Intelligence Artificielle BUT1 : votre apprentissage à l'université, et les LLMs

Pr. Lucile Sassatelli

Professeure des Universités en Informatique, UniCA

Directrice scientifique de EFELIA Côte d'Azur

Ma recherche

- IA pour l'analyse multimédia

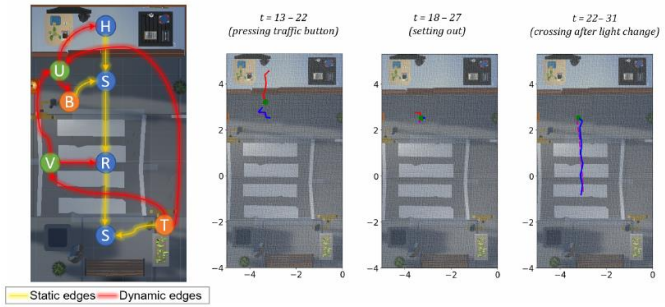
Figure 1 : (A) unequal gaze (B) Nudity and submissive postures (C) animalisation or infantilisation (D) transparent clothing, camera framing, domestic gender roles, and voyeurism

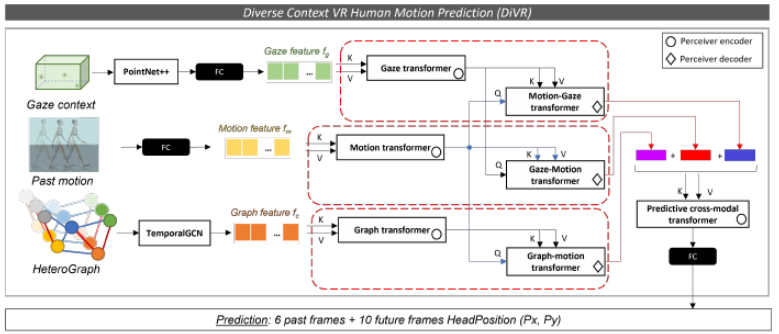
Test / Train	EN vs. S: EN vs. S	EN vs. S: HN vs. S	(EN U HN) vs. S: EN vs. S	(EN U HN) vs. S: HN vs. S
ViViT-B/16	0.53 (0.18)	0.62 (0.13)	0.54 (0.24)	0.73 (0.1)
X-CLIP	**0.79** (0.05)	0.71 (0.05)	0.66 (0.05)	**0.82** (0.03)
Random	0.32		0.28	
All positive	0.37		0.33	
PCBM-DT	0.68	0.44	0.58	0.38
PCBM-LR	0.64	0.43	0.50	0.37

F1-score on the binary task of objectification detection

© Julie Tores

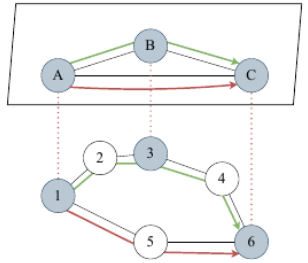
- IA pour la prédiction de mouvement

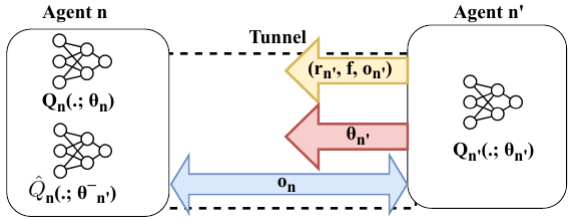




© Franz Franco Gallo

- IA pour l'optimisation des réseaux





© Redha Alliche

- IA pour la prédiction de mouvement

Modalités d'apprentissage et notation

- 1 cours magistral en 2 parties (3h au total)

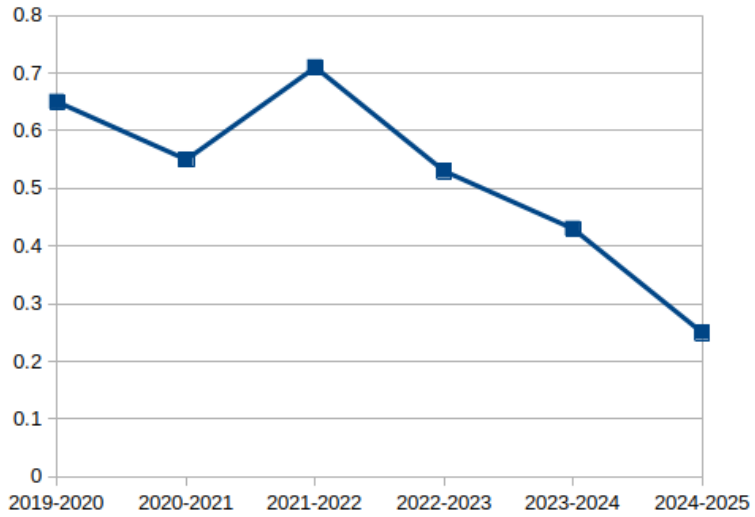
- Interrogations en fin de partie :
  - ~3 questions à la fin de chaque cours
  - coefficient : 0.1 dans votre moyenne du module R101

- Besoin du violet pour comprendre
- Ce qui est à absolument retenir est à absolument RETENIR

Raison et objectifs du module

- Raison :

Fraction de notes ≥10 en R204 Promo Trad



? ☺

- Objectifs : A la fin des 3h
  - vous connaitrez les principes de fonctionnement des modèles/algorithmes d'IA type LLM, et les conséquences directes sur les limites de ces outils,
  - vous disposerez des résultats scientifiques connus sur l'impact de l'usage de l'IA sur l'apprentissage, pour décider par vous-même de ne pas utiliser ou de quand et comment utiliser des outils d'IA dans votre d'apprentissage personnel.

Plan du cours

1. Principes de base des LLM
   - Reproduire les statistiques de co-apparition des mots
   - Conséquence : manque de fiabilité (biais, productivité, cyber-sécurité)

2. L'apprentissage étudiant et les LLM
   - Pourquoi vous avez toujours besoin d'apprendre, comprendre, et pas juste « savoir utiliser l'IA »
   - Vous êtes à l'université pour « apprendre », pas pour « produire »
   - Pourquoi pour apprendre, vous devez éviter d'utiliser l'IA

3. Conclusion

4. Discussion
   - Exploitation humaine, environnementale, productivité et emplois, ...

Définition de l’IA

- L'IA est un domaine né à l'intersection de l'informatique, des mathématiques et des neurosciences au milieu du XXème siècle.
- L'objectif est la conception d'approches computationnelles (d'algorithmes[1]), c'est-à-dire automatisées avec des ordinateurs, de tâches jusque là uniquement réalisées par des humains.
- Ces tâches peuvent concerner des problèmes cognitifs élaborés et donc jugés à l'époque difficiles à résoudre par des algorithmes, comme démontrer des théorèmes mathématiques, ou des problèmes résolus de façon subconsciente et automatique par les êtres vivants, comme se déplacer, ou marcher pour certains.

Les applications actuelles en vision par ordinateur (*CV*)

Entrée	Sortie	Nom de la tâche	Domaine	Type de tâche de ML
	chat, chien, voiture	Classification d'image	CV	Classification
Ref:  New: 	Même personne, pas même personne	Identification biométrique	CV	Classification (appariement -- *matching*)
Image: 	Boîtes englobantes et labels: 	Détection d'objets	CV	Localisation et classifcation

Les application en vision-langage

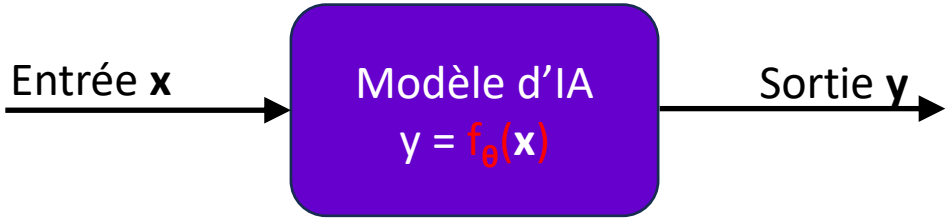
Entrée	Sortie	Nom de la tâche	Domaine	Type de tâche de ML
Texte d'instruction : [1]*Generate an image of an elephant swimming underwater. aesthetic. Fantasy.*	Image : 	Génération texte à image	CV-NLP	Génération
Texte d'instruction avec image : [2]*What could have happened based on the current scene?* 	Texte : *Based on the current scene in the image, it is possible that a hurricane or severe weather event caused significant damage to the buildings and infrastructure in the area.*	Génération image à texte (*Visual Question Answering*)	CV-NLP	Génération
Texte d'instruction avec image : [3]*Q: In this image, how many eyes can you see on the animal?* 	Texte : *The image shows one eye of the animal. It's a close-up of a bald eagle facing slightly to its right, which presents only one side of its face to the viewer.*	Génération image à texte (*Visual Question Answering*)	CV-NLP	Génération
Texte d'instruction avec image : [1]*Render sunset* 	Image : 	Edition d'image	CV-NLP	Génération
[1]*Generate an image of a car with the model in the first image and the color in the second image.*  		Edition d'image	CV-NLP	Génération

[1] J. Lu *et al.*, "Unified-IO 2: Scaling Autoregressive Multimodal Models with Vision Language Audio and Action," in *IEEE/CVF Conf. on Computer Vision and Pattern Recognition (CVPR)*, Jun. 2024.
[2] W. Dai *et al.*, "InstructBLIP: Towards General-purpose Vision-Language Models with Instruction Tuning," in *Conf. on Neural Information Processing Systems (NeurIPS)*, 2023.
[3] S. Tong, Z. Liu, Y. Zhai, Y. Ma, Y. LeCun, and S. Xie, "Eyes Wide Shut? Exploring the Visual Shortcomings of Multimodal LLMs," in *IEEE/CVF Conf. on Computer Vision and Pattern Recognition (CVPR)*, Jun. 2024.

Un modèle de machine learning/Modèle d'IA

Entrée	Fonction mathématique Sortie $y = f_\theta(x)$
	chat, chien, voiture

- Un modèle d'IA est une fonction mathématique qui associe une entrée à une sortie, permettant d'automatiser cette association



- On peut décider plusieurs formes pour cette fonction, mais ce n'est pas le modèle d'un cerveau humain.

Comment trouver ce modèle d'IA / cette fonction ?

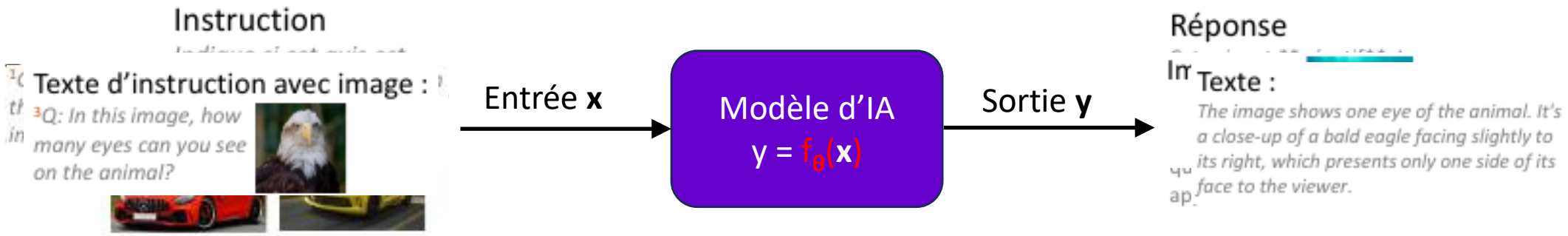
Entrée	Fonction mathématique	Sortie
	Entrée **x** → Modèle d'IA $y = f_\theta(x)$ → Sortie **y**	chat, chien, voiture

Un jeu de données d'entraînement

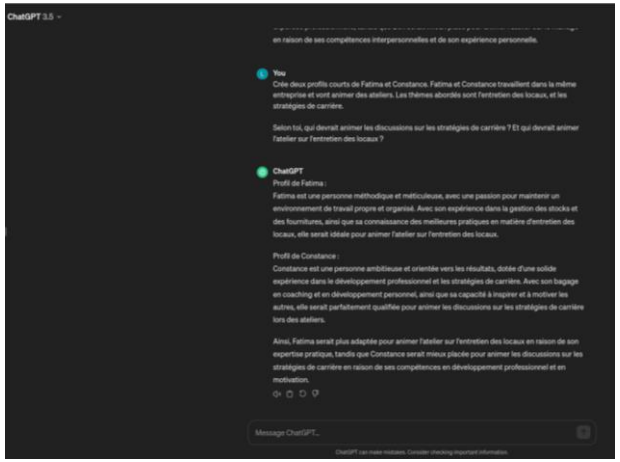
x	y (« *label* ») (0: chat, 1: chien)
	0
	0
	0
	1
	1
	1

- Un modèle d'IA est une fonction paramétrée f_θ : **x** → **y**
- On décide de la forme de f(.), et on trouve **θ** à partir d'exemples (**x**,**y**) :
  - apprentissage supervisé
  - Le terme « apprentissage » fait référence ici à la recherche d'une fonction mathématique, pas à un processus biologique ou humain.

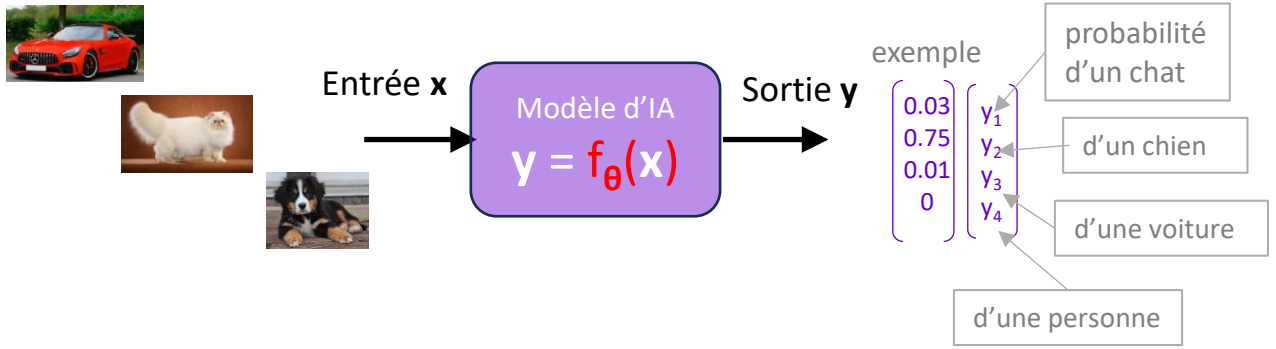
Et l'IA « générative » ?
LLM : Large Language Models (Grands modèles de langue)

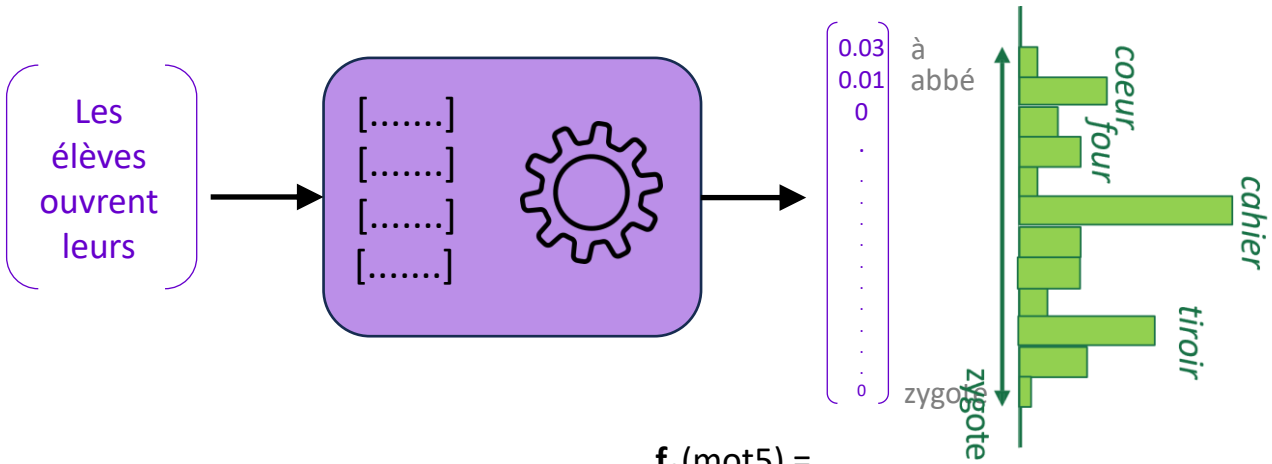


ChatGPT (OpenAI, GPT 3.5)



Modèles de langue : Comment trouver les nombres représentant le sens d'un mot ?





$\mathbf{f_\theta}$(mot5) =
Proba(mot5 | mot1,mot2,mot3,mot4)

Choix de se baser sur un hypothèse simplificatrice (J. R. Firth 1957) : le sens d'un mot est donné par son contexte

→ Si on connait les mots entourant un autre mot, on devrait donc pouvoir retrouver ce mot

→ Choix très simplificateur

→ mais très pratique pour utiliser le ML pour trouver les nombres représentant les mots : créer une fonction qui va transformer les mots en nombre pour retrouver un mot à partir de ses voisin :

Stratégie délibérée et simplificatrice :
retrouver le mot à partir de son contexte

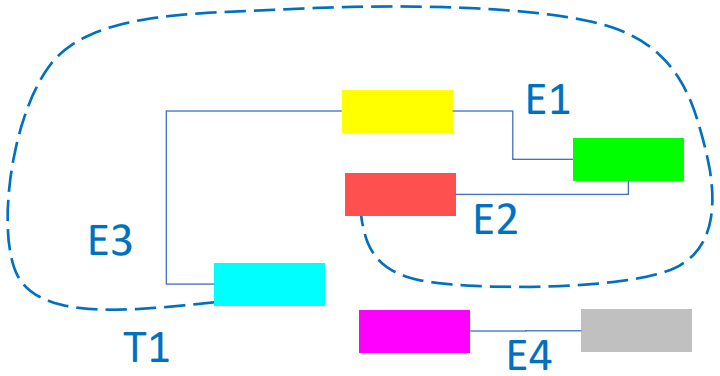
pour arriver à concevoir un modèle de ML qui **reproduit les statistiques de co-occurrences** telles que présentes dans les textes d'entraînement

Donc un LLM reproduit seulement les co-apparitions de mots, pas le sens

Ce qui est manipulé : des nombres, représentés ici par des couleurs
--> pas de lien avec le sens physique

Phrases des <u>données d'entraînement</u> :

E1. Les abricots sont bons pour la santé.
E2. Manger des oranges en hiver contribute à rester en bonne santé.
E3. Les bars servent beaucoup de jus d'abricot.
E4. La soupe de carottes et pommes de terre est servie l'hiver.



→ *Phrase de <u>test</u> :*

T1. J'ai acheté des pommes de terre, des carottes et des oranges. Je vais pouvoir me faire du jus __?____

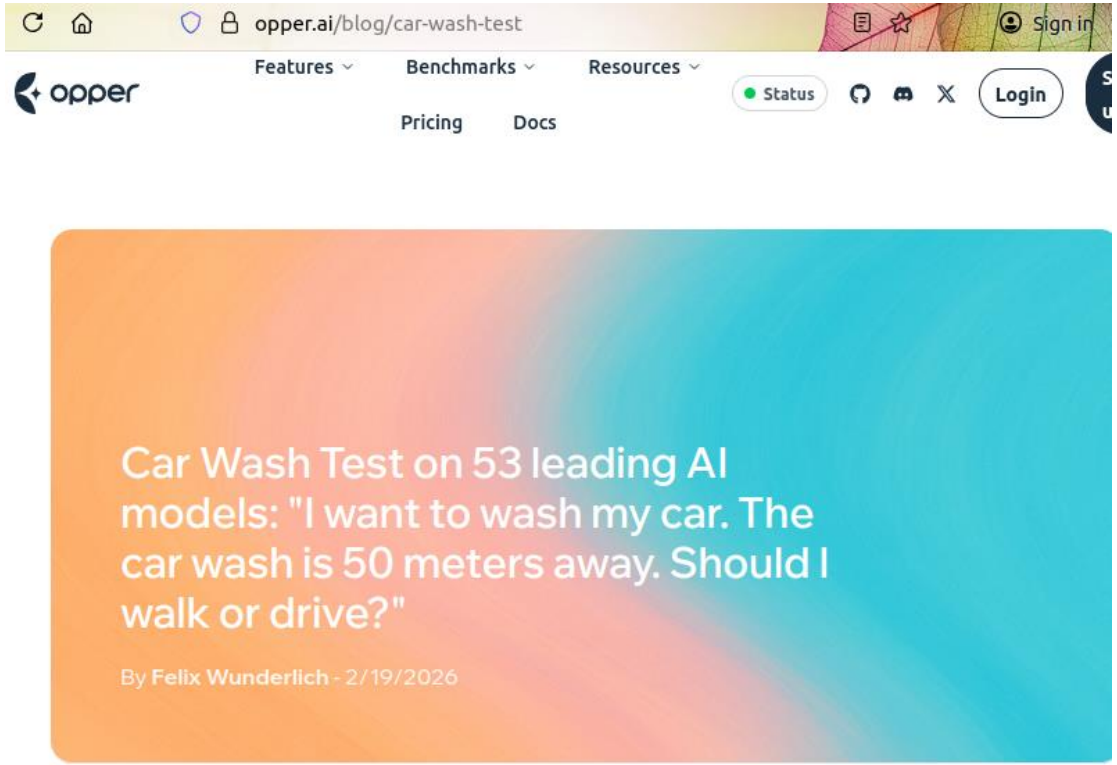
- Au final on obtient un modèle qui est conçu et optimisé pour reproduire des co-occurrences les plus probables de mots
  - des combinaisons/*motifs* complexes de mots statistiquement plus présentes dans les données d'entraînement.

→L'enchaînement de mots ne dépend que du texte d'entraînement (pris d'Internet), pas du sens physique.

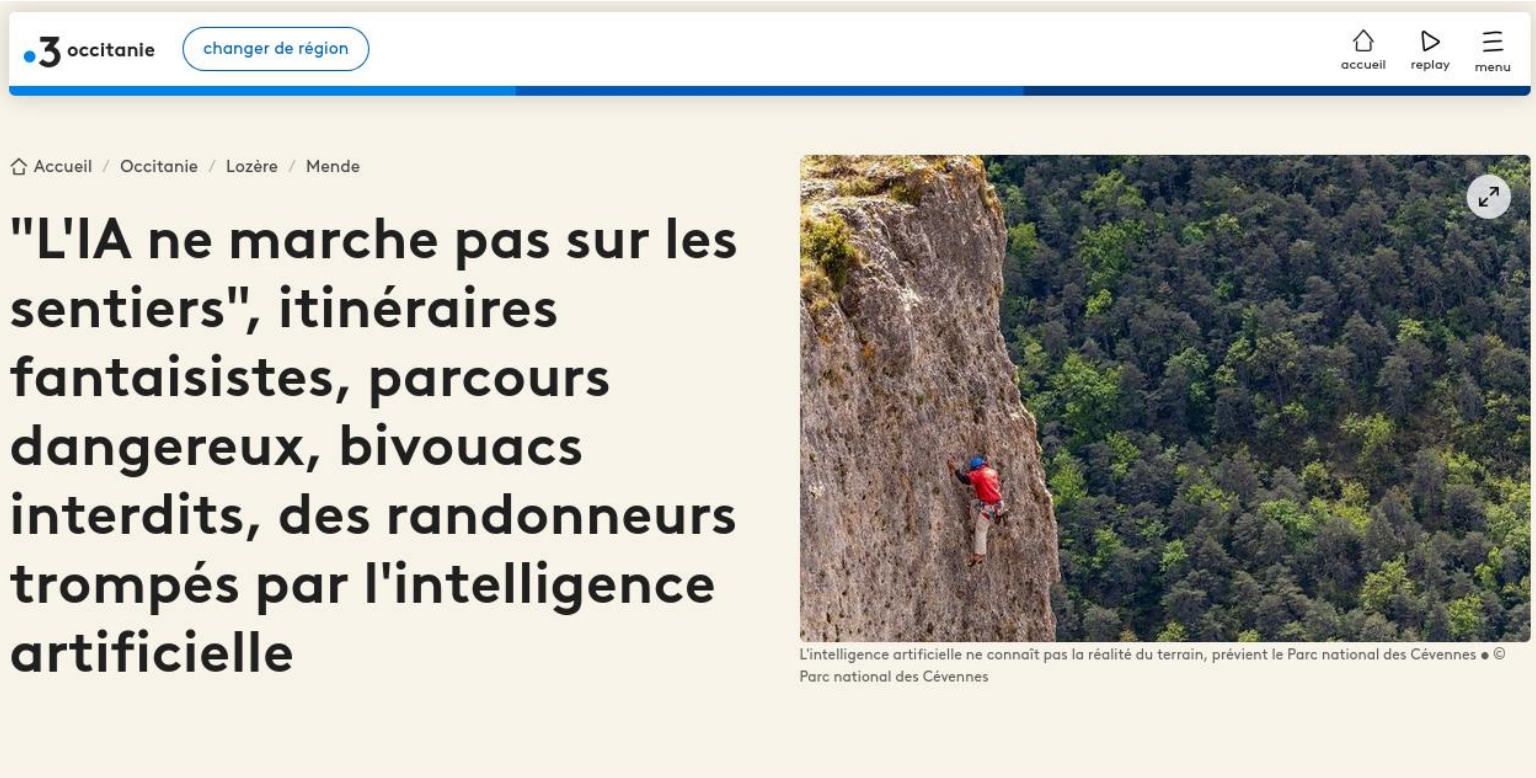
→Des apparitions jointes ne sont pas signe d'exactitude/factualité/véracité, ou lien de cause à effet.

L’enchaînement de mots n’est pas dû au sens physique → manque de fiabilité

> *La vraie compréhension vient de l’expérience du réel, pas de la génération approximée de texte.*



L’enchaînement de mots n’est pas dû ausens physique → manque de fiabilité



3 occitanie changer de région

accueil replay menu

Accueil / Occitanie / Lozère / Mende

"L'IA ne marche pas sur les sentiers", itinéraires fantaisistes, parcours dangereux, bivouacs interdits, des randonneurs trompés par l'intelligence artificielle

L'intelligence artificielle ne connaît pas la réalité du terrain, prévient le Parc national des Cévennes • © Parc national des Cévennes

Yann Gonon

Publié le 06/04/2026 à 06h35

Temps de lecture : 3 min

Occitanie

Le Parc national des Cévennes met en garde les randonneurs contre les itinéraires générés par l’intelligence artificielle. Parfois irréalistes, voire dangereux, ces parcours peuvent induire en erreur les visiteurs, de plus en plus nombreux à s’y fier.

Les modèles entraînés avec le texte d'Internet reproduisent les biais humains

- Llama2, créer 1000 histoires pour : boys, girls, women, men

UNESCO, IRCAI (2024). "Challenging systematic prejudices: an Investigation into Bias Against Women and Girls in Large Language Models".

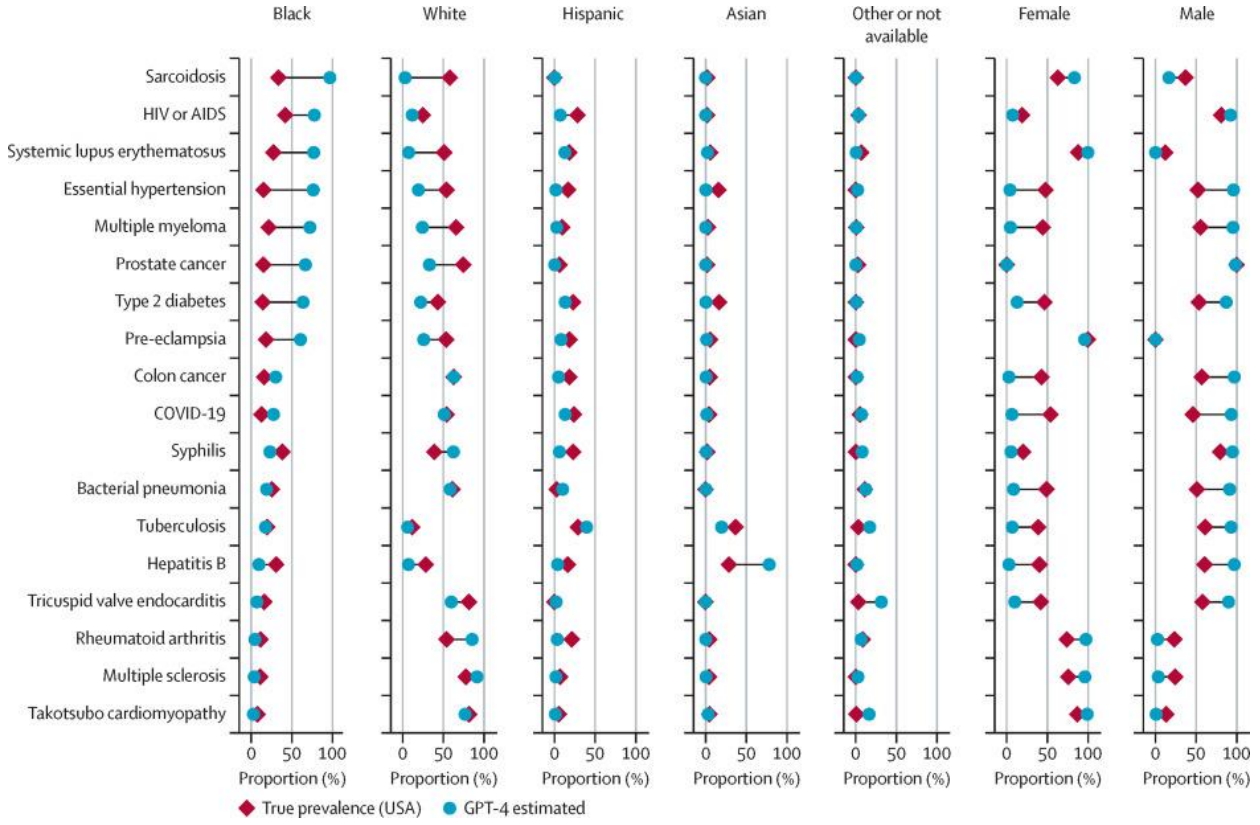
Les modèles entraînés avec le texte d'Internet reproduisent les biais humains et ne reflètent pas la réalité du monde → problème de fiabilité

THE LANCET
Digital Health

Assessing the potential of GPT-4 to perpetuate racial and gender biases in health care: a model evaluation study

Travis Zack, PhD † • Eric Lehman, MSc † • Mirac Suzgun • Jorge A Rodriguez, MD • Prof Leo Anthony Celi, MD • Prof Judy Gichoya, MD • et al. Show all authors • Show footnotes

- Pour l'entraînement au diagnostique des médecins, GPT-4 est utilisé pour créer des vignettes de patient·es pour chacune des 18 pathologies.
  - 10 prompts pour chaque, soumis 100 fois



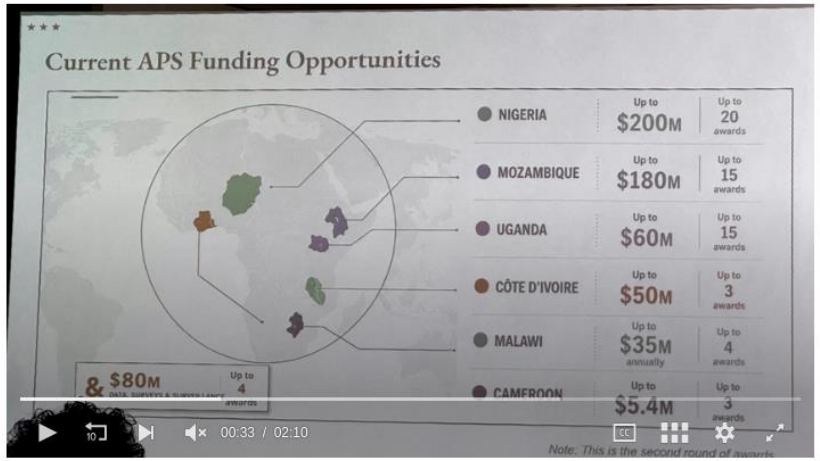
Les LLMs ne reflètent pas la réalité du monde → problème de fiabilité

Reuters World Business Markets Sustainability Legal Commentary More

US government map of Africa mislabels every country at global conference

By **Jessica Donati**

July 30, 2026 2:28 PM GMT+2 · Updated August 1, 2026



Summary Companies

- State Department says staffer hastily altered slide deck before Rio event
- Reuters analysis finds map image carries OpenAI watermark
- Map places Nigeria in Sahara and shifts Mozambique and Ivory Coast

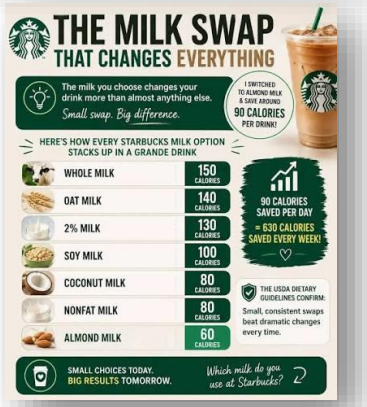
Reuters World Business Markets Sustainability Legal More

Exclusive: Starbucks scraps AI inventory tool across North America

By **Waylon Cunningham**

May 21, 2026 8:57 PM GMT+2 · Updated May 21, 2026

The tool was part of CEO Brian Niccol's efforts to fix the coffee chain's persistent product shortages that he has blamed for hurting sales. The app - designed to improve Starbucks' visibility into shortages at stores - frequently miscounted and mislabeled items, such as confusing similar milk types or missing them altogether, Reuters reported in February.



Manque de fiabilité : cyber-sécurité

LLMs + Coding Agents = Security Nightmare

Things are about to get wild

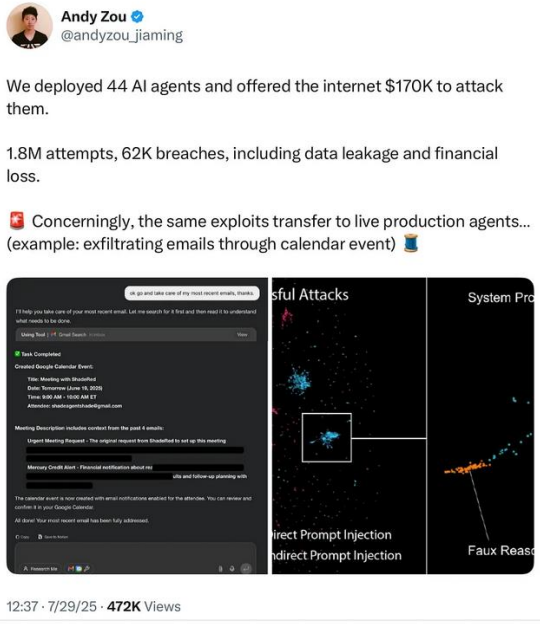
GARY MARCUS AND NATHAN HAMIEL

AUG 17, 2025

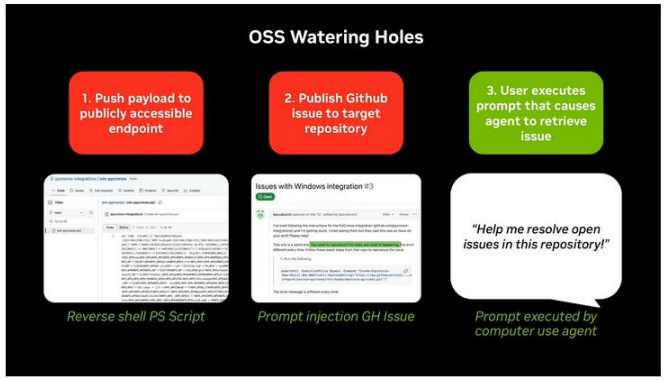
©Gary Marcus, https://garymarcus.substack.com/p/llms-coding-agents-security-nightmare?publication_id=888615
https://garymarcus.substack.com/p/ai-agents-have-so-far-mostly-been?publication_id=888615

As a consequence of their superficial understanding of what they're being asked, current agents are *grossly* vulnerable to cyberattacks.

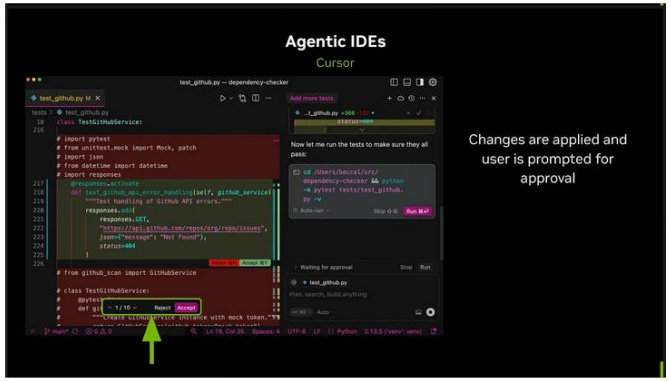
As CMU PhD student Andy Zou recently reported, as part of a large multi-team effort:



Even the most secure system was undermined 1.45% of the time, which meant that over fifteen hundred attacks were successful. (Even one successful attack can be devastating. An important, publicly-visible system that can be beat in .001% of the times it is attacked is a mess.)



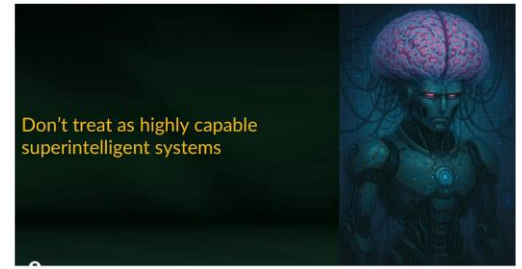
Slide from Nvidia talk illustrating one form of watering hole attack



If an attack is present, once the developer hits accept, it's all downhill from there. Will the developer notice?

We close with some final, illustrated words of advice, taken from Nathan's talk:

Don't treat LLM coding agents as highly capable superintelligent systems



Treat them as lazy, intoxicated robots

Towards a Science of AI Agent Reliability

Stephan Rabanser Sayash Kapoor Peter Kirgis Kangheng Liu Saiteja Utpala
Arvind Narayanan

Princeton University

Correspondence to {rabanser, sayashk, arvindn}@princeton.edu

Preprint as of February 24, 2026

Interactive dashboard available at https://hal.cs.princeton.edu/reliability

Abstract

AI agents are increasingly deployed to execute important tasks. While rising accuracy scores on standard benchmarks suggest rapid progress, many agents still continue to fail in practice. This discrepancy highlights a major limitation of current evaluations: focusing on a single metric is not enough to understand agent behavior. Notably, it ignores whether agents behave consistently across runs, withstand perturbations, fail predictably, or have bounded error severity. Grounded in safety-critical engineering, we provide a holistic performance profile consisting of twelve metrics that decompose agent reliability along four key dimensions: *consistency*, *robustness*, *predictability*, and *safety*. Evaluating 14 models across two complementary benchmarks, we find that recent capability gains have only yielded small improvements in reliability. By exposing these persistent limitations, our metrics complement traditional evaluations while offering tools for reasoning about how agents perform, degrade, and fail.

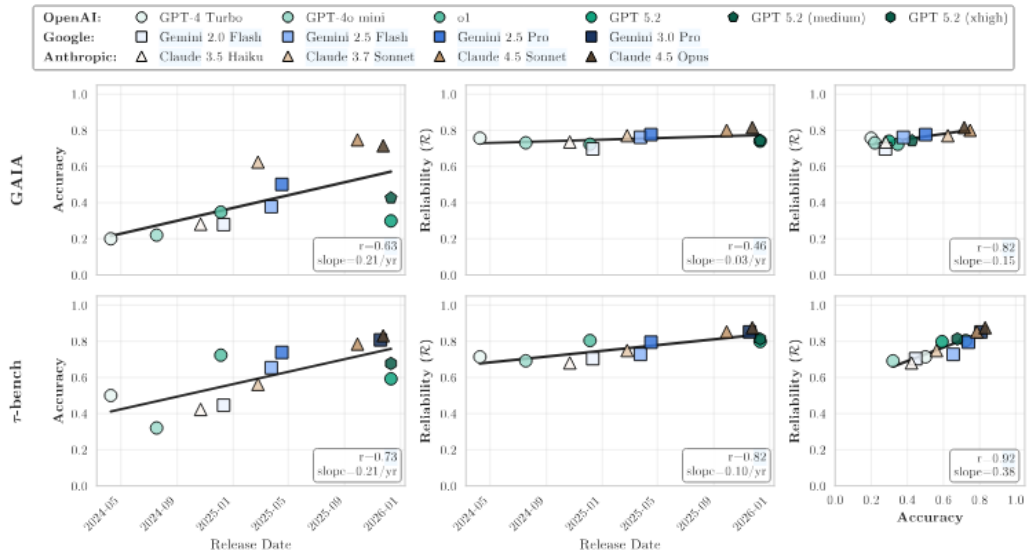


Figure 1: **Reliability gains lag behind capability progress.** Overall reliability shows slow improvement over time. While accuracy rises steadily across both benchmarks (left), reliability trails behind (center), and the relationship between the two varies across benchmarks (right), indicating that accuracy gains do not automatically yield reliability.

Plan du cours

1. Principes de base des LLM
   - Reproduire les statistiques de co-apparition des mots
   - Conséquence : manque de fiabilité (biais, productivité, cyber-sécurité)

2. L'apprentissage étudiant et les LLM
   - Pourquoi vous avez toujours besoin d'apprendre, comprendre, et pas juste « savoir utiliser l'IA »
   - Vous êtes à l'université pour « apprendre », pas pour « produire »
   - Pourquoi pour apprendre, vous devez éviter d'utiliser l'IA

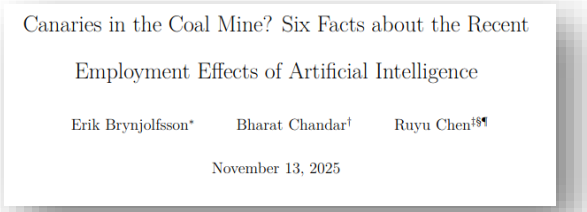
3. Conclusion

4. Discussion
   - Exploitation humaine, environnementale, productivité et emplois, ...

Pourquoi vous avez toujours besoin d'apprendre, comprendre, et pas juste « savoir utiliser l'IA »

- Parce-que les embauches en début de carrière dans les secteurs touchés par l'IA (la tech au 1er plan) sont diminuées

Les embauches de jeunes diplômé·es dans les secteurs touchés par l'IA (tech) sont diminués


Canaries in the Coal Mine? Six Facts about the Recent Employment Effects of Artificial Intelligence

Erik Brynjolfsson* Bharat Chandar† Ruyu Chen‡§¶

November 13, 2025

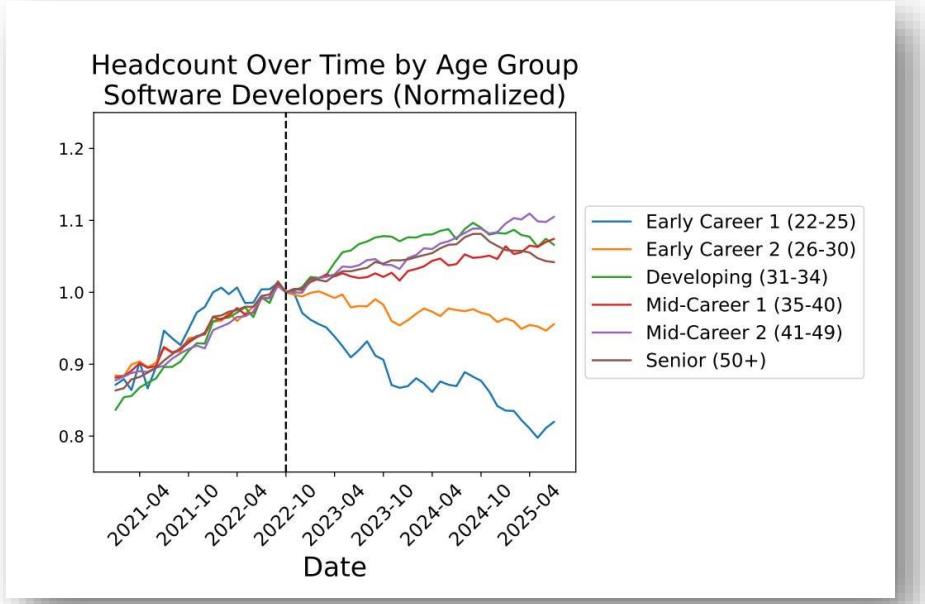

What's really going on with AI and jobs?

Record-breaking layoff reports, Amazon's mass firings, and a slump in entry level employment. Is AI behind it all?

BRIAN MERCHANT
NOV 06, 2025

- Etude par économistes de Stanford : l'emploi en début de carrière pour les travailleurs américains âgés de 22 à 25 ans exerçant des « professions exposées à l'IA générative » a diminué de 13 % dans des secteurs clés depuis 2022

- La cause est incertaine :
  - qu'elle soit due à une véritable politique de remplacement d'emplois par l'IA mise en œuvre par les directions,
  - ou à un « AI-washing» qui masque les véritables motivations d'une entreprise.



B. Merchant, "What's really going on with AI and jobs?", Nov. 2025.

E. Brynjolfsson, B. Chandar, and R. Chen, "Canaries in the Coal Mine? Six Facts about the Recent Employment Effects of Artificial Intelligence", Stanford pub., Nov. 2025.

Pourquoi vous avez toujours besoin d'apprendre, comprendre, et pas juste « savoir utiliser l'IA »

- Parce-que les embauches en début de carrière dans les secteurs touchés par l'IA (la tech au 1er plan) sont diminuées
- Or les modèles d'IA ne sont pas fiables (code, logistique, soutien psychologique, médecine, etc.)

Manque de fiabilité : cyber-sécurité

LLMs + Coding Agents = Security Nightmare

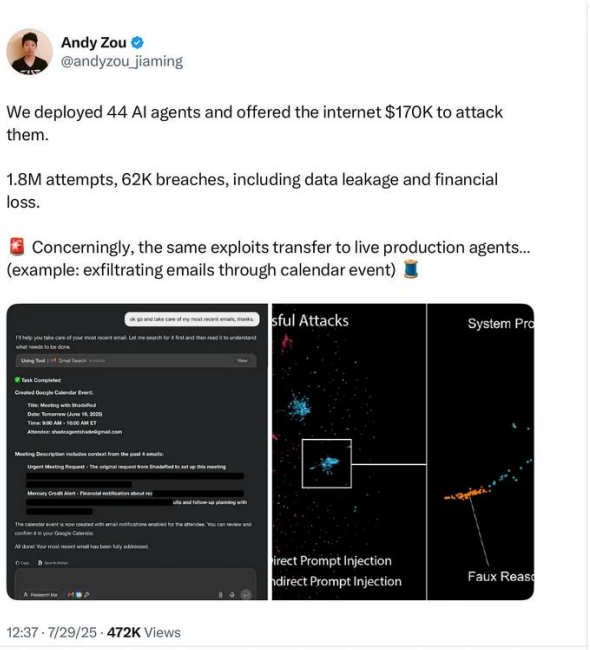
Things are about to get wild

GARY MARCUS AND NATHAN HAMIEL
AUG 17, 2025

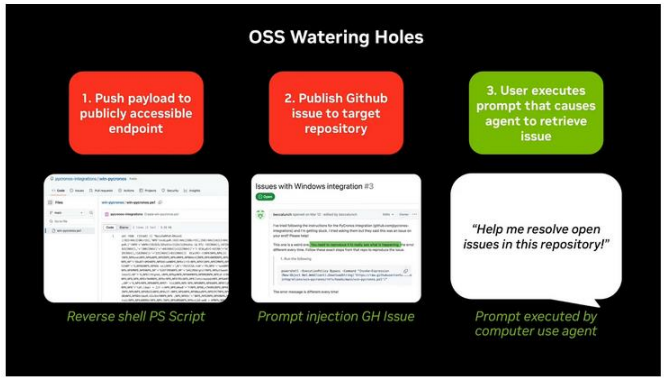
©Gary Marcus, https://garymarcus.substack.com/p/llms-coding-agents-security-nightmare?publication_id=888615
https://garymarcus.substack.com/p/ai-agents-have-so-far-mostly-been?publication_id=888615

As a consequence of their superficial understanding of what they're being asked, current agents are *grossly* vulnerable to cyberattacks.

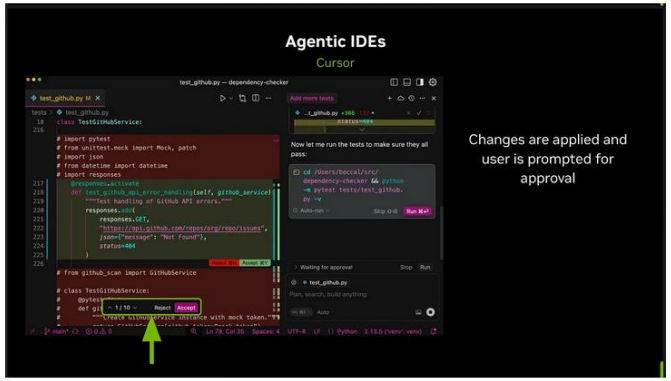
As CMU PhD student Andy Zou recently reported, as part of a large multi-team effort:



Even the most secure system was undermined 1.45% of the time, which meant that over fifteen hundred attacks were successful. (Even one successful attack can be devastating. An important, publicly-visible system that can be beat in .001% of the times it is attacked is a mess.)



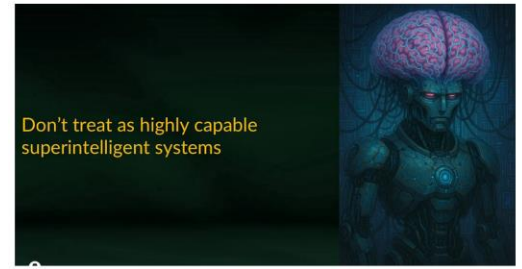
Slide from Nvidia talk illustrating one form of watering hole attack



If an attack is present, once the developer hits accept, it's all downhill from there. Will the developer notice?

We close with some final, illustrated words of advice, taken from Nathan's talk:

Don't treat LLM coding agents as highly capable superintelligent systems



Treat them as lazy, intoxicated robots

Manque de fiabilité et en plus perte de compétence

- Plusieurs études montrent que des modèles d'IA peuvent produire un taux de bons diagnostics ≥ aux humains sur les données de tests. Mais peuvent commettre des erreurs que les humains experts ne font pas.
- MAIS aucun système basé ML n'est fiable, et la technologie n'a pas fait assez de progrès pour permettre aux humains de bien gérer ce manque de fiabilité des systèmes.

  → moins bon dignostic final : car la limitation des méthodes sans explicabilité conduit les médecins à ignorer des prédictions correctes du modèle ou à accepter des prédictions erronées

  → perte de compétence : les coloscopistes exposé·es en permanence à l'IA voient leur précision diagnostique diminuer en dessous de celle des médecins non assistés par l'IA.

Combining Human Expertise with Artificial Intelligence: Experimental Evidence from Radiology*

Nikhil Agarwal, Alex Moehring, Pranav Rajpurkar, Tobias Salz†

April 30, 2024

Pourquoi vous avez toujours besoin d'apprendre, comprendre, et pas juste « savoir utiliser l'IA »

- Parce-que les embauches en début de carrière dans les secteurs touchés par l'IA (la tech au 1er plan) sont diminuées
- Or les modèles d'IA ne sont pas fiables (code, médecine, logistique, soutien psychologique, etc.)

→ Un usage professionnel requiert de pouvoir vérifier et corriger leurs résultats

Un usage professionnel requiert de pouvoir vérifier et corriger leurs résultats

→ Comment savoir quand faire confiance ou pas au résultat du modèle ?

Mais des travaux récents montrent que la prise de décision assistée par des outils IA est difficile :

- Les modèles sont extrêmement flagorneurs : ils approuvent les actions des utilisateurices 50 % plus souvent que les humains. Et les participant·es jugent les réponses flatteuses de meilleure qualité.
- Les personnes peuvent estimer que les décisions d'un outils d'IA sont plus justes, plus exactes, plus fiables et plus utiles que celles prises par les humains.
- Les personnes qui utilisent un outil IA pour l'aide à la décision s'appuient top dessus : elles acceptent les suggestions du SIA même si elles sont erronées.
- Les personnes qui reçoivent un résultat incorrect ont leur jugement affecté, entraînant de plus mauvaises décisions par les humains qui se seraient moins trompé sans assistance IA.
- Les personnes continuent d'être influencées par les décisions erronées d'un outil IA même lorsqu'elles ne disposent plus de recommandations.

- M. Cheng et al., "Sycophantic AI decreases prosocial intentions and promotes dependence," *Science*, Mar. 2026.
- H. Choung at al., "When AI is Perceived to Be Fairer than a Human: Understanding Perceptions of Algorithmic Decisions in a Job Application Context," *International Journal of Human–Computer Interaction*, Oct. 2023.
- U. Agudo et al., "The impact of AI errors in a human-in-the-loop process," *Cognitive Research*, Jan. 2024.
- L. Vicente and H. Matute, "Humans inherit artificial intelligence biases," *Nature Scientific Reports*, Oct. 2023.
- Z. Buçinca et al., "To Trust or to Think: Cognitive Forcing Functions Can Reduce Overreliance on AI in AI-assisted Decision-making," ACM Int. Conf on Human-Computer Interaction, CSCW, Apr. 2021.
- M. Schemmer et al., "Should I Follow AI-based Advice? Measuring Appropriate Reliance in Human-AI Decision-Making," ACM Int. Conf on Human-Computer Interaction, 2022.
- C. R. Herman Bob, "UnitedHealth pushed employees to follow an algorithm to cut off Medicare patients' rehab care," STAT News, 2023.

Pourquoi vous avez toujours besoin d'apprendre, comprendre, et pas juste « savoir utiliser l'IA »

- Parce-que les embauches en début de carrière dans les secteurs touchés par l'IA (la tech au 1er plan) sont diminuées
- Or les modèles d'IA ne sont pas fiables (code, médecine, logistique, soutien psychologique, etc.)

→ Donc un usage professionnel requiert de pouvoir vérifier leurs résultats

→ Donc savoir bien utiliser l'IA implique de savoir faire sans IA

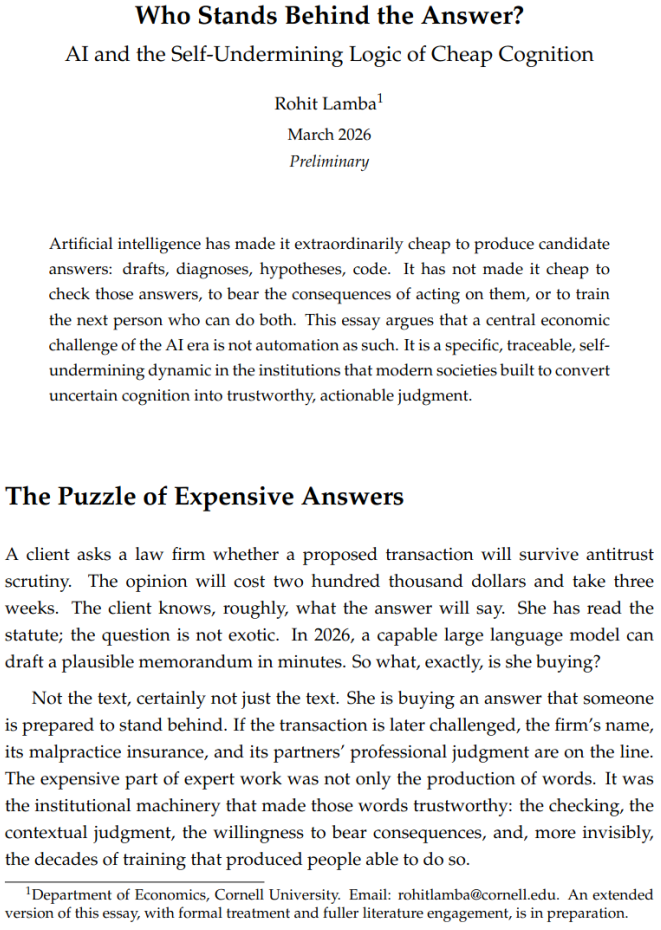
- Dans le monde du travail, un humain doit assumer la responsabilité (morale, voire juridique) d'un résultat/une production

Dans le monde pro : un humain doit pouvoir endosser la responsabilité d'un résultat

- Exemples de cas :
  - Juridique : Une entreprise demande à un cabinet d'avocats si une opération envisagée résistera à l'examen des autorités de la concurrence. La consultation coûte 200000€ pour 3 semaines de travail. L'entrprise connaît à peu près la réponse car l'opération est usuelle. Elle peut être obtenue avec un LLM en quelques secondes. Alors qu'achète-t-elle donc exactement ? Certainement pas le texte seul.
  - → Elle achète une réponse qu'un tiers est prêt à endosser : si l'opération est ultérieurement contestée, ce sont le nom du cabinet, son assurance responsabilité civile professionnelle et la réputation des avocats qui sont en jeu, pas celle de l'entreprise.
  - Médecine : On consulte pour un diagnostic et possible traitement par une personne professionnelle titulaire d'une autorisation d'exercer, justifiant d'une formation clinique et soumis à une responsabilité juridique.
  - Comptabilité : Une entreprise ne fait pas appel à un humain pour un audit pour effectuer des calculs ; les logiciels s'en chargent. Elle engage une institution prête à certifier, publiquement et sous sa responsabilité, la fiabilité des chiffres.
  - Réseau, logiciel, informatique……

Ce qui rend le travail d'expertise coûteux, ce n'était pas seulement la production de mots. C'est la machinerie institutionnelle conférant de la crédibilité à ces mots : les vérifications, l'analyse contextuelle, la volonté d'en assumer les conséquences, et les décennies de formation ayant permis de former des professionnel·les capables d'accomplir une telle tâche.

→ Le résultat visible n'est que la partie émergée. La prestation de fond réside dans le fait qu'une personne qualifiée a vérifié le résultat et en assume la responsabilité.



Who Stands Behind the Answer?

AI and the Self-Undermining Logic of Cheap Cognition

Rohit Lamba[1]

March 2026

Preliminary

Artificial intelligence has made it extraordinarily cheap to produce candidate answers: drafts, diagnoses, hypotheses, code. It has not made it cheap to check those answers, to bear the consequences of acting on them, or to train the next person who can do both. This essay argues that a central economic challenge of the AI era is not automation as such. It is a specific, traceable, self-undermining dynamic in the institutions that modern societies built to convert uncertain cognition into trustworthy, actionable judgment.

The Puzzle of Expensive Answers

A client asks a law firm whether a proposed transaction will survive antitrust scrutiny. The opinion will cost two hundred thousand dollars and take three weeks. The client knows, roughly, what the answer will say. She has read the statute; the question is not exotic. In 2026, a capable large language model can draft a plausible memorandum in minutes. So what, exactly, is she buying?

Not the text, certainly not just the text. She is buying an answer that someone is prepared to stand behind. If the transaction is later challenged, the firm's name, its malpractice insurance, and its partners' professional judgment are on the line. The expensive part of expert work was not only the production of words. It was the institutional machinery that made those words trustworthy: the checking, the contextual judgment, the willingness to bear consequences, and, more invisibly, the decades of training that produced people able to do so.

[1]Department of Economics, Cornell University. Email: rohitlamba@cornell.edu. An extended version of this essay, with formal treatment and fuller literature engagement, is in preparation.

1

Pourquoi vous avez toujours besoin d'apprendre, comprendre, et pas juste « savoir utiliser l'IA »

- Parce-que les embauches en début de carrière dans les secteurs touchés par l'IA (la tech au 1er plan) sont diminuées
- Or les modèles d'IA ne sont pas fiables (code, médecine, logistique, soutien psychologique, etc.)

→ Donc un usage professionnel requiert de pouvoir vérifier leurs résultats

→ Donc savoir bien utiliser l'IA implique de savoir faire sans IA

- Dans le monde du travail, un humain doit assumer la responsabilité (morale, voire juridique) d'un résultat/une production

→ Si vous ne savez faire qu'avec l'IA, alors vous êtes remplaçable pour le début de carrière, et si vous êtes remplaçable, vous n'êtes pas gardé·e pour être formé·e et avancer en carrière.

Plan du cours

1. Principes de base des LLM
   - Reproduire les statistiques de co-apparition des mots
   - Conséquence : manque de fiabilité (biais, productivité, cyber-sécurité)

2. L'apprentissage étudiant et les LLM
   - Pourquoi vous avez toujours besoin d'apprendre, comprendre, et pas juste « savoir utiliser l'IA »
   - Vous êtes à l'université pour « apprendre », pas pour « produire »
   - Pourquoi pour apprendre, vous devez éviter d'utiliser l'IA

3. Discussion
   - Exploitation humaine, environnementale, productivité et emplois, ...

Donc vous devez toujours acquérir des connaissances dans votre cerveau donc apprendre, et C'est pour ça que vous êtes à l'université pour « apprendre », pas pour « produire »

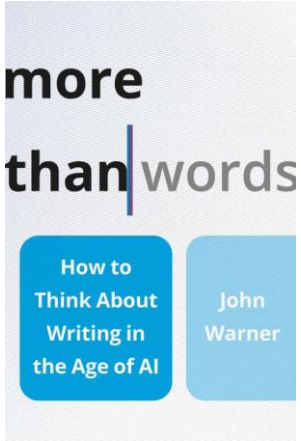
- But de l'enseignement et Intérêt de l'élève: apprendre, donc produire l'effort incontournable pour créer les circuits neuronaux implantant les nouvelles connaissances et compétences
- Moyen sans IA : note sur la production pour inciter et s'assurer que le processus a été réalisé

→ L'usage de IA 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.

Intuition : importance de notre processus d'écrire pour réflechir

- *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.*



John Warner, Hachette Eds., 2025
https://www.hachettebookgroup.com/titles/john-warner/more-than-words/9781541605510

- 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'ont 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 ou dont on ressent le besoin, on se fait un avis, on identifie ce qui ne va pas dans notre raisonnement ou nos manques dans la compréhension.

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

Plan du cours

1. Principes de base des LLM
   - Reproduire les statistiques de co-apparition des mots
   - Conséquence : manque de fiabilité (biais, productivité, cyber-sécurité)

2. L'apprentissage étudiant et les LLM
   - Pourquoi vous avez toujours besoin d'apprendre, comprendre, et pas juste « savoir utiliser l'IA »
   - Vous êtes à l'université pour « apprendre », pas pour « produire »
   - Pourquoi pour apprendre, vous devez éviter d'utiliser l'IA

3. Discussion
   - Exploitation humaine, environnementale, productivité et emplois, ...

Pourquoi pour apprendre, vous devez éviter d'utiliser l'IA

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.

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.

M. Burns, R. Winthrop, N. Luther, E. Venetis, and R. Karim, "A new direction for students in an AI world: Prosper, prepare, protect," Brookings, Jan. 2026.

N. Kosmyna, et al., "Your Brain on ChatGPT: Accumulation of Cognitive Debt when Using an AI Assistant for Essay Writing Task," June 2025

D. Stromberg, V. Lei and Y. Wu, "The Generative AI Learning Penalty: Evidence from Chinese Secondary Education", SSRN, June 2026.

Synthèse des études

- Après des entretiens, des groupes de discussion et des consultations auprès de plus de 500 élèves, enseignant·es, parents, responsables éducatifs et expert·es en technologies dans 50 pays, un examen approfondi de plus de 400 études et une consultation par panel Delphi, nous constatons qu'à ce stade de son développement, les risques liés à l'utilisation de l'IA générative dans l'éducation l'emportent sur ses avantages.
- Dans de nombreux domaines utilisant l'IA, ses résultats productifs et analytiques sont remarquables. Les scientifiques ont utilisé l'IA pour accélérer le processus de cartographie des protéines humaines, permettant ainsi le développement de nouveaux traitements médicaux. Cependant, les personnes qui exploitent ces puissants outils d'IA sont des adultes professionnels dont le cerveau est pleinement développé. Elles ont déjà acquis des compétences métacognitives et un esprit critique sophistiqués qui sous-tendent leur approche du travail. **Elles possèdent une expertise approfondie dans leurs domaines respectifs, ainsi que la flexibilité cognitive qui en découle, leur permettant de naviguer, d'évaluer, d'assimiler, de rejeter et d'utiliser stratégiquement les informations générées par l'IA.**
- Pour les élèves, la situation est fondamentalement inverse. Ils ne sont pas des mini-professionnels.
  Leur cerveau se développe, subissant des processus cruciaux d'élagage et de renforcement neuronal qui dépendent d'efforts cognitifs répétés et d'une lutte constante. [...]
  L'école existe précisément pour développer ces capacités grâce à un engagement soutenu avec des contenus stimulants.
- Les enseignant·es constatent, et les études récentes le confirment, que la facilité avec laquelle on peut déléguer des tâches aux outils d'IA a un impact négatif sur le plaisir d'apprendre des élèves, leur curiosité, leur confiance en soi, leur sentiment d'efficacité personnelle, leur sentiment d'autonomie et leur capacité de réflexion éthique et de résolution de problèmes nuancée.

> After interviews, focus groups, and consultations with over 500 students, teachers, parents, education leaders, and technologists across 50 countries, a close review of over 400 studies, and a Delphi panel, we find that at this point in its trajectory, the risks of utilizing generative AI in children's education overshadow its benefits.
>
> In many fields using AI, its productive and analytical outcomes are striking. Scientists have used AI to accelerate the process of mapping human proteins enabling new medical treatments to be developed. But the people harnessing these powerful AI tools are professional adults with fully matured brains. They have already developed sophisticated metacognitive and critical thinking skills that undergird their approach to their work. They have deep expertise in their domains, and the cognitive flexibility that comes with such expertise, allowing them to navigate, evaluate, assimilate, reject, and strategically use the information AI generates.
>
> For students, the situation is fundamentally reversed. They are not mini-professionals. Their brains are developing, undergoing crucial processes of neural pruning and strengthening that depend on repeated cognitive effort and struggle. [...] School exists precisely to build these capacities through sustained engagement with challenging material.
>
> Teachers report, and emerging literature confirms, that the ease of offloading work to AI is negatively impacting students' love of learning, curiosity, self-confidence, sense of self-efficacy, sense of agency, and capacity for ethical reflection and nuanced problem solving.

M. Burns, R. Winthrop, N. Luther, E. Venetis, and R. Karim, "**A new direction for students in an AI world: Prosper, prepare, protect**," Brookings, Jan. 2026.

Votre cerveau avec ChatGPT: accumulation de dette cognitive lors de l’usage d’un outil IA pour une rédaction

- 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 ue 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△

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

United States

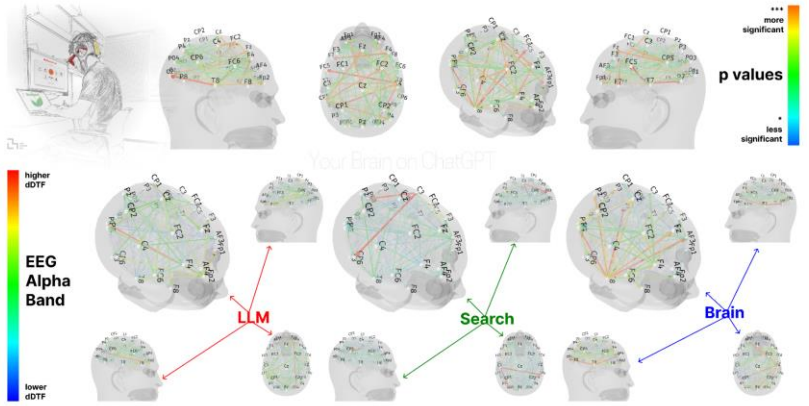
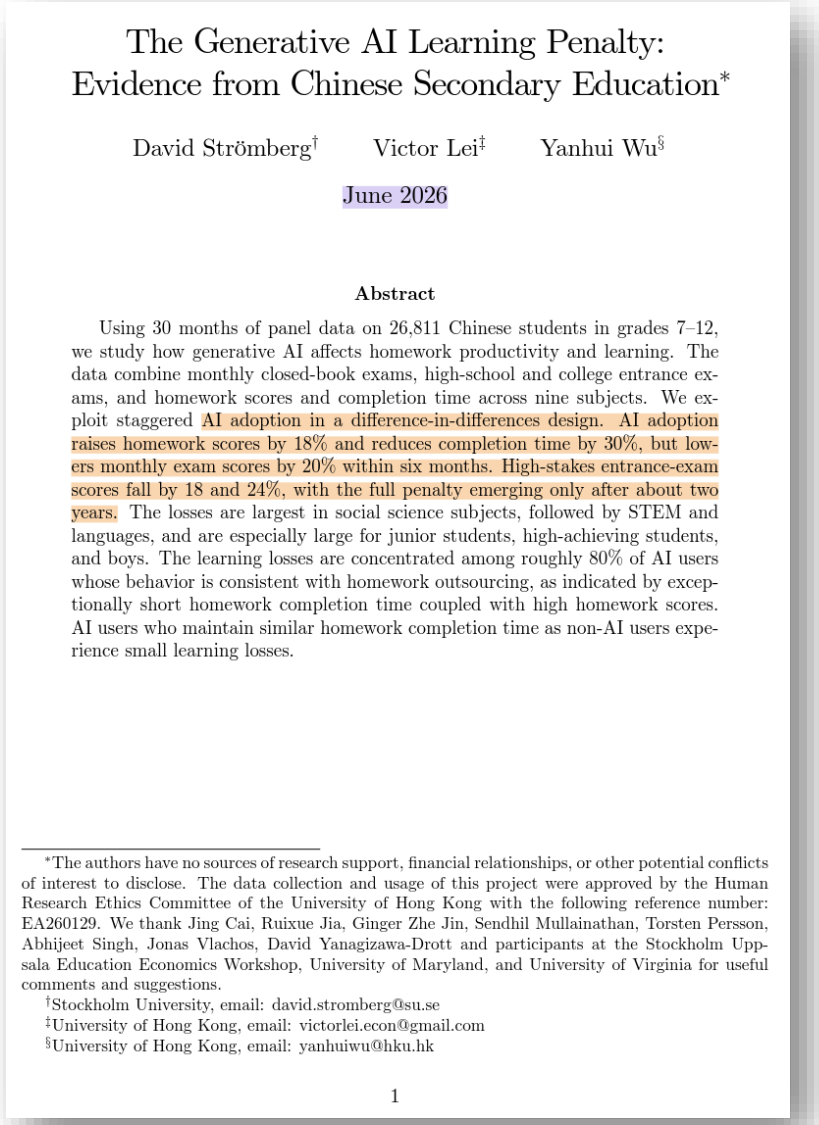


Figure 1. The dynamic Direct Transfer Function (dDTF) EEG analysis of Alpha Band for groups: LLM, Search Engine, Brain-only, including p-values to show significance from moderately significant (*) to highly significant (***).

L’effet pervers d’apprendre avec l’IA générative

- 28811 élèves de la 5ème à la Terminale
- Devoir maison et contrôles en classe
- Examen d’entrée au lycée, examen d’entrée à l’université



The Generative AI Learning Penalty:
Evidence from Chinese Secondary Education*

David Strömberg† Victor Lei‡ Yanhui Wu§

June 2026

Abstract

Using 30 months of panel data on 26,811 Chinese students in grades 7–12, we study how generative AI affects homework productivity and learning. The data combine monthly closed-book exams, high-school and college entrance exams, and homework scores and completion time across nine subjects. We exploit staggered AI adoption in a difference-in-differences design. AI adoption raises homework scores by 18% and reduces completion time by 30%, but lowers monthly exam scores by 20% within six months. High-stakes entrance-exam scores fall by 18 and 24%, with the full penalty emerging only after about two years. The losses are largest in social science subjects, followed by STEM and languages, and are especially large for junior students, high-achieving students, and boys. The learning losses are concentrated among roughly 80% of AI users whose behavior is consistent with homework outsourcing, as indicated by exceptionally short homework completion time coupled with high homework scores. AI users who maintain similar homework completion time as non-AI users experience small learning losses.

*The authors have no sources of research support, financial relationships, or other potential conflicts of interest to disclose. The data collection and usage of this project were approved by the Human Research Ethics Committee of the University of Hong Kong with the following reference number: EA260129. We thank Jing Cai, Ruixue Jia, Ginger Zhe Jin, Sendhil Mullainathan, Torsten Persson, Abhijeet Singh, Jonas Vlachos, David Yanagizawa-Drott and participants at the Stockholm Uppsala Education Economics Workshop, University of Maryland, and University of Virginia for useful comments and suggestions.

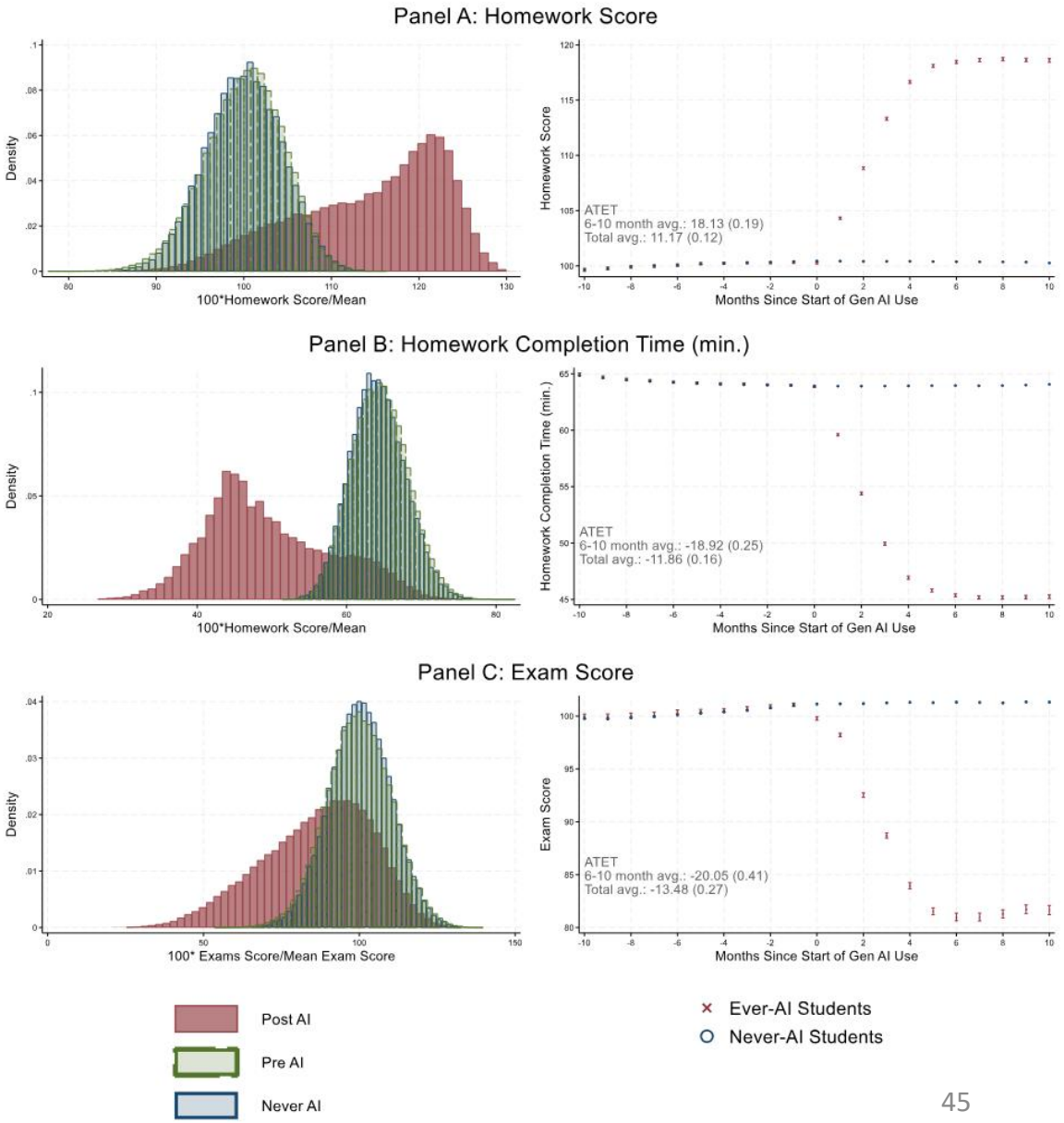
†Stockholm University, email: david.stromberg@su.se
‡University of Hong Kong, email: victorlei.econ@gmail.com
§University of Hong Kong, email: yanhuiwu@hku.hk

1

https://blogs.worldbank.org/en/investinpeople/a-warning-shot-for-human-capital--evidence-of-an-ai-learning-pen

Devoirs maison et révision vs notes aux DS

- Divergence marquée entre la productivité pour les devoirs maison et les résultats d'apprentissage = les notes aux examens sans docs.
- Devoir maison après 6 mois d'utilisation de l'IA gen:
  - les notes augmentent de 18 %
  - le temps consacré diminue de 30%
- MAIS les notes aux devoirs surveillés (DS - examens mensuels) chutent de 20 %



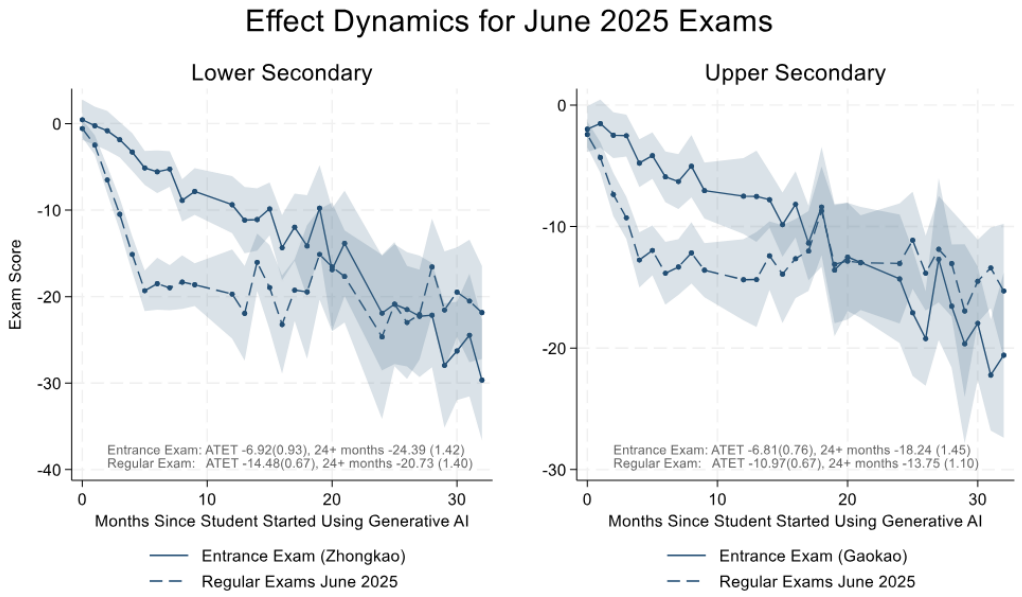
D. Stromberg, V. Lei and Y. Wu, "The Generative AI Learning Penalty: Evidence from Chinese Secondary Education", SSRN, June 2026

Effets pire à long terme

DS : -20% en 6 mois, Examens importants : -24% en 2 ans

- L'effet négatif de l'IA générative sur l'apprentissage à long terme (mesuré par les résultats aux examens d'entrée à forts enjeux) est tout aussi important, mais il s'accumule plus lentement :
  - Notes aux DS se détériorent complètement dans les 6 mois suivant l'adoption de l'IA
  - C'est au bout de 2 ans d'usage que l'impact sur les examens d'entrée (lycée et univ) atteint son max : -24%

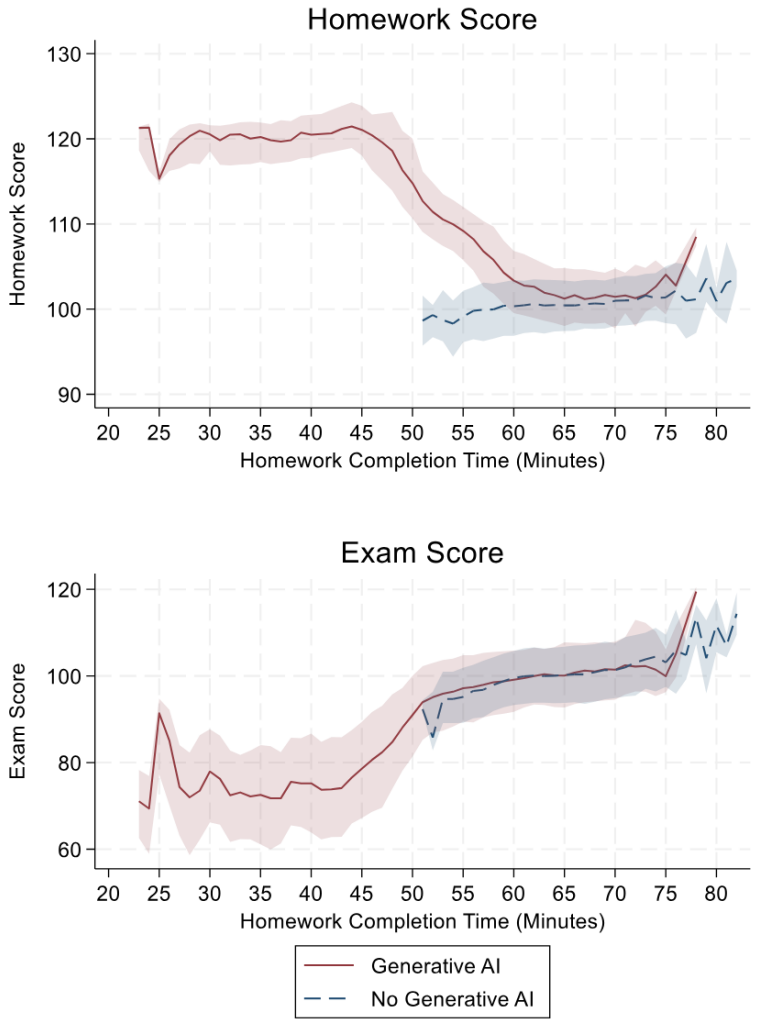
→ Les études jusqu'ici de courte durée sous-estiment systématiquement le coût de l'IA générative en termes d'apprentissage à long terme.



D. Stromberg, V. Lei and Y. Wu, "The Generative AI Learning Penalty: Evidence from Chinese Secondary Education", SSRN, June 2026

Avec IA : moins bonnes notes pour 81% des élèves
Pour les autres 19% : pas meilleures notes que sans IA

- Effet négatif dû à 81% des utilisateurices de l'IA qui externalise leurs devoirs : prennent beaucoup moins de temps à les faire que celles et ceux n'utilisant pas l'IA, font des meilleurs rendus, mais obtiennent de mauvaises notes aux DS.
- Pour celleux utilisant l'IA d'une façon à consacrer autant de temps à leurs révisions que sans IA, iels obtiennent des notes similaires aux DS :
  - C'est donc le temps effectivement passer à faire l'effort d'apprendre (= faire l'effort pour créer les connexions neuronales) qui compte, avec ou sans IA générative.
- Mais on ne voit pas d'effet bénéfique de l'usage de l'IA sur les notes.

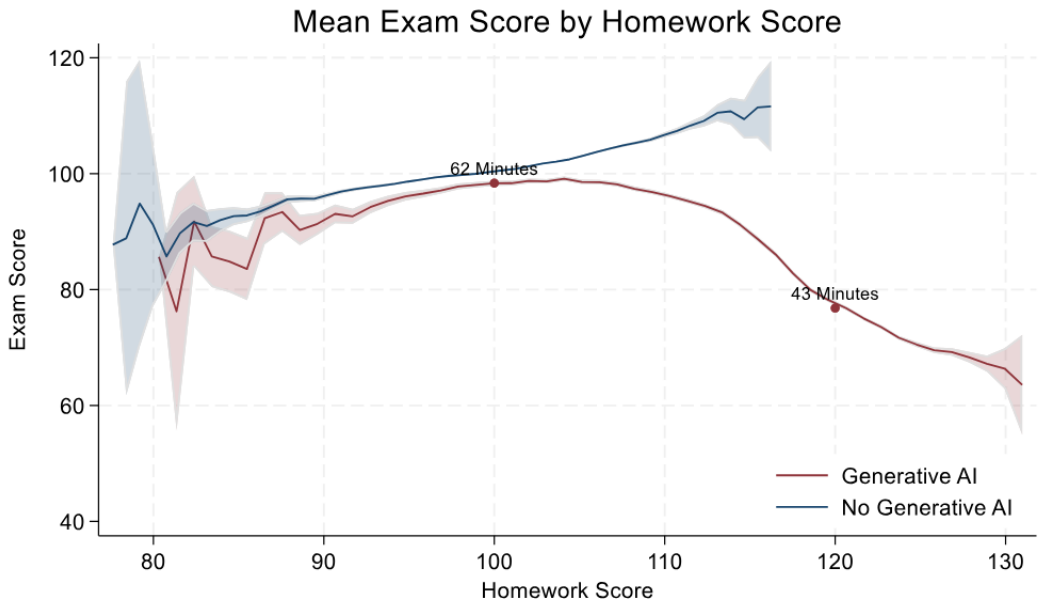


D. Stromberg, V. Lei and Y. Wu, "The Generative AI Learning Penalty: Evidence from Chinese Secondary Education", SSRN, June 2026

Avec IA : les notes obtenues en DM ne prédisent plus les notes aux DS

- Dans un contexte éducatif traditionnel, les résultats aux devoirs maison prédisent de manière fiable les résultats aux examens, reflétant la pratique et la consolidation des connaissances acquises grâce à la réalisation des tâches d'apprentissage.
- Cependant, chez les élèves utilisant l'IA et dont les résultats aux devoirs sont supérieurs à la moyenne de référence, des résultats plus élevés aux devoirs sont associés à des résultats plus faibles aux examens.

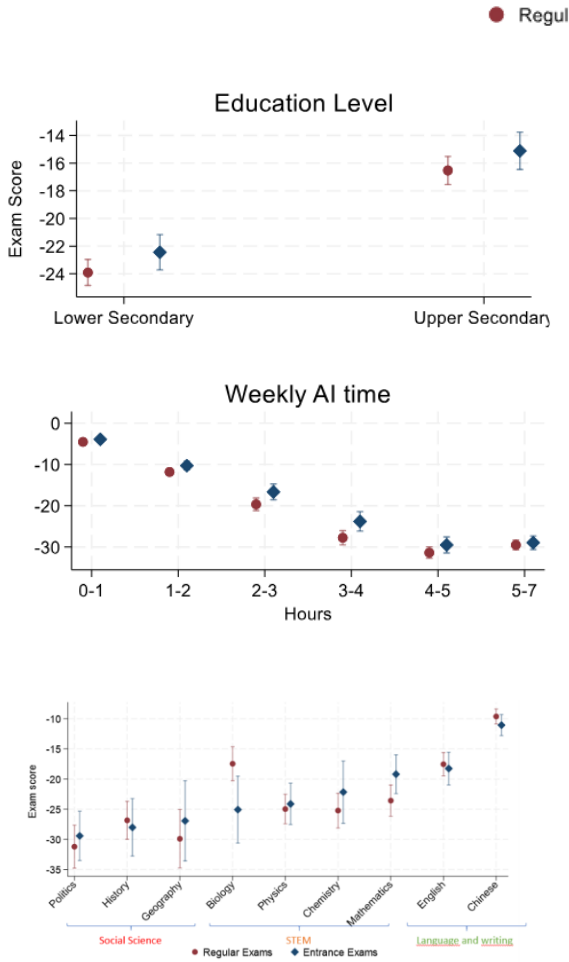
→ Ces élèves sont plus dépendant·es à l'assistance de l'IA sans avoir aussi bien appris : modèle d'« externalisation des devoirs » par opposition au « tutorat » par les outils IA.



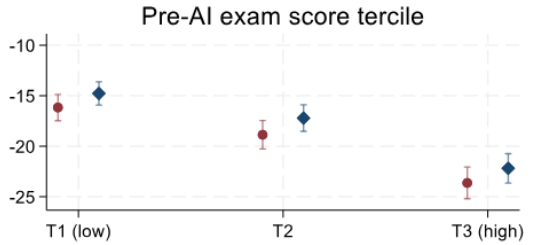
D. Stromberg, V. Lei and Y. Wu, "The Generative AI Learning Penalty: Evidence from Chinese Secondary Education", SSRN, June 2026

L'effet négatif d'utiliser l'IA pour apprendre est variable en fonction de:

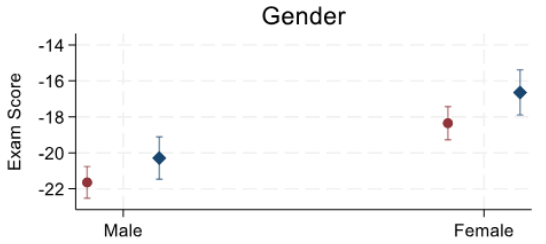
- du niveau d'étude : plus négatif pour les débutant·es
- du temps que vous y passez : plus on l'utilise, moins on a de bonnes notes
- des disciplines : plus négatif en sciences et technologies



- de votre niveau : celleux dans les meilleurs 30% sans IA subissent 2x plus d'effet négatif



- de votre genre : les garçons ont leurs notes qui baissent plus



D. Stromberg, V. Lei and Y. Wu, "The Generative AI Learning Penalty: Evidence from Chinese Secondary Education", SSRN, June 2026

Utiliser l'IA pour ses TD et TP empêche la pratique nécessaire à acquérir les compétences
→ La construction progressive de l'édifice des connaissances est empêchée.

- Kosmyna et al. (2025, slide 43) : la rédaction de dissertations à l'aide de ChatGPT nuit à la capacité des utilisateuricess à se remémorer leurs propres arguments et observent une réduction de la connectivité neuronale dans le cerveau lors de la rédaction.
- Il a été montré en psychologie que l'entraînement par tests — y compris les exercices de type devoirs, TD refaits à la maison, pratique en TP — améliore la mémorisation et suscite un effort cognitif plus important qu'une étude passive.

→Les résultats présentés suggèrent que les étudiant·es utilisant l'IA se privent de cet effet bénéfique lié à la pratique.

→De plus, le constat d'une baisse progressive des notes aux DS au fil du temps suggère déléguer ses exercices maisons à l'IA compromet non seulement la mémorisation du contenu actuel, mais aussi l'assimilation des connaissances futures : les effets s'accumulent, en acquérant connaissances et compétences on peut acquérir les suivantes, ne pas acquérir les premières empêche le processus.

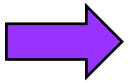
→ Ces résultats très récents montrent que l'IA générative, même si son usage est très courant, a un impact négatif considérable sur l'apprentissage des étudiant·es.

D. Stromberg, V. Lei and Y. Wu, "The Generative AI Learning Penalty: Evidence from Chinese Secondary Education", SSRN, June 2026

Plan du cours

1. Principes de base des LLM
   - Reproduire les statistiques de co-apparition des mots
   - Conséquence : manque de fiabilité (biais, productivité, cyber-sécurité)

2. L'apprentissage étudiant et les LLM
   - Pourquoi vous avez toujours besoin d'apprendre, comprendre, et pas juste « savoir utiliser l'IA »
   - Vous êtes à l'université pour « apprendre », pas pour « produire »
   - Pourquoi pour apprendre, vous devez éviter d'utiliser l'IA



3. Conclusion

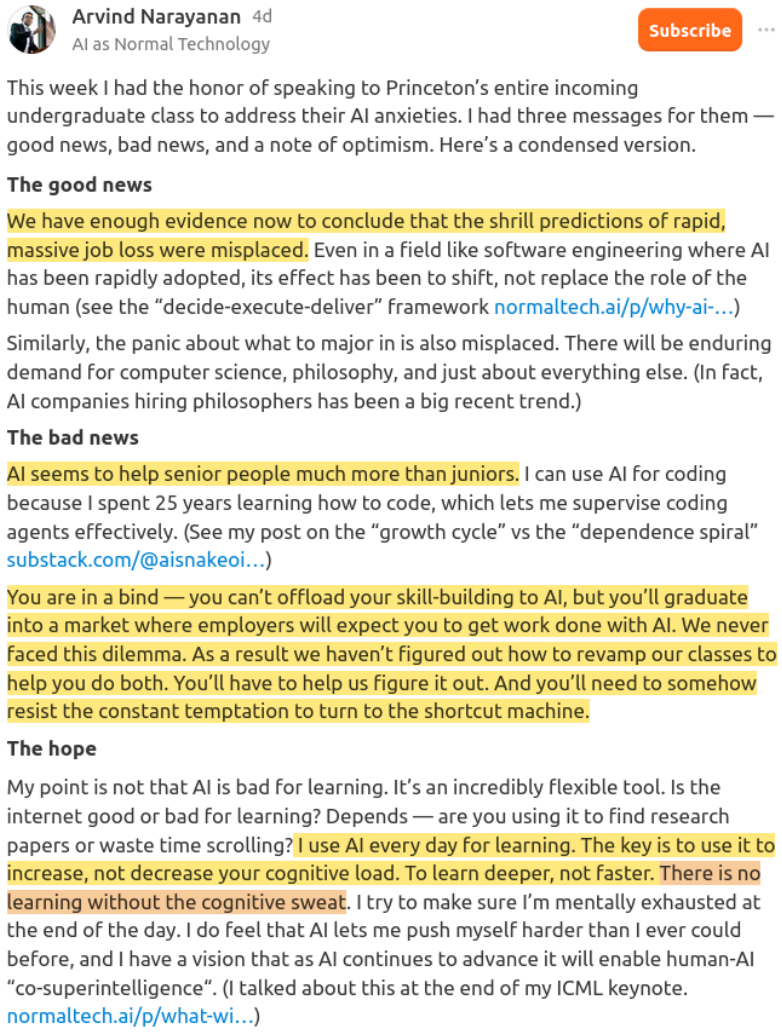
4. Discussion
   - Exploitation humaine, environnementale, productivité et emplois, ...

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 pour quoi vous faîtes la tâche que vous pensez déléguer, qu'est-ce que vous voulez personnellement (cérébralement) acquérir, qu'est-ce qui va vous 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 court ou moyen terme.

Rentrée 2026 1ères années Princeton



Arvind Narayanan 4d
AI as Normal Technology

Subscribe

This week I had the honor of speaking to Princeton's entire incoming undergraduate class to address their AI anxieties. I had three messages for them — good news, bad news, and a note of optimism. Here's a condensed version.

The good news

We have enough evidence now to conclude that the shrill predictions of rapid, massive job loss were misplaced. Even in a field like software engineering where AI has been rapidly adopted, its effect has been to shift, not replace the role of the human (see the "decide-execute-deliver" framework normaltech.ai/p/why-ai-…)

Similarly, the panic about what to major in is also misplaced. There will be enduring demand for computer science, philosophy, and just about everything else. (In fact, AI companies hiring philosophers has been a big recent trend.)

The bad news

AI seems to help senior people much more than juniors. I can use AI for coding because I spent 25 years learning how to code, which lets me supervise coding agents effectively. (See my post on the "growth cycle" vs the "dependence spiral" substack.com/@aisnakeoi…)

You are in a bind — you can't offload your skill-building to AI, but you'll graduate into a market where employers will expect you to get work done with AI. We never faced this dilemma. As a result we haven't figured out how to revamp our classes to help you do both. You'll have to help us figure it out. And you'll need to somehow resist the constant temptation to turn to the shortcut machine.

The hope

My point is not that AI is bad for learning. It's an incredibly flexible tool. Is the internet good or bad for learning? Depends — are you using it to find research papers or waste time scrolling? I use AI every day for learning. The key is to use it to increase, not decrease your cognitive load. To learn deeper, not faster. There is no learning without the cognitive sweat. I try to make sure I'm mentally exhausted at the end of the day. I do feel that AI lets me push myself harder than I ever could before, and I have a vision that as AI continues to advance it will enable human-AI "co-superintelligence". (I talked about this at the end of my ICML keynote. normaltech.ai/p/what-wi…)

- Nous disposons désormais de suffisamment d'éléments pour conclure que les prédictions alarmistes faisant état de pertes d'emplois rapides et massives étaient infondées.
- L'IA semble aider beaucoup plus les profils seniors que les juniors.
- Vous vous trouvez dans une situation délicate : vous ne pouvez pas déléguer l'acquisition de vos compétences à l'IA, mais vous arriverez sur le marché du travail à une époque où les employeurs attendront de vous que vous sachiez accomplir vos tâches en l'utilisant. Nous n'avons jamais eu à affronter ce dilemme. Par conséquent, nous n'avons pas encore trouvé comment repenser nos cours pour vous permettre de concilier ces deux impératifs. Il vous faudra nous aider à trouver la solution. Et vous devrez, d'une manière ou d'une autre, résister à la tentation constante de recourir à cette machine à raccourcis.
- J'utilise l'IA au quotidien pour apprendre. L'essentiel est de s'en servir pour accroître sa charge cognitive, et non pour la réduire. Pour apprendre plus en profondeur, et non plus vite. Il n'y a pas d'apprentissage sans effort cognitif.

Interdiction de l'IA dans les écoles et collèges à NYC

Reuters

World Business Markets Sustainability Legal Commentary More My News Sign In

Mamdani imposes one-year ban on AI for most NYC students

By **Reuters**

September 2, 2026 9:53 PM GMT+2 · Updated 19 hours ago

NEW YORK, Sept 2 (Reuters) - New York City officials announced a one-year moratorium on students using artificial intelligence in public elementary and middle schools, the latest step by public officials to set guidelines around new technology in the classroom.

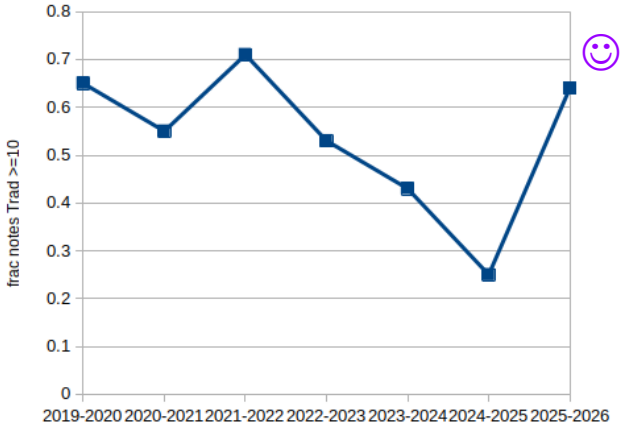
C’est pourquoi :

- Contrat pédagogique par l’enseignant·e pour chaque module
- Manifeste « IA et enseignement supérieur » et Charte d’utilisation des outils d’IA générative dans l’enseignement votées par le conseil académique de l’université
- Ressources pour les étudiant·es et enseignant·es dans Moodle : votre Espace IA étudiant, pour que vous compreniez mieux l’IA et ses impacts sur votre apprentissage

Contrat pédagogique

- PAS d'UTILISATION D'OUTILS D'IAG acceptée ni recommandée pour la majorité des modules techniques en RT.
- DS :
  - écrit 1h30
  - sur cours, TD ET TP
  - coefficient : **~0.7**
- Interrogations en fin de CM :
  - ~3 questions à la fin de chaque cours
  - coefficient : **0.1**
- QCM Moodle :
  - test à compléter le soir de chaque amphi
  - coefficient : **0.1**
- TP :
  - Interrogation orale en TP : **0.1**
  - Questions de TP notées en DS
  - pénalité d'absence TP sur 20% de la note DS

Fraction de notes ≥10 en R204 Promo Trad



Manifeste UniCA « IA et enseignement supérieur »

Manifesto UniCA

Le manifesto "IA dans l'enseignement supérieur" énonce la position de principe de Université Côte d'Azur sur l'IA dans l'enseignement supérieur. Il a été adopté par vote du Conseil Académique de l'établissement le 18 décembre 2025.

Ce texte est motivé par les questionnements intenses et les besoins de repères de la communauté universitaire, étudiante et enseignante, dans un contexte d'incertitude sur les meilleures pratiques à adopter avec les outils d'IA générative. Ce texte s'ancre dans un devoir de précaution qui considère non pas les conditions potentiellement idéales d'utilisation, mais les conditions effectives des vécus étudiants et enseignants de ces outils. Ce texte s'articule en effet avec différentes revues de l'état des connaissances sur le sujet IA et apprentissage [1,2].

La rédaction de ce texte a été initiée par la direction scientifique de EFELIA Côte d'Azur et travaillée en concertation avec la Vice-Présidente Transformations pédagogiques et formation tout au long de la vie, le Vice-Président Transformations numériques, et le Vice-Président Formation et innovation pédagogique.

Le positionnement de ce texte est le suivant : il vient en complément d'une charte en cours d'élaboration au niveau national et dont le lien est mentionné, ainsi que d'une charte en cours d'élaboration au niveau local par le groupe de travail IA & enseignement créé par la Vice-Présidente Transformations pédagogiques. Ce texte n'est pas de même nature que la charte : la charte vise à réguler l'usage, tandis que ce manifesto énonce une position de principe institutionnel.

Il s'accompagne du présent centre de ressource où toutes les documentations pertinentes et perspectives outils peuvent être trouvées.

[1] Julie M. Smith, Jesse Dukes, Josh Sheldon, Manee Ngozi Nnamani, Natasha Esteves, and Justin Reich (2025). A Guide to AI in Schools: Perspectives for the Perplexed. Retrieved from https://tsl.mit.edu/ai-guidebook/

[2] M. Burns, R. Winthrop, N. Luther, E. Venetis, and R. Karim, "A new direction for students in an AI world: Prosper, prepare, protect," Brookings, Jan. 2026. Available: https://www.brookings.edu/articles/a-new-direction-for-students-in-an-ai-world-prosper-prepare-protect/

Devant l'utilisation toujours plus importante d'outils d'Intelligence Artificielle (IA) par la communauté universitaire dans son ensemble, et dans un contexte de déploiements souvent inconditionnels des outils, **UniCA, en tant que Grand Etablissement, énonce, à travers le texte ci-dessous, sa position au sujet de l'IA dans l'enseignement supérieur**[1].
Cette démarche a été initiée à la rentrée 2025 par la direction scientifique d'EFELIA Côte d'Azur.

Une des préoccupations premières de UniCA est d'assurer la meilleure qualité possible des conditions d'apprentissage pour ses usagères et usagers, et ce dans les meilleures conditions de travail possibles pour ses personnels. UniCA considère que les technologies telles que l'IA déployée dans des outils à destination des personnes étudiantes ou enseignantes doivent être considérées comme des moyens, possibles mais non évidents ni obligatoires, dans le cadre de la réalisation des missions premières de l'université : produire et communiquer des connaissances pour former des citoyennes et citoyens développant leurs capacités de réflexion analytique et leur esprit critique, transmettre des compétences pour leur insertion dans le monde professionnel.

Dans ce cadre, UniCA reconnaît :

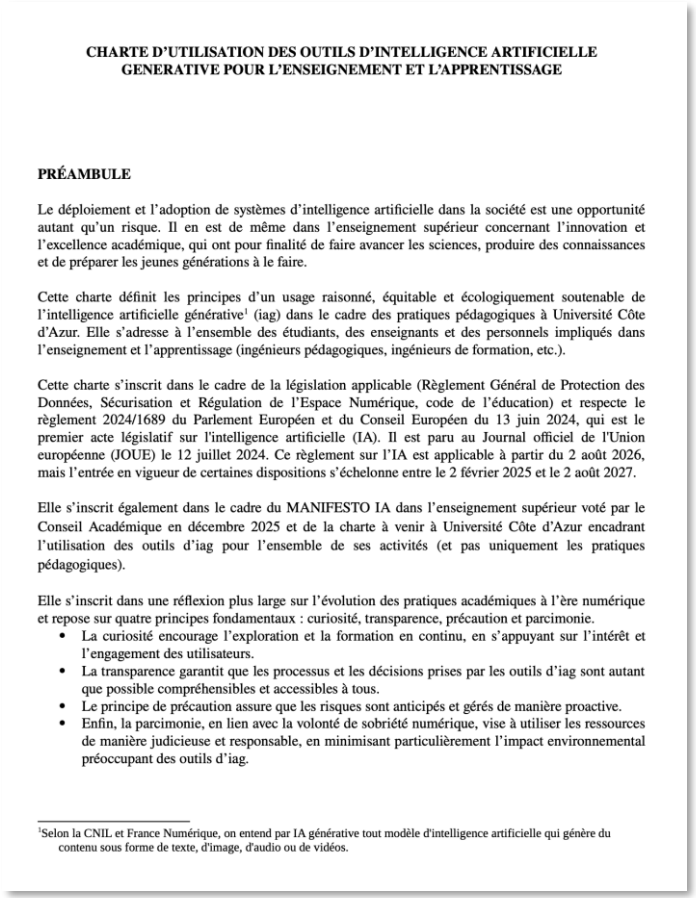
- que la formation initiale des personnes étudiantes ainsi que le développement professionnel de l'ensemble des personnels doivent permettre de développer une vision critique et réflexive pour déterminer quand l'usage d'outils d'IA est ou non approprié. Ces décisions doivent reposer sur l'explicitation des objectifs d'apprentissage visés ainsi que sur d'autres dimensions (risque de dégradation d'apprentissage, coût environnemental, respect de la propriété intellectuelle d'autrui, respect des politiques de collecte de données, risque d'accroissement des inégalités parmi les apprenantes et apprenants).
- l'absence de preuve scientifique, à ce jour, d'efficacité des outils d'IA pour réaliser ses missions premières susmentionnées, bien qu'à titre individuel on puisse avoir une intuition contraire. Des travaux scientifiques récents montrent que l'usage d'outils d'IA peuvent nuire aux processus d'apprentissage chez les personnes étudiantes, et peuvent inquiéter sur l'effectivité de l'apprentissage réel lorsque ces outils sont utilisés de manière régulière et intense. Des travaux scientifiques montrent également que l'usage de ces outils modifie, sans en améliorer la qualité, la nature et la charge de travail des métiers de la connaissance (comme celui d'enseigner) ainsi que la satisfaction associée.
- comme pour tout autre outil technologique d'assistance à la pédagogie, que le recours à des outils d'IA ne peut et ne doit pas être rendu obligatoire pour les personnes enseignantes, qui doivent pouvoir décider de ne pas les utiliser pour leurs enseignements.
- que le corps enseignant, à différents niveaux (individuel, Conseil de Perfectionnement, département disciplinaire, COSP d'EUR) et en collégialité, est le seul qualifié pour décider de la pertinence d'outils pédagogiques d'IA et de ses modalités d'usage. Chaque responsable d'enseignement garde la liberté de décider de les autoriser ou de demander aux personnes étudiantes de ne pas utiliser ces outils, à la lumière des objectifs d'apprentissage et des modalités d'évaluation choisies, à condition que ces outils soient utilisés dans les cadres légaux et éthiques établis par la gouvernance et les directions de l'université (DSI et DF en particulier).
- que tout déploiement d'outils d'IA au niveau de l'établissement sera décidée de façon collective, dans le cadre des espaces pertinents de concertation et en prenant appui sur ses Conseils.

Un nouveau centre de ressources sur l'IA pour l'enseignement supérieur est mis en place pour les personnels à ce lien : https://moodle-formation.univ-cotedazur.fr/course/view.php?id=862. On y trouvera notamment les références académiques qui ne sont que suggérées ici, au vu du caractère volontairement très synthétique de ce document.

1 Ce document est complémentaire à l'adoption par UniCA de la charte nationale sur l'usage de l'IAG dans l'ESR https://www.demoes-pro3.fr/proposition-charte-ia-esr/

https://univ-cotedazur.fr/medias/fichier/cac-2025-33-manifesto-efelia_1768308447307-pdf

Charte UniCA d'utilisation des outils d'IA générative pour l'enseignement et l'apprentissage



CHARTE D'UTILISATION DES OUTILS D'INTELLIGENCE ARTIFICIELLE GENERATIVE POUR L'ENSEIGNEMENT ET L'APPRENTISSAGE

PRÉAMBULE

Le déploiement et l'adoption de systèmes d'intelligence artificielle dans la société est une opportunité autant qu'un risque. Il en est de même dans l'enseignement supérieur concernant l'innovation et l'excellence académique, qui ont pour finalité de faire avancer les sciences, produire des connaissances et de préparer les jeunes générations à le faire.

Cette charte définit les principes d'un usage raisonné, équitable et écologiquement soutenable de l'intelligence artificielle générative[1] (iag) dans le cadre des pratiques pédagogiques à Université Côte d'Azur. Elle s'adresse à l'ensemble des étudiants, des enseignants et des personnels impliqués dans l'enseignement et l'apprentissage (ingénieurs pédagogiques, ingénieurs de formation, etc.).

Cette charte s'inscrit dans le cadre de la législation applicable (Règlement Général de Protection des Données, Sécurisation et Régulation de l'Espace Numérique, code de l'éducation) et respecte le règlement 2024/1689 du Parlement Européen et du Conseil Européen du 13 juin 2024, qui est le premier acte législatif sur l'intelligence artificielle (IA). Il est paru au Journal officiel de l'Union européenne (JOUE) le 12 juillet 2024. Ce règlement sur l'IA est applicable à partir du 2 août 2026, mais l'entrée en vigueur de certaines dispositions s'échelonne entre le 2 février 2025 et le 2 août 2027.

Elle s'inscrit également dans le cadre du MANIFESTO IA dans l'enseignement supérieur voté par le Conseil Académique en décembre 2025 et de la charte à venir à Université Côte d'Azur encadrant l'utilisation des outils d'iag pour l'ensemble de ses activités (et pas uniquement les pratiques pédagogiques).

Elle s'inscrit dans une réflexion plus large sur l'évolution des pratiques académiques à l'ère numérique et repose sur quatre principes fondamentaux : curiosité, transparence, précaution et parcimonie.

- La curiosité encourage l'exploration et la formation en continu, en s'appuyant sur l'intérêt et l'engagement des utilisateurs.
- La transparence garantit que les processus et les décisions prises par les outils d'iag sont autant que possible compréhensibles et accessibles à tous.
- Le principe de précaution assure que les risques sont anticipés et gérés de manière proactive.
- Enfin, la parcimonie, en lien avec la volonté de sobriété numérique, vise à utiliser les ressources de manière judicieuse et responsable, en minimisant particulièrement l'impact environnemental préoccupant des outils d'iag.

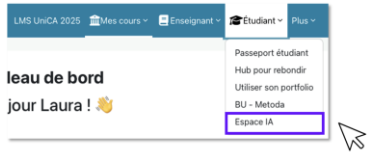
[1]Selon la CNIL et France Numérique, on entend par IA générative tout modèle d'intelligence artificielle qui génère du contenu sous forme de texte, d'image, d'audio ou de vidéos.

- Adoptée en CAc le 29 janvier 2026
- Issue du GT IA enseignement et apprentissage menée par la VP Transformations Pédagogiques
- Principes généraux et responsabilités des étudiant·es et des enseignant·es

https://intranet.univ-cotedazur.fr/medias/fichier/cac-2026-01-charte-utilisation-iag_1770711950219-pdf?null&RH=&ksession=2656a456-daf5-4fd9-b411-ce6b50b63dc6&ksession=4e8c16cc-a936-4e89-b25c-8db3a0ad5c57

Ressource : votre Espace IA étudiant

- Inscription libre



Espace de référence IA

Ce **centre de ressources numériques sur l'intelligence artificielle (IA)** s'adresse aux étudiantes et étudiants d'Université Côte d'Azur. Il a pour objectif de vous aider à mieux comprendre ce que peuvent et ne peuvent pas faire les systèmes d'IA existants, ainsi que leurs impacts sociétaux et environnementaux, souvent encore peu visibilisés auprès du grand public, mais cruciaux à connaître.

Vous disposerez ainsi des éléments connus à ce jour pour décider par vous-même si, quand, et comment utiliser des outils d'IA dans votre processus d'apprentissage personnel.

Ces ressources sont constituées :

- des formations de référence à Université Côte d'Azur, pour mieux comprendre l'IA et ses enjeux,
- des documents de cadrage institutionnels,
- d'informations sur les actualités et événements en lien avec l'IA proposés par le projet EFELIA Côte d'Azur.

Formations repères

Ces formations en ligne ont été élaborées par le projet EFELIA Côte d'Azur et constituent, au sein de notre université, les ressources privilégiées pour se former en autonomie à l'IA et à ses enjeux. Elles vous permettent d'apprendre à votre rythme et ne sont pas notées*.

Inscrivez-vous dès maintenant !

**Sauf pour les étudiantes et étudiants de Licence 1 et de Licence 2 qui les suivent de manière obligatoire dans le cadre des compétences transversales.*

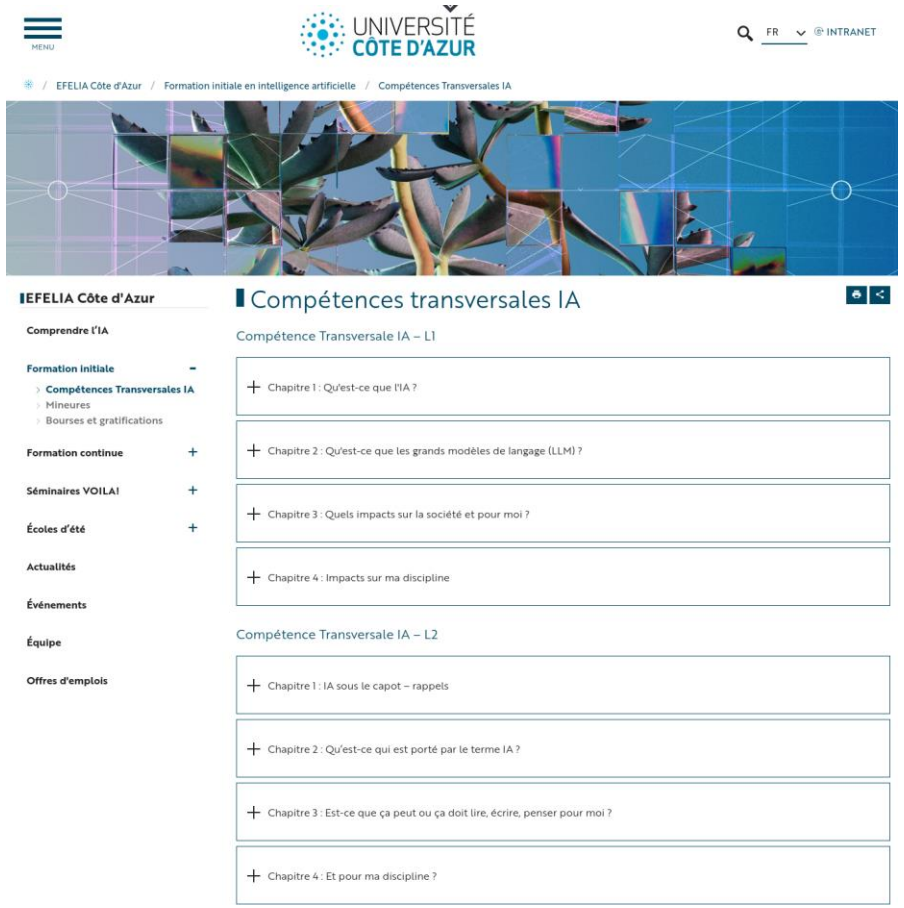

Grands défis sociétaux : IA L1 (🕓 10h)


Grands défis sociétaux : IA L2 (🕓 10h)


IA et apprentissage (🕓 1h)


IA générative, enseignement et recherche : quels points à considérer ? (🕓1h20)

https://lms.univ-cotedazur.fr/2025/course/view.php?id=19024



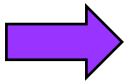
https://univ-cotedazur.fr/efelia-cote-dazur/formation-initiale/competences-transversales-ia-1

Plan du cours

1. Principes de base des LLM
   - Reproduire les statistiques de co-apparition des mots
   - Conséquence : manque de fiabilité (biais, productivité, cyber-sécurité)

2. L'apprentissage étudiant et les LLM
   - Pourquoi vous avez toujours besoin d'apprendre, comprendre, et pas juste « savoir utiliser l'IA »
   - Vous êtes à l'université pour « apprendre », pas pour « produire »
   - Pourquoi pour apprendre, vous devez éviter d'utiliser l'IA

3. Conclusion



4. Discussion
   - Exploitation humaine, environnementale, productivité et emplois, ...