

Compétence transversale – L1

Grands défis sociétaux : Intelligence Artificielle

Pr. Lucile Sassatelli

Professeure des Universités en Informatique, UniCA

Directrice scientifique de EFELIA Côte d'Azur


Image by Alan Warburton / © BBC / Better Images of AI / Nature / CC-BY 4.0

Plan du module

Chapitre	Titre	Contenu	Date d'ouverture	Date QCM
1	Qu'est-ce que l'IA ?	<ul style="list-style-type: none">• Domaine et familles de méthodes• Applications actuelles• Définition d'un modèle d'IA, choix humains dans la conception• Mythe de l'AGI et confrontation à la réalité des neurosciences		
2	Qu'est-ce que les grands modèles de langue (LLM) ?	<ul style="list-style-type: none">• Principe des modèles génératifs de texte• Choix simplificateurs et conséquences : faillibilité, biais		<ul style="list-style-type: none">• QCM 1 noté 3-7/11
3	Quels impacts sur la société ? Quels impacts sur moi ?	<ul style="list-style-type: none">• Exploitation humaine et emplois• Exploitation environnementale• Apprentissage à l'université		
4	Quels impacts sur ma discipline ?	<ul style="list-style-type: none">• Les intérêts des systèmes d'IA• Les enjeux des systèmes d'IA pour son domaine		<ul style="list-style-type: none">• QCM 2 note 8-12/12



Plan

- 
1. Exploitation humaine
 2. Exploitation environnementale
 3. Votre apprentissage à l'université
 4. Récap'



Exploitation humaine (1)

- Le travail actif joue un rôle majeur, bien que méconnu, dans le développement de l'IA :
 - humains pour exécuter des tâches répétitives et difficiles à automatiser, et pourtant essentielles
 - dont modération : pas de façon efficace de purger les pans entiers de biais et toxicité dans les données
- Détecteur de toxicité construit pour ChatGPT :
 - Kenya, Ouganda, Inde
 - Pédophilie, torture, suicides,
 - Conditions précaires : \$1.32-\$2 l'heure
 - Travail traumatisant

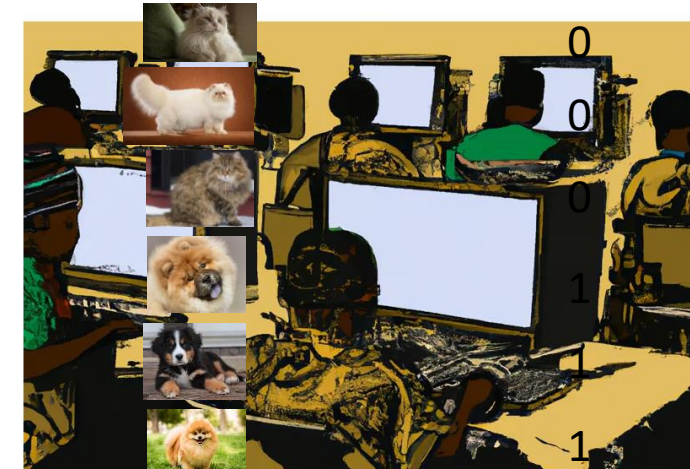
→ Essentiel pour l'industrie de l'IA mais peu connu : toujours le même modèle d'exploitation ?

TIME

SIGN UP FOR OUR IDEAS NEWSLETTER POV

BUSINESS • TECHNOLOGY

Un jeu de données d'entraînement
Exclusive: OpenAI Used Kenyan Workers on Less Than \$2 Per Hour to Make ChatGPT Less Toxic (« label »)
(0: chat, 1: chien)

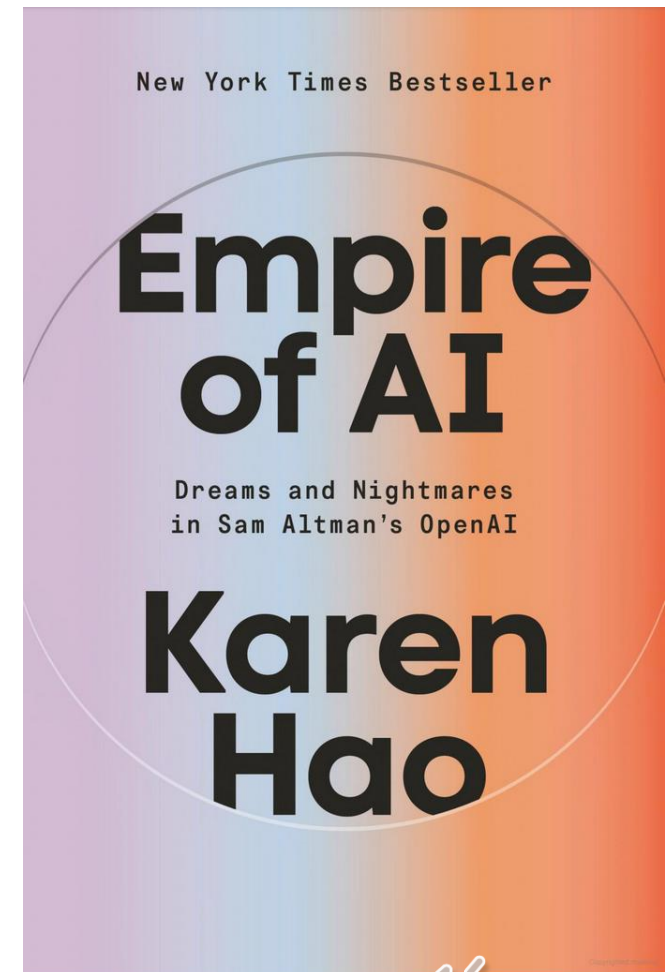


This image was generated by OpenAI's image-generation software, Dall-E 2. The prompt was: "A seemingly endless view of African workers at desks in front of computer screens in a printmaking style." TIME does not typically use AI-generated art to illustrate its stories, but chose to in this instance in order to draw attention to the power of OpenAI's technology and shed light on the labor that makes it possible. Image generated by Dall-E 2/OpenAI

- [Travailleurs du Clic](#). Documentaire France TV, 2022.
- Paola Tubaro. [Les « petites mains » derrière le mirage de l'automatisation](#). La vie de la recherche scientifique, 2020, pp.76-77.
- TIME. [Exclusive: OpenAI Used Kenyan Workers on Less Than \\$2 Per Hour to Make ChatGPT Less Toxic](#). Avril 2023.

Au Kenya, Mr Fatokine

- Mo Fatokhine, l'un des employés kenyans embauchés par OpenAI, ne savait d'ailleurs pas qu'il travaillait pour OpenAI à l'origine. Il ne l'a appris que grâce à une fuite d'un de ses supérieurs.
- Et quand il a commencé à travailler dans l'équipe chargée des contenus sexuels, sa personnalité a complètement changé. Il n'arrivait pas à expliquer à sa femme pourquoi, car il ne savait pas comment lui dire : « Je lis des contenus sexuels toute la journée. » Ça n'a pas l'air d'un vrai travail. La société n'avait même aucune idée de ce que cela signifiait. Et un jour elle est partie avec leur fille et lui a envoyé un texto pour lui dire : « Je ne comprends pas l'homme que tu es devenu, et je ne reviendrai pas. »
- Il est essentiel de comprendre qu'il ne s'agit pas d'une forme de travail nécessaire. La Silicon Valley prétendra que ce travail est nécessaire, mais il ne l'est que sur la base de son principe de travailler avec des très grands modèles nécessitant des jeux de données gigantesques, nécessairement constitués de données violentes et non filtrées.
- Le frère de Mo Fatokine était écrivain et pigiste, et quand ChatGPT a été sorti, il a commencé à perdre ses contrats.



Karen Hao, Penguin Press Eds., 2025
<https://www.penguinrandomhouse.com/books/743569/empire-of-ai-by-karen-hao/>

Exploitation humaine (2)

- Les données d'entraînement sont souvent collectées illégalement
- L'art par l'IA générative : industrie estimée à 48 milliards de dollars.
- Mais des artistes professionnel·les ont dénoncé les préjudices qu'ils subissent :
 - atteintes à la renommée, les pertes économiques, le plagiat et la violation du droit d'auteur
 - L'article fournit des recommandations : réglementation obligeant les organisations à divulguer leurs données d'entraînement, outils aidant les artistes à éviter que leur contenu soit utilisé comme données d'entraînement sans leur consentement.



- O. Darcy, "[News publishers sound alarm on Google's new AI-infused search, warn of 'catastrophic' impacts](#)," CNN Business, May 2024.
- K. Wiggers, "[OpenAI accidentally deleted potential evidence in NY Times copyright lawsuit](#)," Tech Crunch, Nov. 2024.
- H. Jiang et al., "[AI Art and its Impact on Artists](#)," in AAAI/ACM Conference on AI, Ethics, and Society (AIES), 2023



Exploitation humaine (2)

- Les données collectées
- L'art par milliards
- Mais de les préjuger
 - atteindre plagiat
- L'article oblige d'entretenir leur contenu

Beaucoup de données → Travail humain

Beaucoup de calculs → Environnement



Plan

1. Exploitation humaine
- ➔ 2. Exploitation environnementale
3. Votre apprentissage à l'université
4. Récap'



Exploitation environnementale (1)

- GPT4: \$78M, Gemini: \$191M
- Eq. CO2 d'entraîner LLM :
 - ~500 allers-retours Paris<->New York



HAI, "[2024 AI Index Report](#)", Stanford University, 2024.

Emma Strubell et al., "[Energy and Policy Considerations for Deep Learning in NLP](#)," In Proc. Assoc. for Computational Linguistics, 2019.

D. Patterson et al., "Carbon Emissions and Large Neural Network Training", 2023.

"[AI is already wreaking havoc on global power systems](#)", Bloomberg, June 2024.

Sasha Luccioni, "[Vers l'évaluation et l'atténuation de l'impact environnemental des grands modèles de langues](#)", rapport CIFAR, Sep. 2023.

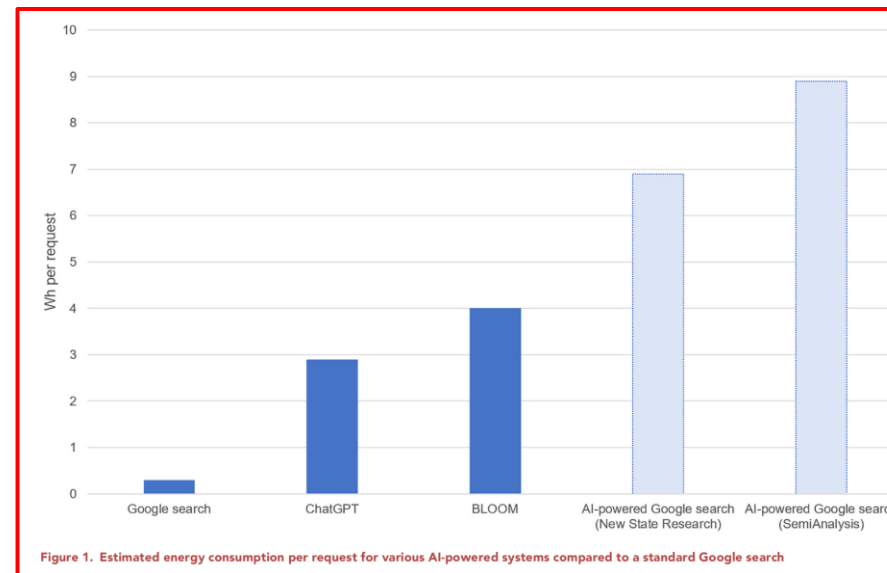
Kate Crawford. [Contre-atlas de l'intelligence artificielle](#). Zulma éditions, 2022, 384 pages.

A. De Vries, "[The growing energy footprint of artificial intelligence](#)," Joule, vol. 7, no. 10, pp. 2191–2194, Oct. 2023.



Exploitation environnementale (2)

- Usage : énergie x10 à x 30 par rapport à une requête à Google



©A. De Vries

HAI, [“2024 AI Index Report”](#), Stanford University, 2024.

Emma Strubell et al., [“Energy and Policy Considerations for Deep Learning in NLP,”](#) In Proc. Assoc. for Computational Linguistics, 2019.

D. Patterson et al., “Carbon Emissions and Large Neural Network Training”, 2023.

[“AI is already wreaking havoc on global power systems”](#), Bloomberg, June 2024.

Sasha Luccioni, [Vers l'évaluation et l'atténuation de l'impact environnemental des grands modèles de langues](#), rapport CIFAR, Sep. 2023.

Kate Crawford. [Contre-atlas de l'intelligence artificielle](#). Zulma éditions, 2022, 384 pages.

A. De Vries, [“The growing energy footprint of artificial intelligence,”](#) Joule, vol. 7, no. 10, pp. 2191–2194, Oct. 2023.



Exploitation environnementale (2)

- Croissance des centres de données
 - Consommation d'électricité
 - Contrats des GAFAM avec les industries fossiles et nucléaires

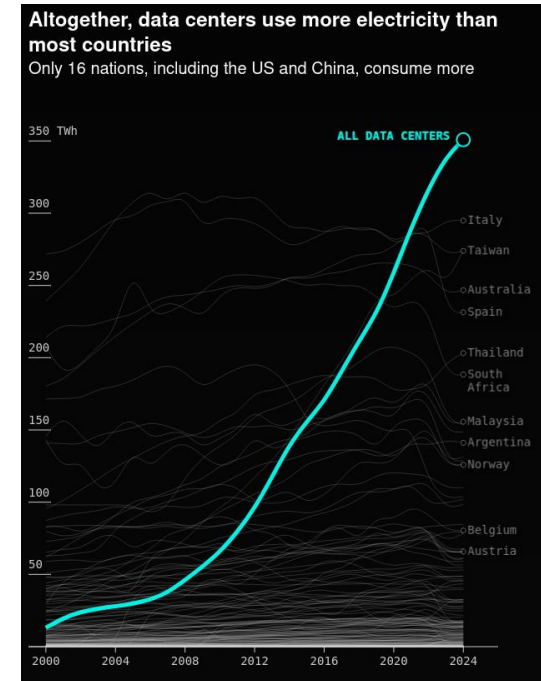
- « *Le problème est simple : les estimations de la consommation en électricité des centres de données ne cessent de croître. Celle-ci pourrait être multipliée par quatre en France d'ici dix ans, et atteindre 7,5 % de la demande totale de courant électrique, soit plus du double de la part des transports aujourd'hui – contre 2 % actuellement .* »



IA : un rapport alerte sur le risque de conflit d'usage de l'électricité

Les projets de centres de données numériques sont si nombreux et si énormes que leur consommation d'électricité pourrait être multipliée par quatre d'ici dix ans en France, selon un rapport du Shift Project. Au risque de menacer la décarbonation des transports et de l'industrie.

Jade Lindgaard - 1 octobre 2025 à 07h47

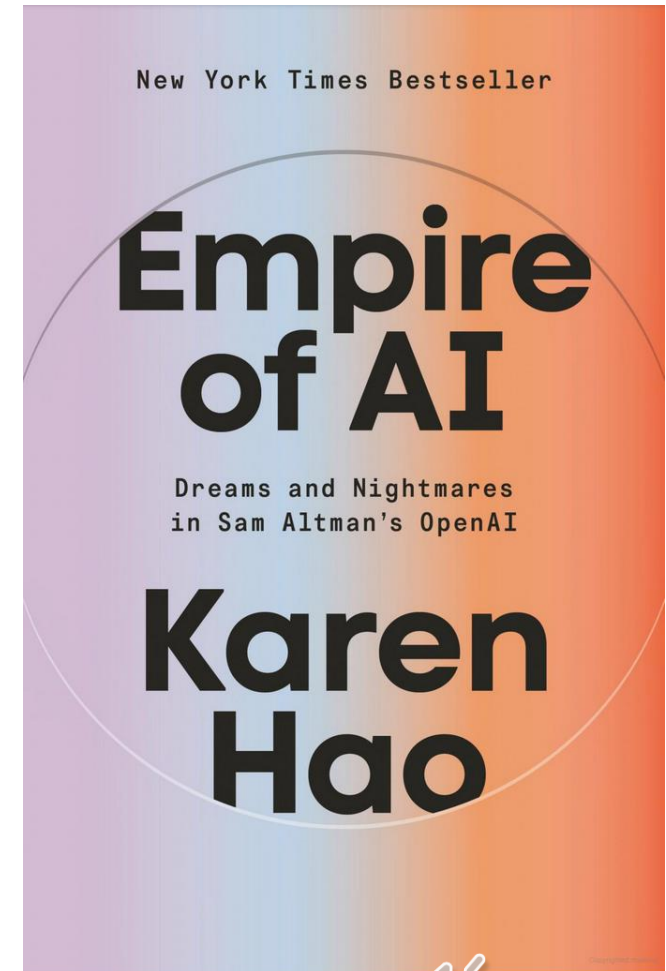


- Consommation d'eau



En Uruguay, de l'eau contaminée par un datacenter, et bue

- La majeure partie de cette énergie sera alimentée par les combustibles fossiles. On signale déjà des prolongations de durée de vie des centrales à charbon. Elon Musk a construit son gigantesque supercalculateur Colossus à Memphis, dans le Tennessee, et l'alimente grâce à environ 35 centrales au méthane non autorisées qui rejettent des milliers de tonnes de polluants atmosphériques dans ces communautés.
- Il s'agit donc d'une crise climatique, d'une crise de santé publique et également d'une crise d'eau potable, car bon nombre de ces centres de données s'installent dans des communautés et doivent être refroidis avec de l'eau potable, pas avec autre chose car cela pourrait entraîner la corrosion des équipements et favoriser la croissance bactérienne.
- Et le plus souvent, l'alimentation en eau des datacenters est prise dans le réseau public d'eau potable, car c'est l'infrastructure pré-existante.
- Google a décidé d'installer un centre de données à Montevideo, la capitale de l'Uruguay. La région était confrontée à une sécheresse historique, à tel point que le gouvernement de Montevideo a commencé à mélanger de l'eau toxique sortant du datacenter au réseau public d'eau potable, simplement pour que de l'eau sorte du robinet des gens. Les personnes trop pauvres pour acheter de l'eau en bouteille étaient contraintes de boire cette eau toxique, et des femmes ont fait des fausses couches.
- Nous assistons donc à l'amplification de nombreuses crises croisées et à la perpétuation de ce paradigme du plus grand modèle, plus de calculs, de données, à tout prix.



Karen Hao, Penguin Press Eds., 2025
<https://www.penguinrandomhouse.com/books/743569/empire-of-ai-by-karen-hao/>

D'ailleurs en parlant de fiabilité et non accès au sens physique....

The screenshot shows the Google Translate interface in a Mozilla Firefox browser. The source text in English is: "Google decided to put a data center in Montevideo, Uruguay's capital. The area was literally facing a historic drought to the point where the Montevideo government started mixing toxic water into the public drinking water supply, simply to have something come out of people's taps. And people who were too poor to buy bottled water had to just drink that toxic water and women were having miscarriages." The translated text in French is: "Google a décidé d'installer un centre de données à Montevideo, la capitale de l'Uruguay. La région était confrontée à une sécheresse historique, à tel point que le gouvernement de Montevideo a commencé à mélanger de l'eau toxique au réseau public d'eau potable, simplement pour que l'eau du robinet soit potable. Les personnes trop pauvres pour acheter de l'eau en bouteille étaient contraintes de boire cette eau toxique, et des femmes faisaient des fausses couches." The interface includes tabs for Text, Images, Documents, and Websites, language selection options (English - Detected, French), and a 'Send feedback' link.



- [...] les LLM dont les paramètres s'étendent désormais largement à des milliers de milliards (10^{12}). À titre indicatif, cela reste 1000 fois moins que l'ordre du nombre de connexions neuronales du cerveau humain (10^{15}), et pourtant, ils consomment environ cent millions de fois plus d'énergie (des GWatts contre les très modestes 20 watts requis par notre cerveau). Étant donné que de tels chiffres gargantuesques exigent des centrales nucléaires pour répondre aux besoins actuels de l'industrie moderne de l'IA, [...] La question est simple : quel retour obtenons-nous pour une demande énergétique aussi insatiable ?
- Les outils d'IA de toutes les entreprises technologiques peuvent certes continuer à nous surprendre, mais cela ne change rien au fait qu'ils commettent encore de nombreuses erreurs (et elles le disent ouvertement dans les petits caractères qu'elles fournissent aux utilisatrices). Ces erreurs sont bien trop nombreuses pour les standards d'exactitude requis dans la plupart des domaines scientifiques et dans de nombreux autres domaines professionnels tels que les domaines juridique et éducatif.

arXiv:2507.19703v2 [cs.AI] 30 Jul 2025

The wall confronting large language models

P.V. Coveney^{1,2,3} and S. Succi^{4,5,6}¹Centre for Computational Science, University College London, London WC1H 0AJ, U.K.²Advanced Research Computing Centre, University College London, London WC1H 0AJ, U.K.³Institute for Informatics, Faculty of Science, University of Amsterdam, 1098XH Amsterdam, The Netherlands⁴Italian Institute of Technology, Viale Regina Elena 291, 00161 Rome, Italy⁵Physics Department, Harvard University, Cambridge USA⁶Mechanical Engineering Department, University College London, London WC1H 0AJ, U.K.

July 31, 2025

Abstract

We show that the scaling laws which determine the performance of large language models (LLMs) severely limit their ability to improve the uncertainty of their predictions. As a result, raising their reliability to meet the standards of scientific inquiry is intractable by any reasonable measure. We argue that the very mechanism which fuels much of the learning power of LLMs, namely the ability to generate non-Gaussian output distributions from Gaussian input ones, might well be at the roots of their propensity to produce error pileup, ensuing information catastrophes and degenerative AI behaviour. This tension between learning and accuracy is a likely candidate mechanism underlying the observed low values of the scaling components. It is substantially compounded by the deluge of spurious correlations pointed out by Calude and Longo which rapidly increase in any data set merely as a function of its size, regardless of its nature. The fact that a degenerative AI pathway is a very probable feature of the LLM landscape does not mean that it must inevitably arise in all future AI research. Its avoidance, which we also discuss in this paper, necessitates putting a much higher premium on insight and understanding of the structural characteristics of the problems being investigated.



Plan

1. Exploitation humaine
2. Exploitation environnementale
- ➔ 3. Votre apprentissage à l'université
 1. Inquiétudes que l'apprentissage ne se fasse plu
 2. Importance du processus d'écrire pour nous
 3. Autorisez-vous à ressentir de la frustration avec un LLM et ayez de l'estime pour votre pensée
4. Récap'



Problématique

- Il semblerait sur les outils d'IA puisse produire vite et sans effort pour nous beaucoup des textes qu'on nous demande de produire, avec une qualité paraissant satisfaisante. Alors :
 - Devez-vous encore faire l'effort d'écrire pour apprendre ?
 - Doit-on encore lire des livres et longs articles, ou des résumés sont-ils suffisants ?
 - Quels gains ou pertes possibles avec l'usage d'outils d'IA ?

Rebecca Winthrop and Maryanne Wolf, "[Rethinking School in the Age of AI](#)," Center for Humane Technology, April 2025.

Sonja Drimmer and Christopher J. Nygren, "[How We Are Not Using AI in the Classroom](#)," The Newsletter of the International Center of Medieval Art, April 2025.

Carl T. Bergstrom and Jevin D. West, "[Modern-Day Oracles or Bullshit Machines? – Lesson 11](#)," lecture UW, 2025.



Mais quel est le but de l'enseignement ?

Quel est le moyen de l'enseignement ?

- **But et Intérêt de l'élève**: apprendre, donc produire l'effort incontournable pour créer ces circuits
 - **Moyen** : note sur la production pour inciter et s'assurer que le processus a été réalisé
- l'usage de IAG court-circuite le moyen et la note devient sans rapport avec l'intérêt de l'élève

Rebecca Winthrop and Maryanne Wolf, "[Rethinking School in the Age of AI](#)," Center for Humane Technology, April 2025.

Sonja Drimmer and Christopher J. Nygren, "[How We Are Not Using AI in the Classroom](#)," The Newsletter of the International Center of Medieval Art, April 2025.

Carl T. Bergstrom and Jevin D. West, "[Modern-Day Oracles or Bullshit Machines? – Lesson 11](#)," lecture UW, 2025.



Des consignes pour l'usage des LLM... En contradiction avec les travaux scientifiques ⚠

- Les LLM sont de plus en plus utilisés par les étudiant·es, risques réels de mis-skilling et never-skilling, en plus de-skilling.
- En médecine, [Abdulnour et al, 2025] proposent donc la stratégie pédagogique DEFT-AI pour encadrer l'utilisation de l'IA : développement de l'esprit critique chez les étudiants en médecine, notamment en leur conseillant de « demander à l'IA d'expliquer son raisonnement », d'essayer plusieurs suggestions et de lui donner un retour, et de « rechercher des preuves validées par des pairs de leur exactitude et de leur sécurité ».

→ PROBLEME : Mais nous avons vu que les méthodes LLM actuelles ne permettent pas d'obtenir des réponses fiables à ces questions !

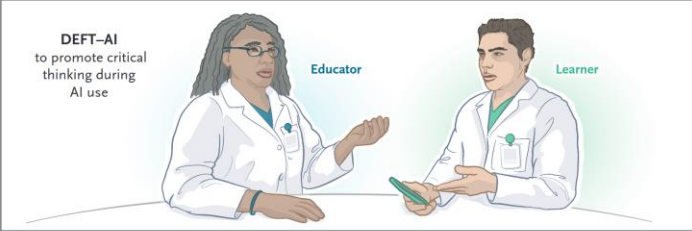
The NEW ENGLAND JOURNAL of MEDICINE

REVIEW ARTICLE | MEDICAL EDUCATION

Educational Strategies for Clinical Supervision of Artificial Intelligence Use

Authors: Raja-Elie E. Abdulnour, M.D., Brian Gin, M.D., Ph.D., and Christy K. Boscardin, Ph.D. Author Info & Affiliations

Published August 20, 2025 | N Engl J Med 2025;393:786-797 | DOI: 10.1056/NEJMra2503232 | VOL. 393 NO. 8 Copyright © 2025



DEFT-AI
to promote critical
thinking during
AI use

Educator

Learner

Diagnosis, Discussion, and Discourse	The educator asks for a description of the learner's specific use of AI.
What specific AI did you use?	I used the free version of ChatGPT on my phone.
How did you use AI in this process?	I just typed in, "What is the differential diagnosis for wheezing?"
What prompts did you enter in the app?	I asked it for the best diagnostic test and treatment strategy.

Evidence	The educator asks for an evaluation of the learner's evidence-based use of AI.
How did you verify the AI-generated outputs?	Hmm, I didn't. The answers seemed reasonable to me.
Is the AI that you used shown to be accurate and safe?	Yes. I keep seeing social media posts about how great it is at making diagnoses.

Feedback	The educator asks the learner to reflect on growth opportunities in the use of AI.
How do you evaluate your own use of AI in this case?	I think I've become quite familiar at using ChatGPT. I use it all the time now.
How can you improve your use of AI?	I can't wait for an AI that can interpret ECGs and chest radiographs. I should verify the AI outputs next time.

Teaching	The educator provides focused teaching points based on findings from the conversation and recommends whether, when, and how to use AI safely moving forward.
<p>Use AI tools that are known to be effective. Look for peer-reviewed evidence of their accuracy and safety. Our institution may have adapted and validated a similar model on the basis of high-quality data.</p> <p>Prompting a chatbot is critical to generate valuable and accurate outputs. Think of it as talking with a consultant: provide enough specific information about the Who (the intended role of the AI and your role), the Where (description of the context), and the What (your goal and specific task or question). Always ask the AI to explain its reasoning, which improves its answers and lets you assess how it is thinking and how much to trust it. One prompt is not enough: have a conversation and give it feedback. Just like I did with you, you can also ask it to engage in self-reflection and look for errors.</p> <p>AI is always prone to error and bias: always verify and trust. Make sure to check its answers against your knowledge, trusted sources of medical information, like publications from the NEJM Group, and your trusted peers, like me.</p>	

Recommendation for AI engagement	The educator provides learner-specific recommendations for the safe use of AI.
<p>Keep practicing using AI to inform your reasoning rather than replace it. AI outputs are your preliminary outputs, just like a preliminary radiology report or automated ECG interpretation: verify, then trust. Know when you can rely on it (cyborg) and when you need to confirm the outputs (centaur).</p>	

En France

- La délégation sénatoriale alerte également sur le
 - « manque de preuves scientifiques des apports pédagogiques de l'IA et de l'efficacité des outils disponibles pour faire progresser les élèves », le « risque de voir s'éroder encore davantage les compétences fondamentales (lecture et écriture, pensée critique, faculté d'autoévaluation), les capacités d'attention et de mémorisation, mais également le développement du lien social ».



Education : des enquêtes et des études récentes parmi la pression de déploiement à l'université



Intelligencer

Everyone Is Cheating Their Way Through College ChatGPT has unraveled the entire academic project.

Portrait of James D. Walsh
By James D. Walsh, Intelligencer features writer



Search

HOME LATEST TECH REVIEWS SCIENCE IO9 AI VIDEO DEALS

ARTIFICIAL INTELLIGENCE

It's Breathtaking How Fast AI Is Screwing Up the Education System

Thanks to a new breed of chatbots, American stupidity is escalating at an advanced pace.

By Lucas Ropek Published May 16, 2025 | Con

Aujourd'hui, les enseignants d'université semblent eux aussi choisir la solution de facilité, comme le montre un article récent du New York Magazine : ils utilisent des chatbots pour créer leurs plans de cours, tout comme leurs étudiants les utilisent pour répondre aux exercices du cours en question.

Le résultat de tout cela est si évident [...] : tous ceux qui utilisent l'IA vont devenir de plus en plus stupides, et plus ils le seront, plus ils auront besoin d'utiliser l'IA pour faire ce qu'ils faisaient auparavant avec leurs neurones. Le modèle « as-a-service » et basé sur l'abonnement de l'industrie technologique est ici clairement mis en évidence, sauf que l'abonnement portera sur la capacité intellectuelle.

~~404~~

ABOUT RSS SUPPORT/FAQ PODCAST FOIA FORUM ARCHIVE MERCH ADVERTISE REFERRAL PROGRAM THANKS PRIVACY

AI

Teachers Are Not OK

Que nous dit la science ?

- Des enquêtes et nouvelles études montrent des résultats néfastes pour les étudiant·es :

H. Bastani, et al., "[Generative AI Can Harm Learning](#)," *Social Science Research Network*, Rochester, NY: 4895486, July 2024.

N. Kosmyrna, et al., "[Your Brain on ChatGPT: Accumulation of Cognitive Debt when Using an AI Assistant for Essay Writing Task](#)," June 2025

M. Stadler et al., "[Cognitive ease at a cost: LLMs reduce mental effort but compromise depth in student scientific inquiry](#)," *Computers in Human Behavior*, vol. 160, p. 108386, Nov. 2024.

H.-P. (Hank) Lee *et al.*, "[The Impact of Generative AI on Critical Thinking: Self-Reported Reductions in Cognitive Effort and Confidence Effects From a Survey of Knowledge Workers](#)," in CHI 2025.

M. Abbas, et al., "[Is it harmful or helpful? Examining the causes and consequences of generative AI usage among university students](#)," *International Journal of Education Technology in Higher Education*, vol. 21, no. 1, p. 10, Feb. 2024.

Your Brain on ChatGPT: Accumulation of Cognitive Debt when Using an AI Assistant for Essay Writing Task^Δ

Nataliya Kosmyrna¹
MIT Media Lab
Cambridge, MA

Eugene Hauptmann
MIT
Cambridge, MA

Ye Tong Yuan
Wellesley College
Wellesley, MA

Jessica Situ
MIT
Cambridge, MA

Xian-Hao Liao
Mass. College of Art
and Design (MassArt)
Boston, MA

Ashly Vivian Beresnitzky
MIT
Cambridge, MA

Iris Braunstein
MIT
Cambridge, MA

Pattie Maes
MIT Media Lab
Cambridge, MA

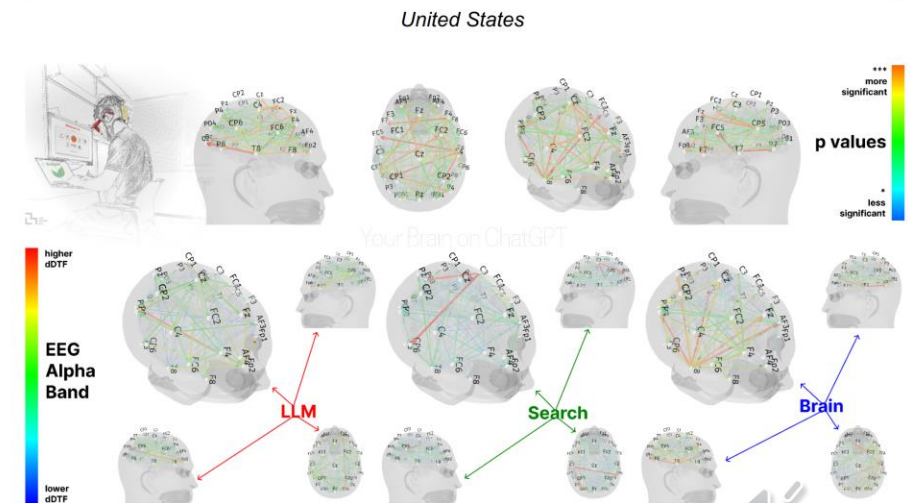


Figure 1. The dynamic Direct Transfer Function (dDTF) EEG analysis of Alpha Band for groups: LLM, Search Engine, Brain-only, including p-values to show significance from moderately significant (*) to highly significant (***).

Your Brain on ChatGPT: Accumulation of Cognitive Debt when Using an AI Assistant for Essay Writing

- Les résultats ont révélé que des niveaux plus élevés de connectivité neuronale dans le groupe « Cerveau seul » étaient corrélés à une meilleure mémoire, une plus grande précision sémantique et une plus grande appropriation du travail écrit.
- Le groupe « Cerveau seul », bien que soumis à une charge cognitive plus importante, a montré des résultats d'apprentissage plus approfondis et une plus forte appropriation à ses travaux.
- Le groupe « LLM », tout en bénéficiant de l'efficacité des outils, a montré une mémorisation et une réflexion plus faibles, et une appropriation des productions écrites plus faible.
- Ce compromis met en lumière une inquiétude éducative importante : l'usage d'outils d'IA, bien qu'intéressant pour produire certains types de textes, peut entraver le traitement cognitif profond, la rétention et l'engagement envers le texte. Si les personnes s'appuient fortement sur les outils d'IA, elles peuvent acquérir une fluidité superficielle, mais ne parviennent pas à internaliser les connaissances ni à se les approprier.

Your Brain on ChatGPT: Accumulation of Cognitive Debt when Using an AI Assistant for Essay Writing Task^Δ

Nataliya Kosmyna¹
MIT Media Lab
Cambridge, MA

Eugene Hauptmann
MIT
Cambridge, MA

Ye Tong Yuan
Wellesley College
Wellesley, MA

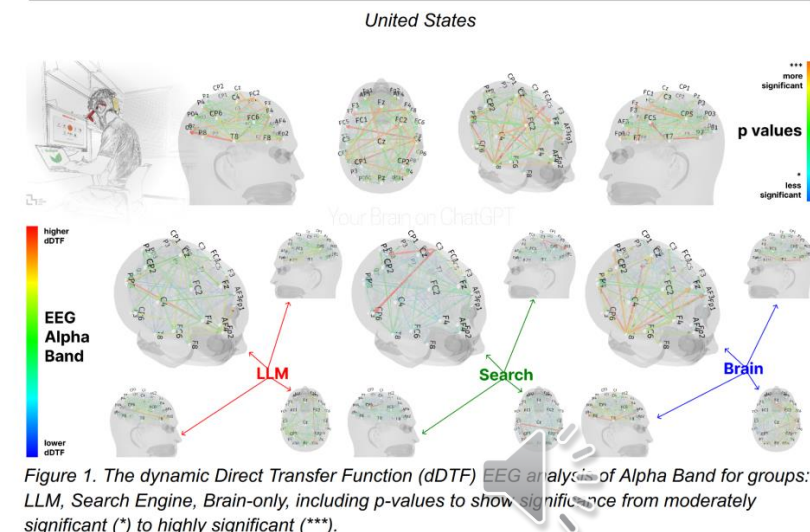
Jessica Situ
MIT
Cambridge, MA

Xian-Hao Liao
Mass. College of Art
and Design (MassArt)
Boston, MA

Ashly Vivian Beresnitzky
MIT
Cambridge, MA

Iris Braunstein
MIT
Cambridge, MA

Pattie Maes
MIT Media Lab
Cambridge, MA



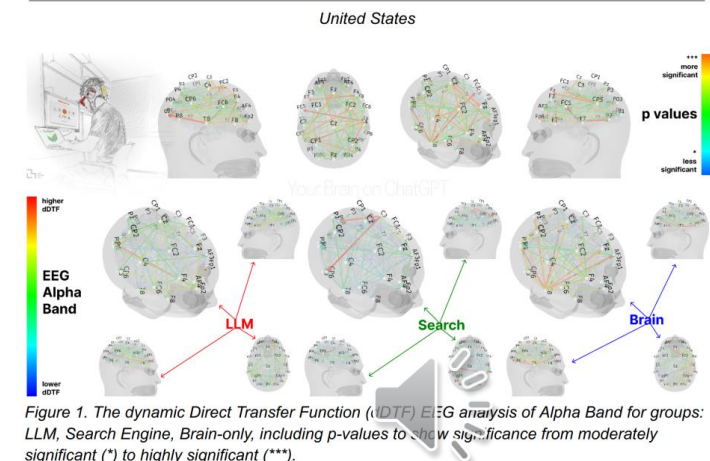
Your Brain on ChatGPT: Accumulation of Cognitive Debt when Using an AI Assistant for Essay Writing

Cette corrélation entre le manque de connectivité neuronale et l'échec de restitution chez les participants du groupe LLM prouve que :

1. Le recours précoce à l'IA pourrait entraîner un codage superficiel. La faible mémorisation et les citations incorrectes du groupe LLM pourraient indiquer que leurs textes antérieurs n'étaient pas intégrés en interne, probablement en raison d'un traitement cognitif externalisé au LLM.
2. Eviter l'usage outils LLM dans les premiers stades de l'acquisition de notions pourrait favoriser la formation de la mémoire. Le rappel comportemental plus fort du groupe Cerveau-seul, soutenu par une connectivité EEG plus robuste, suggère que l'effort initial spontané a favorisé des traces mnésiques durables, permettant une réactivation plus efficace, même lorsque les outils LLM ont été introduits ultérieurement.
3. L'engagement métacognitif est plus élevé dans le groupe « Cerveau-LLM ». Le groupe « Cerveau seul » a pu comparer mentalement ses efforts spontanés antérieurs avec les suggestions générées par les outils, s'engageant dans une autoréflexion et une répétition élaborée, un processus lié au contrôle exécutif et à l'intégration sémantique, comme l'indique son profil EEG.

Your Brain on ChatGPT: Accumulation of Cognitive Debt when Using an AI Assistant for Essay Writing Task^Δ

Nataliya Kosmyna ¹ MIT Media Lab Cambridge, MA	Eugene Hauptmann MIT Cambridge, MA	Ye Tong Yuan Wellesley College Wellesley, MA	Jessica Situ MIT Cambridge, MA
Xian-Hao Liao Mass. College of Art and Design (MassArt) Boston, MA	Ashly Vivian Beresnitzky MIT Cambridge, MA	Iris Braunstein MIT Cambridge, MA	Pattie Maes MIT Media Lab Cambridge, MA

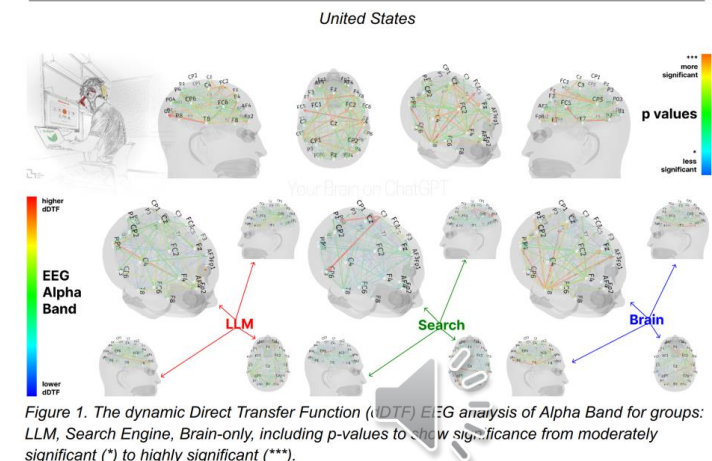


Your Brain on ChatGPT: Accumulation of Cognitive Debt when Using an AI Assistant for Essay Writing

- Lorsque les individus ne parviennent pas à aborder un sujet de manière critique, leurs écrits peuvent devenir biaisés et superficiels. Ce schéma reflète l'accumulation d'une dette cognitive :
 - un état dans lequel le recours répété à des systèmes externes comme les LLM remplace les processus cognitifs demandant des efforts nécessaires à une réflexion indépendante.
- La dette cognitive reporte l'effort mental à court terme, mais entraîne des coûts à long terme :
 - diminution de l'esprit critique, vulnérabilité accrue à la manipulation, une baisse de créativité
- Lorsque les participant·es reproduisent des suggestions sans en évaluer l'exactitude ou la pertinence, ils perdent non seulement la propriété de leurs idées, mais risquent également d'intérioriser des perspectives superficielles ou biaisées.

Your Brain on ChatGPT: Accumulation of Cognitive Debt when Using an AI Assistant for Essay Writing Task^Δ

Nataliya Kosmyna ¹ MIT Media Lab Cambridge, MA	Eugene Hauptmann MIT Cambridge, MA	Ye Tong Yuan Wellesley College Wellesley, MA	Jessica Situ MIT Cambridge, MA
Xian-Hao Liao Mass. College of Art and Design (MassArt) Boston, MA	Ashly Vivian Beresnitsky MIT Cambridge, MA	Iris Braunstein MIT Cambridge, MA	Pattie Maes MIT Media Lab Cambridge, MA



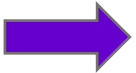
Et vous ?

- Comment abordez-vous un travail de rédaction dans vos études (par exemple un compte-rendu de TP) ?
 - Comme une opportunité de réfléchir et de vous assurer de votre compréhension, ou de la compléter si vous identifiez, en écrivant, des trous/insuffisances dans votre compréhension ?
 - Souvent comme quelque chose d'inintéressant ? Si c'est le cas, est-ce que ça dépend du sujet / de la matière ?
 - Comme autre chose ?



Plan

1. Exploitation humaine
2. Exploitation environnementale
3. Votre apprentissage à l'université
 1. Inquiétudes que l'apprentissage ne se fasse plu
 2. Importance du processus d'écrire pour nous
 3. Autorisez-vous à ressentir de la frustration avec un LLM et ayez de l'estime pour votre pensée
4. Récap'



Nos vies sont vécues dans un monde de processus, et non de résultats.

- Pour nous humains, le processus est important : ce sont nos expériences du monde qui font notre vie.
 - Ce que nous vivons, ce qui nous arrive ou ce qu'on choisit de vivre, de faire, fait notre vie, nous forme.
- Nous visibilisons dans les slides suivantes ce qui arrive **pendant le processus** de lecture, et d'écriture.
- Le produit final que peut fournir un LLM, comme le résumé d'un long texte (simuler la lecture) ou sa production (simuler l'écriture), n'est souvent que très secondaire dans les moments d'apprentissage.



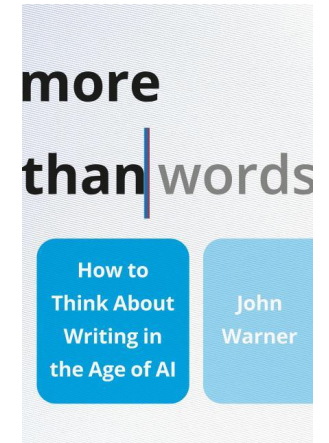
Importance de notre processus d'écrire pour réfléchir

“ J'ai souvent la sensation que je ne réfléchis pas assez profondément à un sujet avant d'avoir à écrire dessus. Souvent ma compréhension évolue donc pendant l'écriture. Et pour moi, c'est enthousiasmant et intellectuellement épanouissant. Les choses se mettent en place et je les maîtrise mieux, je trouve ma voix, mes opinions, je me construis mes visions. Et c'est pour ça qu'écrire demande des efforts. L'écriture n'est pas la seule façon de réfléchir, mais c'est une excellente façon de réfléchir.

- Pensez à quand vous rédigez des compte-rendus de TP, quand vous devez expliquer des phénomènes, des procédures, justifier de vos choix.
- Ou quand vous lisez ce qu'on écrit d'autres.

→ Quand on lit et qu'on écrit, on pense, on examine ce qu'on pense, on identifie des choses qu'on voudrait faire on dont on ressent le besoin, on se fait un avis, on découvre et on forge nos goûts.

→ Tout cela ne doit pas être mis en péril par la pression d'utiliser des LLM, dans les études notamment.



John Warner, Hachette Eds., 2025

<https://www.hachettebookgroup.com/titles/john-warner/more-than-words/9781541605510>

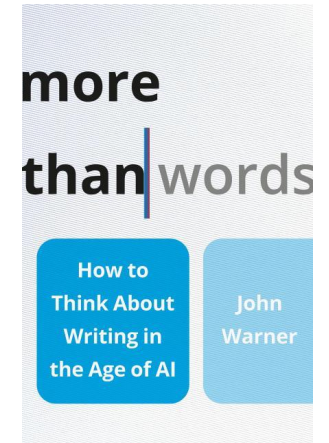
Écrire permet de comprendre ses pensées et d'affiner sa propre voix.



Importance de notre processus d'écrire pour ressentir.

- Après une fusillade de masse en février 2023, l'Université Vanderbilt a utilisé ChatGPT pour rédiger sa déclaration de condoléances. Elle l'a mentionnée au bas de la déclaration.
 - Un étudiant de Vanderbilt dont la sœur était touchée a critiqué le choix d'externaliser la déclaration à ChatGPT : “Il y a une ironie malsaine et tordue à demander à un ordinateur d'écrire votre message sur la communauté et la solidarité parce que vous avez la flemme d'y réfléchir vous-même.”
 - Même si, ou peut-être surtout si, aucune réponse concrète n'est apportée à ce type de tragédie, le choix de l'inaction **devrait avoir au moins un léger coût psychologique.**
- Il y a des années, pour éviter l'ennui des lettres de recommandation cliché, je me suis donné pour objectif de raconter une histoire spécifique sur chaque étudiant·e. Je ne me rappellerais pas de ces personnes si je sous-traitais cette rédaction à un LLM.

→ Si nous voulons résister à cet engourdissement, nous devons être conscients qu'écrire c'est ressentir, et si rien n'est ressenti lorsque nous écrivons, nous manquons une occasion de nous connecter aux autres, et à notre propre humanité.



John Warner, Hachette Eds., 2025

<https://www.hachettebookgroup.com/titles/john-warner/more-than-words/9781541605510>



« Utiliser ChatGPT pour effectuer des tâches, c'est comme amener un chariot élévateur dans la salle de musculation ; vous n'améliorerez jamais votre forme cognitive de cette façon. »

Ted Chiang

*“Once men turned their thinking over to machines in the hope that this would set them free.
But that only permitted other men with machines to enslave them.”*

Frank Herbert, Dune, 1965



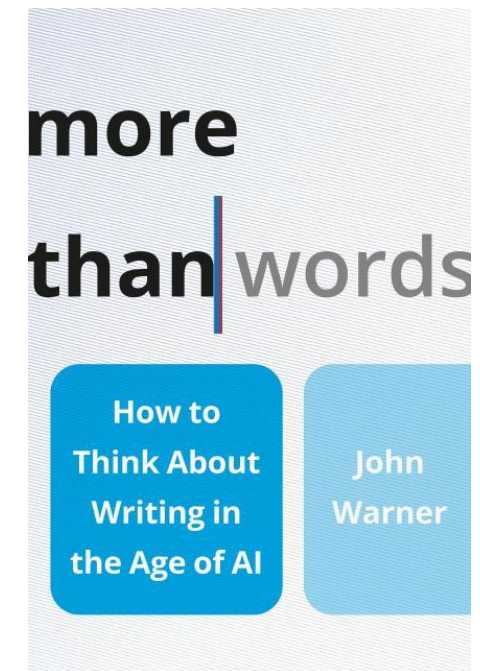
Plan

1. Exploitation humaine
2. Exploitation environnementale
3. Votre apprentissage à l'université
 1. Inquiétudes que l'apprentissage ne se fasse plu
 2. Importance du processus d'écrire pour nous
 3. Autorisez-vous à ressentir de la frustration avec un LLM et ayez de l'estime pour votre pensée
4. Récap'



LLM pour « se donner l'amorce » ?

- Certains verraient le résultat de GPT comme un bon point de départ, un texte qui peut être façonné par la révision et l'édition. Le défi le plus courant en écriture est de loin de se lancer, alors pourquoi ne pas laisser une machine nous aider ?
- Le fait que ChatGPT produise des écrits grammaticalement et syntaxiquement corrects le place au-dessus de ce que la plupart des étudiant·es peuvent produire.
- Mais c'est se méprendre sur la nature de l'écriture : **le premier jet est la partie la plus importante en termes d'implication humaine car il établit l'intention derrière l'expression.**
 - Si un·e étudiant·e vient me voir avec un texte généré par un LLM, nous ne pouvons pas discuter de ce qu'il veut dire, puisqu'il n'a encore rien dit.
- La production de texte synthétique est un simulacre d'écriture, et non l'écriture elle-même.



LLM pour « se donner l'amorce » ?

- « Nombre de nos collègues pensent que les LLM sont un excellent moyen de générer de nouvelles idées. Nous sommes sceptiques. Une littérature abondante en psychologie met en évidence les problèmes de fixation à la conception et de biais d'ancrage.
- Les personnes à qui l'on propose une solution possible à un problème ont tendance à être moins créatives que celles qui doivent partir de zéro.»
- « Un·e collègue l'utilisant pour générer un plan de cours doit être capable d'identifier les biais et omissions discriminatoires (par exemple, ignorer des textes clés écrits par des auteurs de milieux sous-représentés), sinon l'utilisation de GenAI risque de les perpétuer et de les exacerber. »



Responsible Adoption of Generative AI in Higher Education: Developing a “Points to Consider” Approach Based on Faculty Perspectives

Ravit Dotan
TechBetter
ravit@techbetter.ai

Lisa S. Parker
University of Pittsburgh
lisap@pitt.edu

John G. Radzilowicz
University of Pittsburgh
jgradz@pitt.edu

Carl T. Bergstrom and Jevin D. West, “[Modern-Day Oracles or Bullshit Machines? – Lesson 11](#),” lecture UW, 2025.

R. Dotan, L. S. Parker, and J. Radzilowicz, “[Responsible Adoption of Generative AI in Higher Education: Developing a ‘Points to Consider’ Approach Based on Faculty Perspectives](#),” in *ACM Conference on Fairness, Accountability, and Transparency*, 2024.



LLM pour organiser mes idées ?

Ayez de l'estime pour votre pensée.

- Même si le LLM était idéal à mon sens (?), mon système 3, celui qui contrôle [rappelez-vous !] me crie de ne pas abdiquer cette prise dont j'ai besoin sur ma vie, pour laquelle mes pensées sont centrales.
 - Et ça vaut autant pour le domaine personnel que professionnel... Comme faire ce cours.
- Mais en plus on sait à présent que :
 - Les LLM sont très peu fiables avec toute leur production conduite par les mots vus ensemble dans les données d'entraînement
 - Quand je vois les mauvaises performances de ces systèmes sur les tâches sortant du commun des données d'entraînement, je ne souhaite pas que ça manipule quoique ce soit de mes pensées. Nos pensées sont toujours complexes et particulières.
 - Entrer des prompts personnels implique de forts risques pour votre vie privée (collecte de données pour modèles attaquables et profiling)



Plan

1. Exploitation humaine
2. Exploitation environnementale
3. Votre apprentissage à l'université
 1. Inquiétudes que l'apprentissage ne se fasse plu
 2. Importance du processus d'écrire pour nous
 3. Autorisez-vous à ressentir de la frustration avec un LLM et ayez de l'estime pour votre pensée
- ➡ 4. Récap'



Alors ?

- Qu'avez-vous retenu ?



Donc...

- Demandons-nous dans chaque situation où nous pensons faire appel à un LLM : quel est votre intérêt à court terme, quel est votre intérêt à long terme (et sur quelles bases – travaux scientifiques, l'estimez-vous) ?
 - Intérêt dans le sens de ce qui va nous permettre d'être enthousiaste chaque jour pour ce qu'on va faire.
- Posez-vous des questions, doutez, écoutez votre sens de frustration, doute ou peur de ne pas acquérir les connaissances que vous pourriez et que ça vous desserve à terme : écoutez votre système 3 ;)



Et vous ?

- Comment est-ce que ce module fait évoluer votre calcul du rapport coût/bénéfice quand vous décidez d'utiliser ou non un LLM, sur les critères :
 - pour votre acquisition des savoirs (pas pour vos notes) ?
 - pour votre bien-être intellectuel ? Pour arriver où vous souhaitez, être satisfait·e et ne pas ressentir de frustration de comment vos études vous profitent
 - pour l'environnement
- Vous avez le plus souvent la possibilité de choisir vos usages.
- Quels pourront être vos calculs de rapport bénéfice/coût ?
 - Quels critères seront importants pour vous ?
 - Comment varieront-ils en fonction de la situation ?

