie explore un espace de raison abstrait, similaire en cela aux
50 mathématiques, et le recours aux IA ne serait qu'un moyen d'accélérer l'exploration et la cartographie de cet espace. Cependant, cela revient à s'engager sur une voie où les philosophes humains, devenus moins efficaces que leurs collègues artificiels, pourraient finir marginalisés dans la pratique même de la discipline. Pour les autres, la philosophie reste inséparable de l'expérience d'êtres humains situés dans le monde. Pierre Hadot, dans *Qu'est-ce que la philosophie antique* ?
55 (Gallimard, 1995), défendait ainsi l'idée d'une philosophie comme une manière de vivre, un exercice destiné avant tout à transformer celui qui le pratique. Dès lors, une réflexion philosophique déléguée à une machine n'aurait aucun sens.

Plutôt que de trancher globalement la question de ce qui restera propre à l'humain, il convient sans doute de l'aborder à un niveau plus fin. Prenons, par exemple, une question classique en
60 philosophie de l'esprit : que pouvons-nous apprendre en « introspectant » notre expérience subjective ? A première vue, il semble que l'être humain conserve un net avantage sur l'IA, dans la mesure où il est le seul à pouvoir « introspecter » sa propre expérience.

Pourtant, la question de l'introspection se pose aussi pour l'IA. De plus en plus d'études se penchent sur les capacités introspectives et la représentation de soi qui émergent dans les grands
65 modèles de langage. Ceux-ci pourraient avoir une place privilégiée pour s'interroger, à la première personne, sur le vertige de cette introspection artificielle. De même que les êtres humains sont, par définition, les mieux à même de réfléchir à la condition humaine, des « philosophes artificiels » pourraient être les mieux placés pour aborder les questionnements liés à la condition d'IA.

Au-delà du cas particulier de la philosophie, il devient urgent de réfléchir dès maintenant à ce que
70 l'IA fait à la recherche académique et, plus généralement, au travail intellectuel. En France, cette réflexion peine à s'engager, freinée par le scepticisme envers les capacités des LLM qui, s'il se justifiait il y a quelques années, résiste de moins en moins à l'épreuve des faits. La vitesse à laquelle progressent ces capacités, plus que leur niveau actuel, laisse bel et bien augurer un bouleversement dans tous les champs du travail intellectuel.

75 **Antonin Broi** est docteur en philosophie et chercheur à l'université de Genève (Suisse) ; **Thibaut Giraud** est docteur en philosophie et créateur de la chaîne YouTube de vulgarisation philosophique Monsieur Phi. Il est l'auteur de l'ouvrage « La Parole aux machines. Philosophie des grands modèles de langage » (Grasset, 2025).

80 [Antonin Broi \(Philosophe\)](#) et [Thibaut Giraud \(Philosophe\)](#)

https://www.lemonde.fr/idees/article/2026/08/29/il-devient-urgent-de-reflechir-a-ce-que-l-ia-fait-a-la-recherche-academique_6760022_3232.html