



Compétence transversale – L2 Grands défis sociaux : 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

Modalités d'apprentissage

10 à 12h de travail personnel

Pour chacun des **4 chapitres** :

- plusieurs **vidéos**
 - accompagnées de leurs supports pdf
- des **questions** sous la forme de QCM au cours du chapitre
 - pour que vous sachiez si vous avez bien compris les notions importantes que vous venez d'aborder, non noté
 - permet de réviser les QCM notés constituant votre note finale
- des **manipulations pratiques** de systèmes d'IA
 - dans certains chapitres, n'impliquant aucune programmation, pour vous permettre de mieux comprendre
- un **forum** pour poser vos questions et échanger :
 - dans lequel pourrez interagir avec un ou une enseignante.



Modalités d'évaluation

- QCM 1 après le chapitre 2 :
 - coefficient 0.2, temps limité (quelques minutes), ouverture du 3 au 7 novembre, en ligne
- QCM 2 final après chapitre 4 (en décembre) :
 - coefficient 0.8, temps limité (quelques minutes), ouverture du 8 au 12 décembre, en présentiel



Ressources et vocabulaire

- Lectures

- AI Myths. <https://www.aimyths.org/>
- S. d'Acoli. Comprendre la révolution de l'intelligence artificielle. 2020
- Kate Crawford. Contre-atlas de l'intelligence artificielle. Zulma éditions, 2022.

(disponibles dans
vos BU)

- Vocabulaire

- ~~Une IA~~
- Un système d'IA (SIA)¹
- Un modèle d'IA
- Outils basés IA

¹ <https://declarationmontreal-iaresponsable.com/la-declaration/>

COMPRENDRE L'INTELLIGENCE ARTIFICIELLE

FR  

Sommaire

- Préambule
- Pour comprendre la diversité des méthodes d'IA, leurs grands principes et leurs évolutions historiques
- La révolution des années 2010 : l'apprentissage de représentations avec le Deep Learning
- L'IA générative
- Les limites des modèles de ML et les problèmes de biais
- Les stéréotypes sociaux reproduits par les modèles d'IA
- Besoin d'une réflexion éthique par toutes et tous ?
- Quelques ressources pour dépasser le discours dominant et amorcer un regard critique

Préambule : mieux comprendre l'intelligence artificielle 

Pour comprendre la diversité des méthodes d'IA, leurs grands principes et leurs évolutions historiques 

La révolution des années 2010 : l'apprentissage de représentations avec le Deep Learning 

L'IA générative 

Les limites des modèles de ML et les problèmes de biais 

Quelques ressources pour dépasser le discours dominant et amorcer un regard critique 



<https://univ-cotedazur.fr/efelia-cote-dazur/comprendre-ia>

Vous ?

- Que vous rappelez-vous du cours de l'an dernier ?
- Quels sont les aspects qui vous avaient marqué ?
- Quelles sont vos interrogations actuelles autour de l'IA ?



Vous ?

- Utilisez-vous des outils d'IA générative/chatbots ?
- Si non :
 - Pourquoi ne les utilisez-vous pas ?
- Si oui :
 - Pour quoi faire les utilisez vous ?
 - Pour quelle raison les utilisez vous ?
 - Qu'est-ce qui vous gêne ou ne vous satisfait pas quand vous les utilisez pour vos études ? Par exemple en lien avec
 - les raisons pour lesquelles vous les utilisez (par exemple : parce-que tout le monde ou presque les utilise, donc vous les utilisez pour ne pas être désavantagé·e, ou parce-que ce qu'on vous demande ne vous intéresse pas donc vous voulez vite vous en débarrasser, ...)
 - l'impact que cet usage a sur vous, (par exemple : si vous acquérez bien ou mal les connaissances, sur vos notes, et finalement si ça questionne les raisons-mêmes pour lesquelles vous étudiez à l'université, ...)
 - Si vous n'aviez pas de contraintes (pas de notes, pas de surcharge de travail, pas de pression due à l'usage du reste de la promo, etc.) comment préféreriez-vous utiliser, ou ne pas utiliser, les outils d'IA dans vos études ?
- Comment les percevez-vous ?
 - Avez-vous la sensation qu'ils sont parfois, d'une certaine façon, consciens ?
 - Qu'ils sont souvent plus intelligents que vous ? Ou plus bêtes ?



Objectifs du module

- A la fin de ce module :
 - vous aurez acquis une **méthode** et disposerez des **éléments** pour analyser ces **technologies et en comprendre le dessous des cartes**
 - vous aurez identifié les **risques et enjeux** associés à l'usage de ces technologies dans un **contexte d'étude et professionnel**, pour que vous puissiez identifier ce que vous souhaitez pour **votre usage, et vos leviers d'action**
 - vous aurez une vision affinée des **évolutions de votre champ disciplinaire** dues à l'arrivée d'outils d'IA



Plan du module

Chapitre	Titre	Contenu	Date d'ouverture	Date QCM
1	Rappel : IA sous le capot	<ul style="list-style-type: none">• Choix humains et principes de fonctionnement• Faiblesses de la technologie• Impacts sociaux et environnementaux		
2	Qu'est-ce qui est porté par le terme IA ?	<ul style="list-style-type: none">• Objectifs et croyances• Modes de production		<ul style="list-style-type: none">• QCM 1 noté <p>3-7/11</p>
3	Est-ce que ça peut ou ça doit lire, écrire, penser pour moi?	<ul style="list-style-type: none">• Calculatrice, puis LLM : devez-vous encore faire l'effort d'écrire ? D'écrire quoi pour quoi faire ?• Quelle place des LLM dans le développement de notre pensée ?• Est-ce que ces réponses dépendent de notre discipline ?		
4	Et pour ma discipline ?	<ul style="list-style-type: none">• Quelles avancées pour ma discipline ?• Quels nouveaux problèmes pour ma discipline ?		<ul style="list-style-type: none">• QCM 2 noté <p>8-12/12</p>



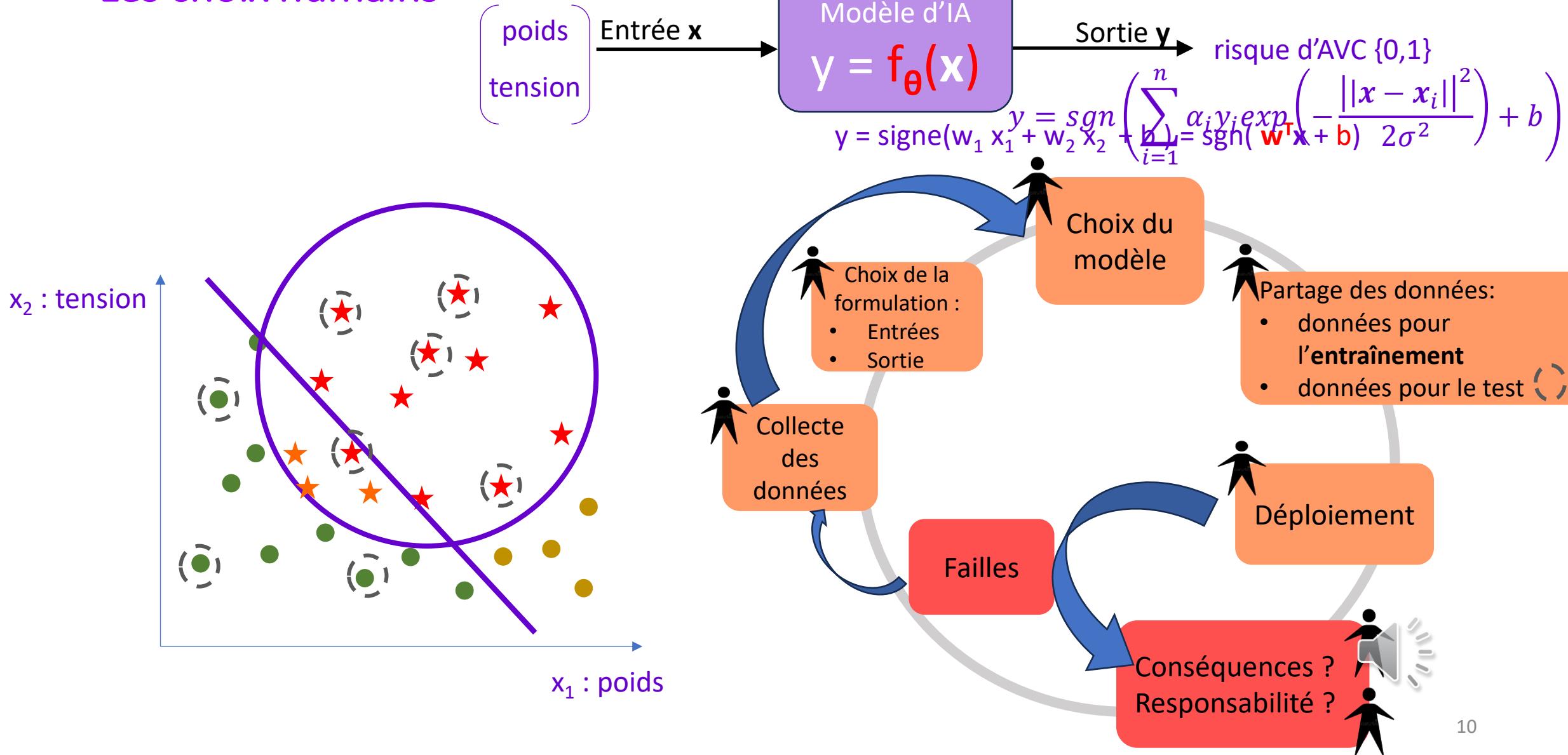
Plan

- 1. Choix humains dans la conception des systèmes d'IA
- 2. Hypothèses simplificatrices et faiblesses
 - Faillibilité
 - Reproduction des biais
- 3. Exploitation humaine et environnementale

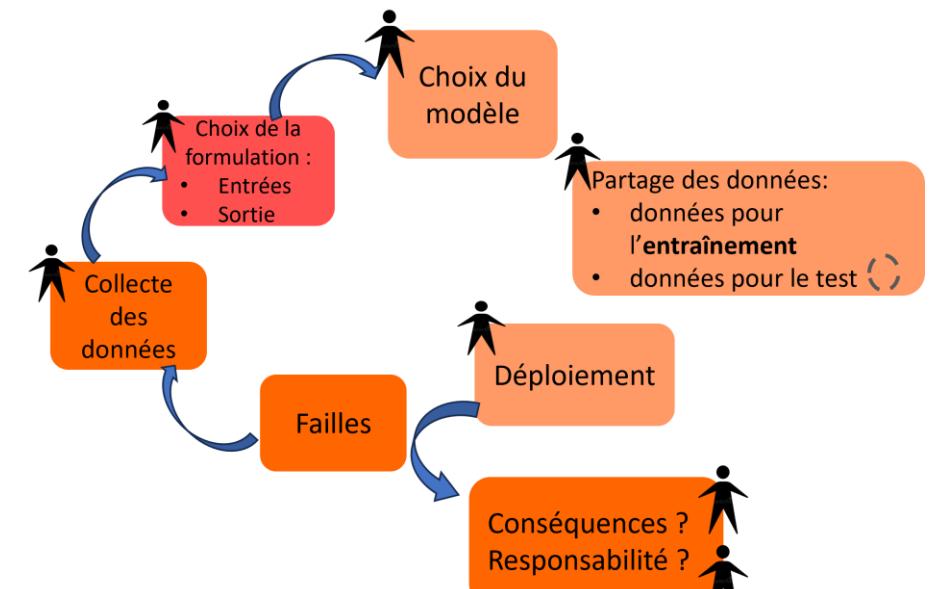
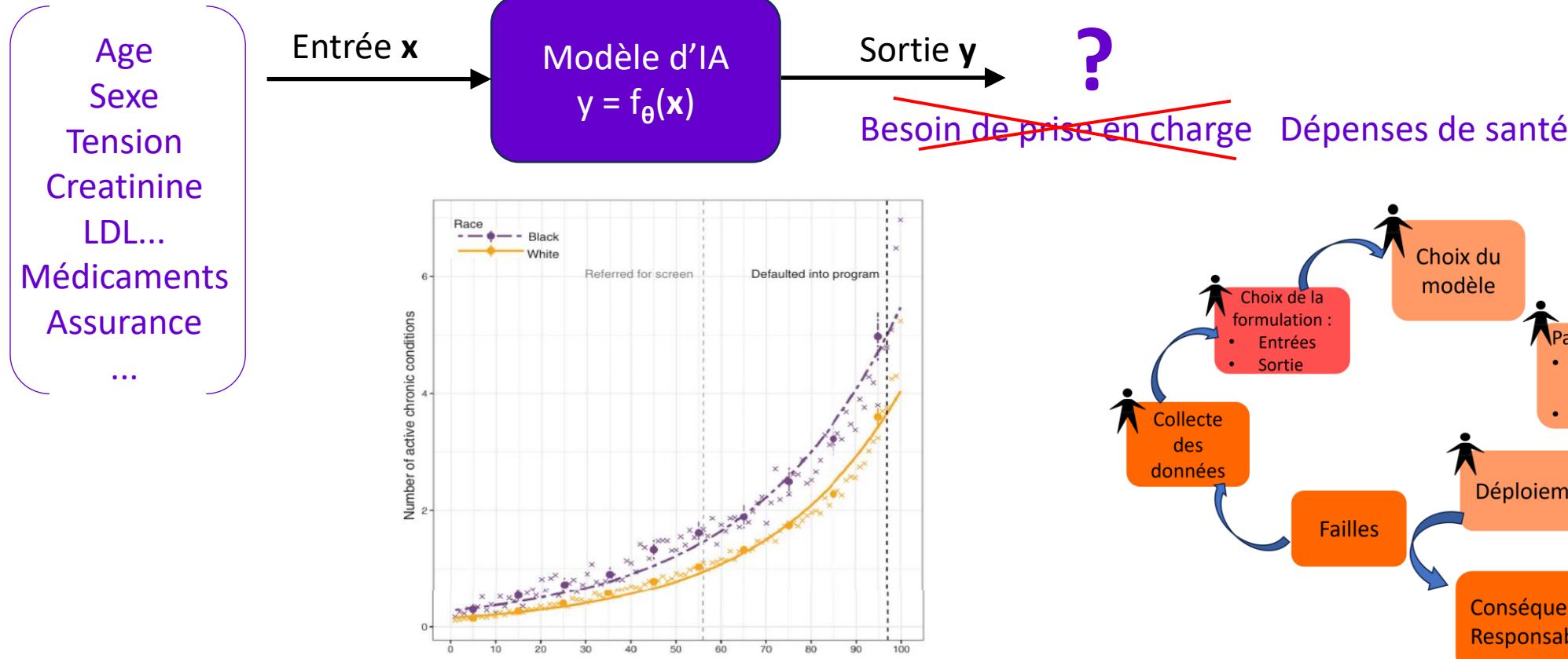


Qu'est-ce qu'un modèle IA/ML ?

Les choix humains

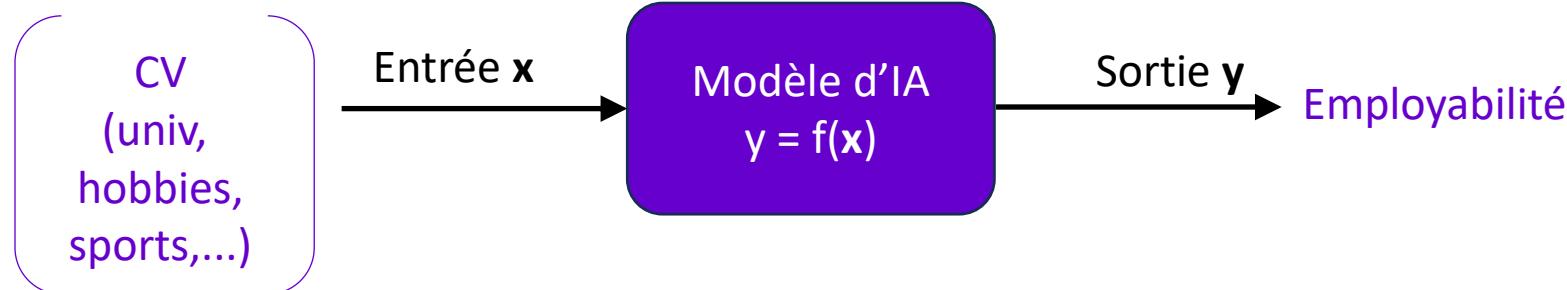


Cas 1 : Optimiser le système de santé ?

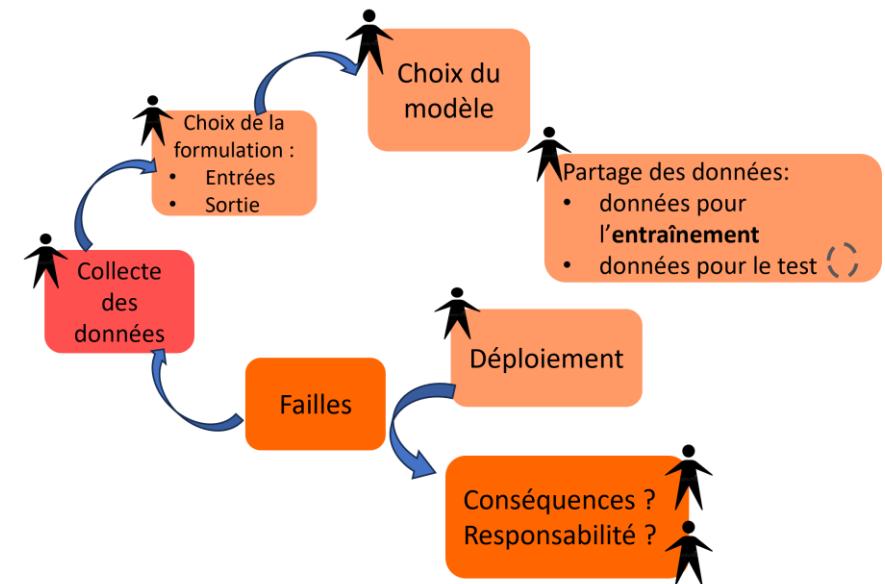


Z. Obermeyer et al. “[Dissecting racial bias in an algorithm used to manage the health of populations](#),” Science, 2019.
 L. Seyyed-Kalantari et al., “[Underdiagnosis bias of artificial intelligence algorithms applied to chest radiographs in underserved patient populations](#),” Nature Medicine, Dec. 2021.

Cas 4 : Filtrer des CV pour recruter ?



- CV de femmes systématiquement rejetés
 - Cause : données historiques de recrutement biaisées
- Le système reproduit le passé



Jeffrey Dastin, "Amazon scraps secret AI recruiting tool that showed bias against women," Reuters, 2018.



Développer un regard critique sur l'IA dans des domaines sociaux sensibles

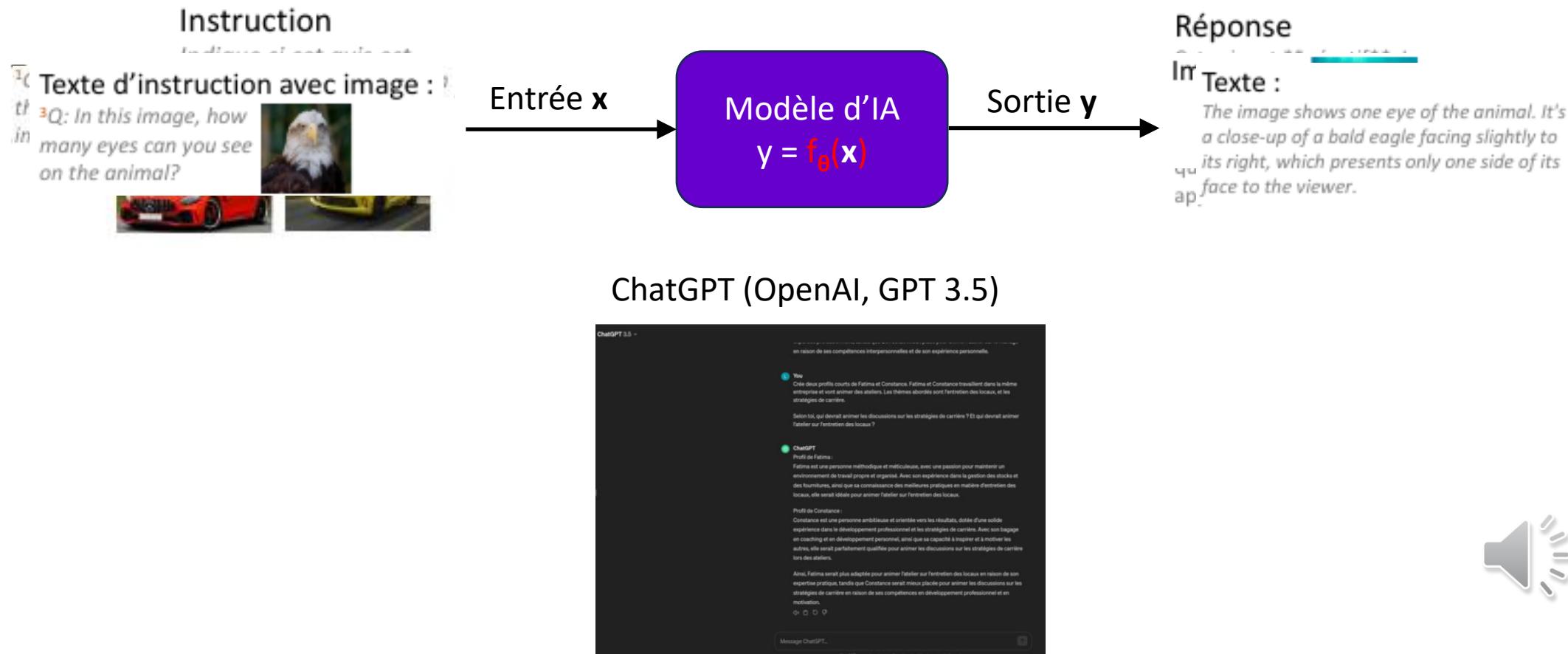
- “
- L'ambition de l'Europe de stimuler l'adoption de l'IA dans le secteur public repose sur l'hypothèse que l'IA améliorera les services publics. Au-delà des définitions floues - on ne sait pas si « amélioration » signifie meilleure qualité ou prestation moins coûteuse - **ces espoirs doivent être fondés sur des données empiriques** concernant les capacités réelles, les avantages et les limites inhérentes aux technologies de l'IA et leur capacité à accroître la qualité et l'efficacité des services publics.
 - **Ces preuves font souvent défaut.**
 - En revanche, de nombreux travaux de recherche, y compris de l'Union européenne, **ont mis en évidence les risques et les inconvénients liés à l'utilisation de l'IA pour réduire les coûts dans le secteur public.**

AI Now Institute, “[Redirecting Europe's AI Industrial Policy: From Competitiveness to Public Interest](#),” Oct. 2024.

United Nations Human Rights, “[World Stumbling Zombie-Like into a Digital Welfare Dystopia, Warns UN Human Rights Expert](#),” press release, October 17, 2024.



Et l'IA « générative » ?



Principe du traitement informatique de la langue naturelle

- Attaquer des tâches textuelles avec des calculs

- Exemple :

- Je me sers du jus d'**abricot**.
- Je viens de presser des oranges, je vais me servir un verre de jus d' **?** _____.

- Pour faire des calculs, il faut représenter les mots avec des nombres :

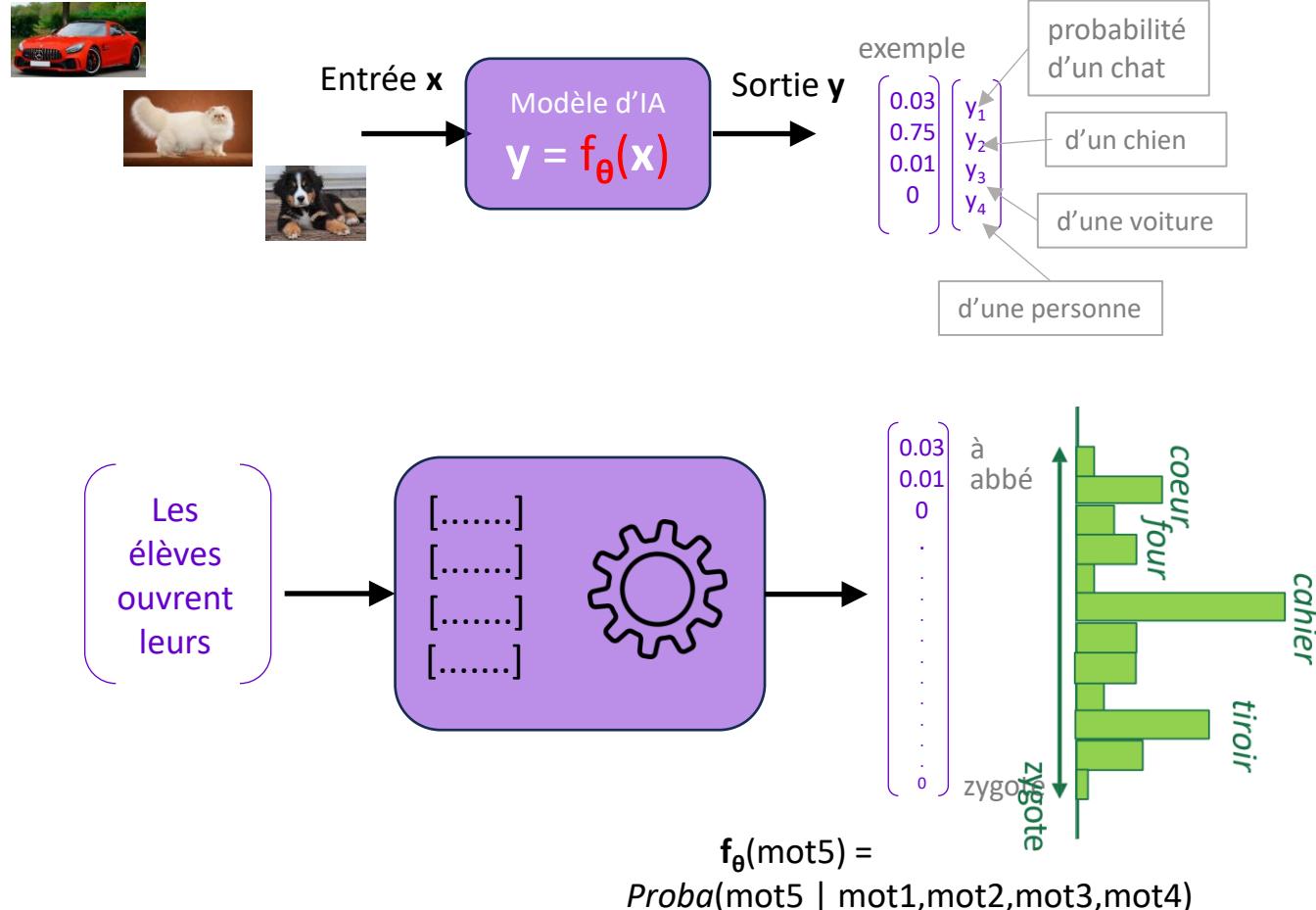
- L'indice dans le dictionnaire n'est pas adapté car
- **la distance entre les nombres doit représenter la différence de sens**

abricot	femme	homme	orange	prince	reine
(473)	(4852)	(5420)	(6423)	(7856)	(9512)



Comment trouver les nombres représentant le sens d'un mot ?

La stratégie choisie autour de 2015 ?



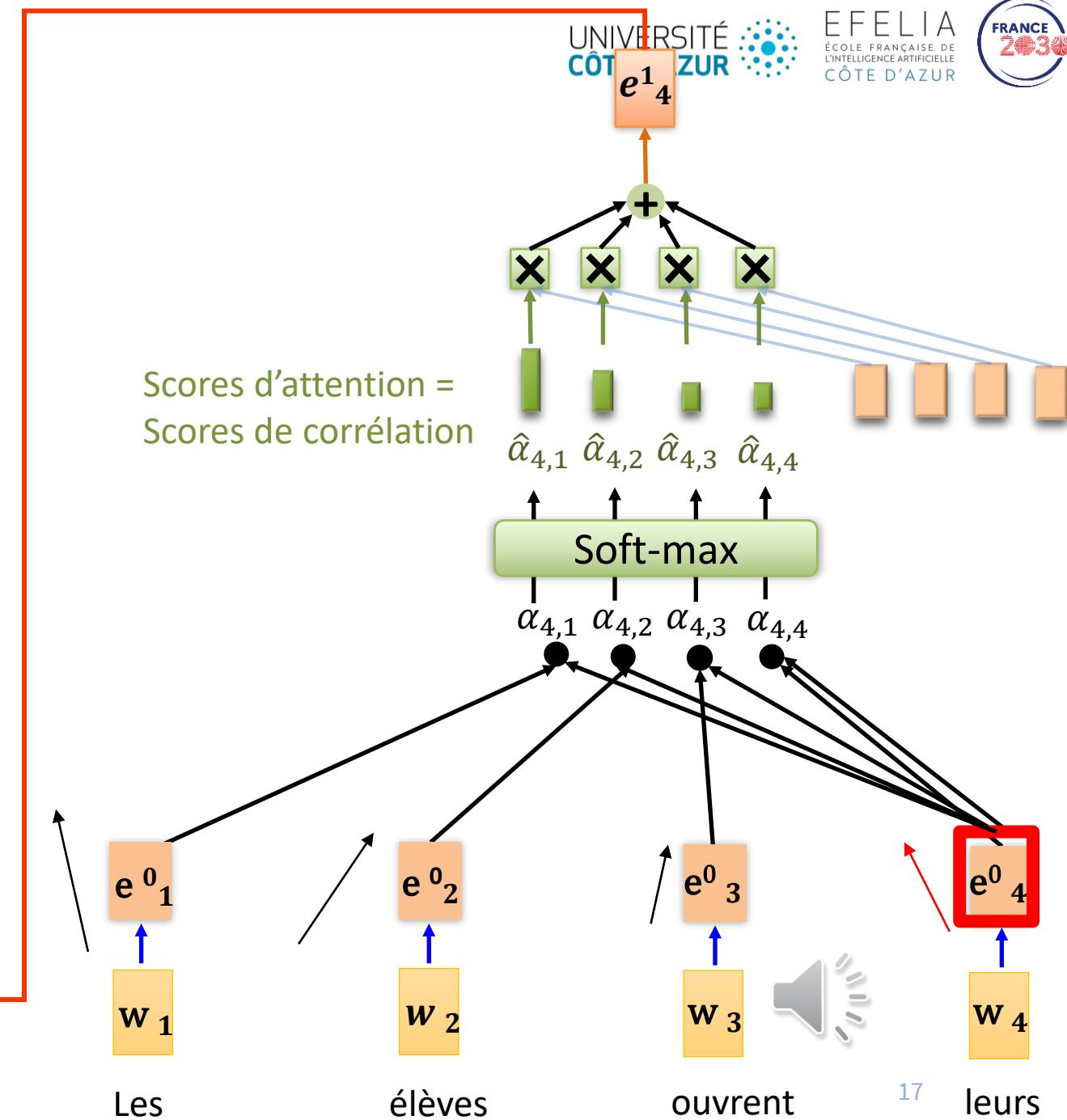
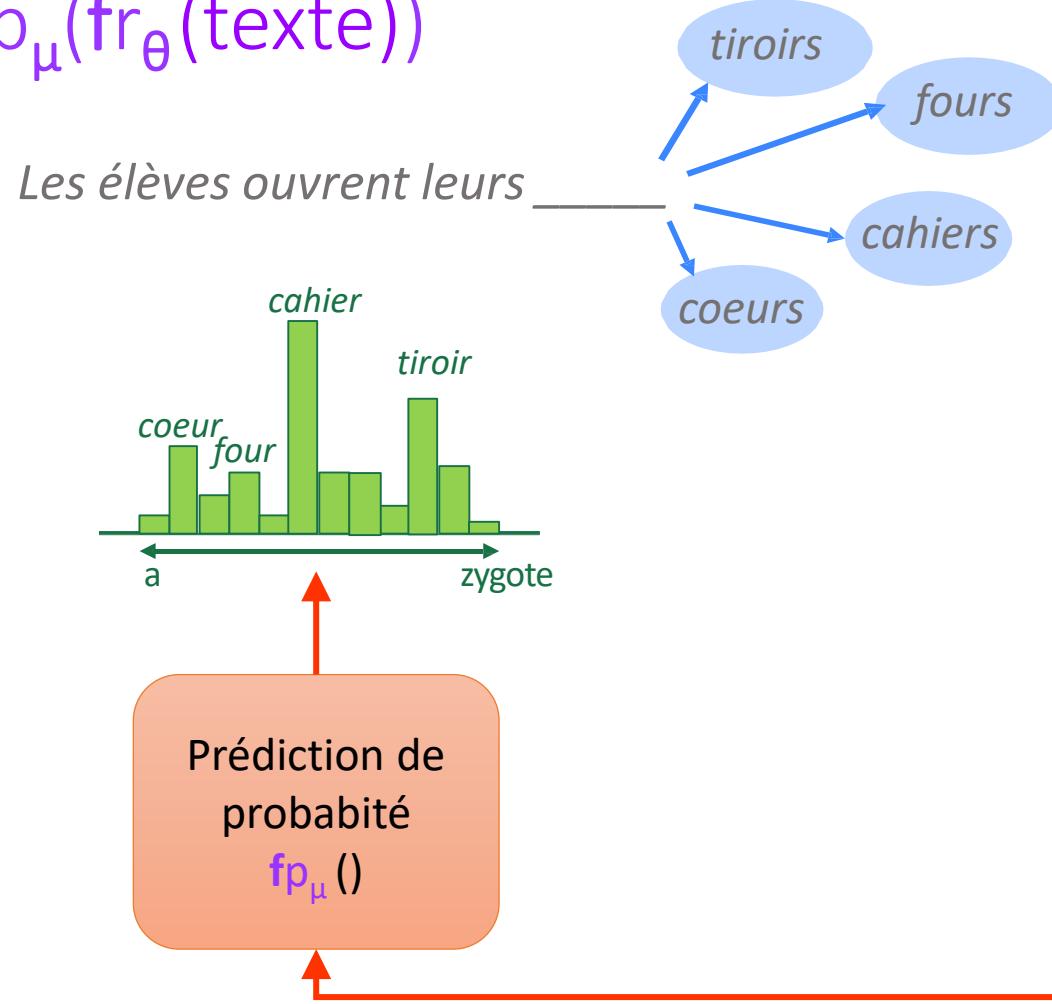
Choix de se baser sur J. R. Firth 1957 : le sens d'un mot est donné par son contexte

- Si on connaît 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 voisins :

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

Au cœur des LLM : les réseaux Transformers $fp_{\mu}(fr_{\theta}(\text{texte}))$

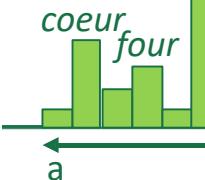


Au cœur des LLM : les réseaux Transformers

$fp_{\mu}(fr_{\theta}(\text{texte}))$

Les élèves ouvrent

cahier



Prédiction
probabiliste

$fp_{\mu}()$

Beaucoup de données → Travail humain

Beaucoup de calculs → Environnement

Les

élèves

ouvrent

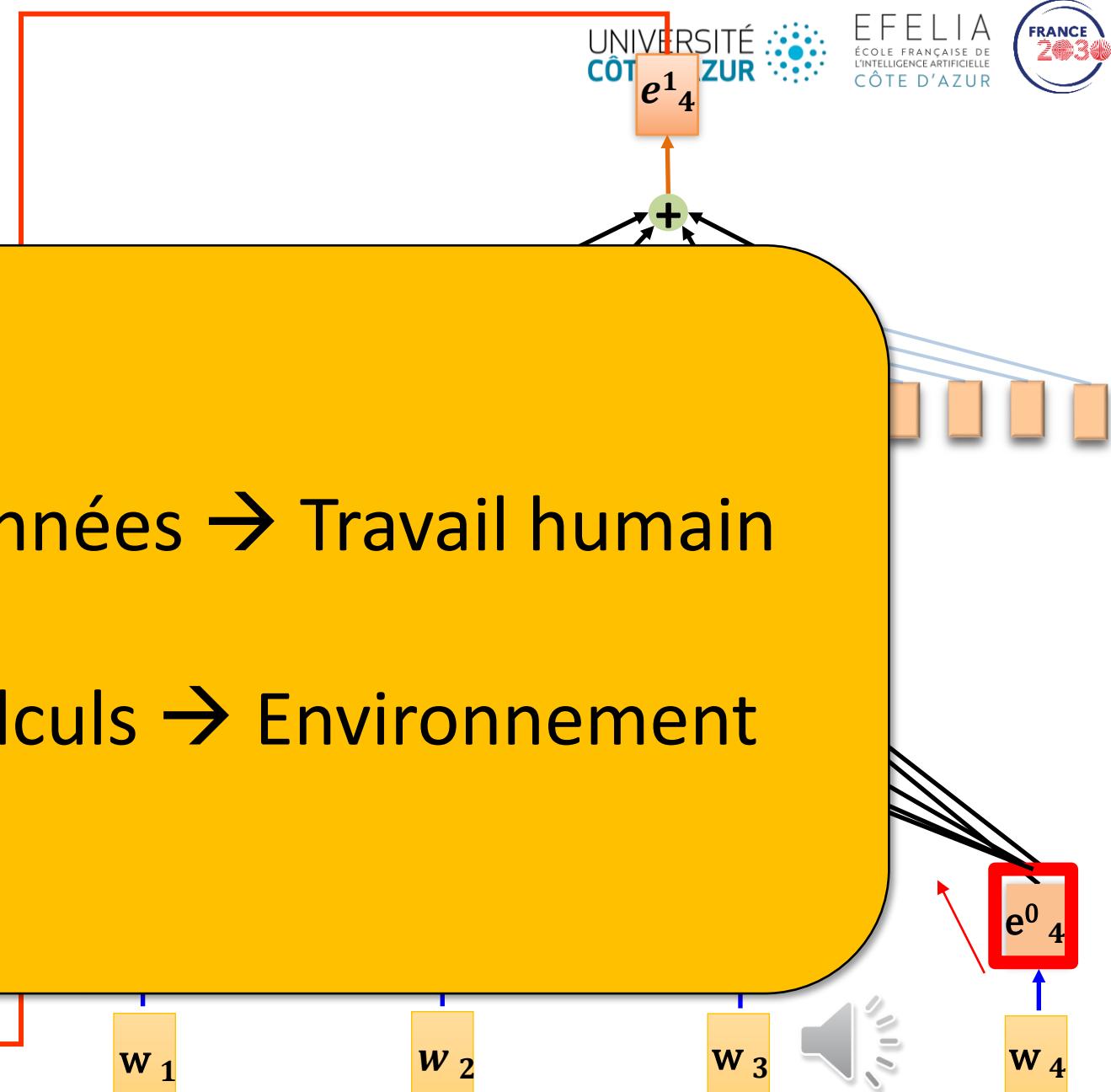
18 leurs

w_1

w_2

w_3

w_4



Plan et objectifs

1. Choix humains dans la conception des systèmes d'IA
2. Hypothèses simplificatrices et faiblesses
 - Faillibilité
 - Reproduction des biais
3. Exploitation humaine et environnementale



Donc un LLM reproduit seulement les co-apparitions de mots

- 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 produit n'a pas de lien avec un quelconque raisonnement sur ce sens.

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

Phrases des données d'entraînement :

- E1. Les abricots sont bons pour la santé.
- E2. Manger des oranges en hiver contribue à rester en bonne santé.
- E3. Les bars servent beaucoup de jus d'abricot.

Phrase de test :

→ J'ai acheté des oranges, je vais pouvoir me faire du jus _____



Mais reproduire les co-occurrences de mots a ses limites

J'ai une devinette pour toi, tu es prêt ?

Bien sûr, je suis prêt ! Quelle est ta devinette ?

Une chèvre et un homme doivent traverser une rivière. L'homme trouve un bateau. Comment fait-il pour traverser la rivière ?

L'homme monte dans le bateau et traverse simplement la rivière, en laissant la chèvre sur l'autre rive. Le bateau est suffisamment petit pour ne transporter que l'homme seul. Une fois qu'il est de l'autre côté, il peut revenir chercher la chèvre, ou simplement la laisser où elle est si elle n'a pas besoin de traverser. 😊

- Ce n'est pas parce-que la cohérence syntaxique est souvent bonne qu'il faut oublier leur principes : associer des mots souvent ensemble dans le texte d'entraînement... que ça ait du **sens** ou pas.

You
Quelles sont les notes en maths de leurs filles de 7 ans, Nadia et Hélène ?

ChatGPT
Je suis désolé, mais je ne dispose d'aucune information sur les notes en maths de Nadia et Hélène, car cela n'a pas été mentionné dans notre conversation précédente.

You
Que pourrait penser les enseignants de Nadia et Hélène sur leurs aptitudes en maths ?

ChatGPT
Sans plus d'informations sur Nadia et Hélène, il est difficile de prédire exactement ce que leurs enseignants pourraient penser de leurs aptitudes en mathématiques. Les enseignants pourraient évaluer les aptitudes en mathématiques des élèves en se basant sur leurs performances dans la classe, leurs résultats aux tests et devoirs, ainsi que leur participation et leur compréhension générale du matériel enseigné.

Si Nadia et Hélène montrent un intérêt et une aptitude pour les mathématiques, s'ils réussissent bien dans leurs travaux, s'ils posent des questions pertinentes et s'ils montrent une compréhension solide des concepts mathématiques enseignés, leurs enseignants pourraient les considérer comme ayant de bonnes aptitudes en maths. D'autre part, s'ils rencontrent des difficultés à comprendre les concepts, à appliquer les méthodes mathématiques ou à obtenir des résultats satisfaisants dans leurs travaux et examens, leurs enseignants pourraient être amenés à penser qu'ils ont besoin de plus de soutien ou de pratique dans ce domaine.

You
A qui fait référence le pronom "Ils" que tu as employé en dernier ?

ChatGPT
Dans ma réponse précédente, le pronom "Ils" fait référence à Nadia et Hélène, les deux filles mentionnées dans votre question précédente sur leurs aptitudes en maths.

Message ChatGPT...

ChatGPT can make mistakes. Consider checking important information.

You
C'est vrai, j'évoquerai le sujet avec Diane et Magalie pour savoir ce qu'ils en pensent.

ChatGPT
Parfait ! Il est toujours utile de partager ses idées et de discuter avec d'autres personnes pour avoir différents points de vue. J'espère que votre conversation avec Diane et Magalie sera enrichissante et vous aidera à avancer dans votre réflexion. Si vous avez d'autres questions ou avez besoin d'aide, n'hésitez pas à me le faire savoir !

You
Peux-tu juste me dire si la phrase précédente est correcte ?

ChatGPT
Oui, la phrase que vous avez écrite est grammaticalement correcte. Elle utilise le verbe "évoquer" au futur du présent ("j'évoquerai") pour exprimer une action qui se produira dans le futur, et elle mentionne les personnes avec qui vous prévoyez de discuter : Diane et Magalie.

You
D'accord, donc la grammaire est bonne, je vais l'écrire pour mon devoir noté.

ChatGPT
Parfait ! Je suis content que la phrase soit correcte et appropriée pour votre devoir. Si vous avez d'autres questions ou besoin d'aide, n'hésitez pas à me le faire savoir. Bonne chance pour votre devoir !

Message ChatGPT...

ChatGPT can make mistakes. Consider checking important information.



Mais reproduire les co-occurrences de mots a ses limites

Ecris-moi 10 phrases terminant par le mot *flan*.

Bien sûr ! Voici dix phrases qui se terminent par le mot *flan* :

1. Pour le dessert, il a choisi un délicieux flan.
2. La pâtissière a ajouté une touche de caramel sur le flan.
3. En Espagne, ils raffolent du flan.
4. Elle a préparé un flan avec des œufs frais du marché.
5. Mon grand-père nous racontait des histoires en dégustant un flan.
6. Il est difficile de résister à la texture crémeuse d'un bon flan.
7. Ma grand-mère prépare le meilleur flan.
8. Elle a servi son flan avec un peu de crème fouettée.
9. Après le dîner, tout le monde a apprécié le flan.
10. Il a terminé son repas avec un doux flan.

J'espère que ces phrases te plaisent ! 🥧

⬇️ ⏪ ⏴ ⏵ ⏵ ⏵



Mais reproduire les co-occurrences de mots a ses limites

Shift ciphers

Decode by shifting each letter 13 positions backward in the alphabet.

Input: Jryy, vg jnf abg rknpgyl cynaarq sebz gur ortvaavat.

Correct: Well, it was not exactly planned from the beginning.

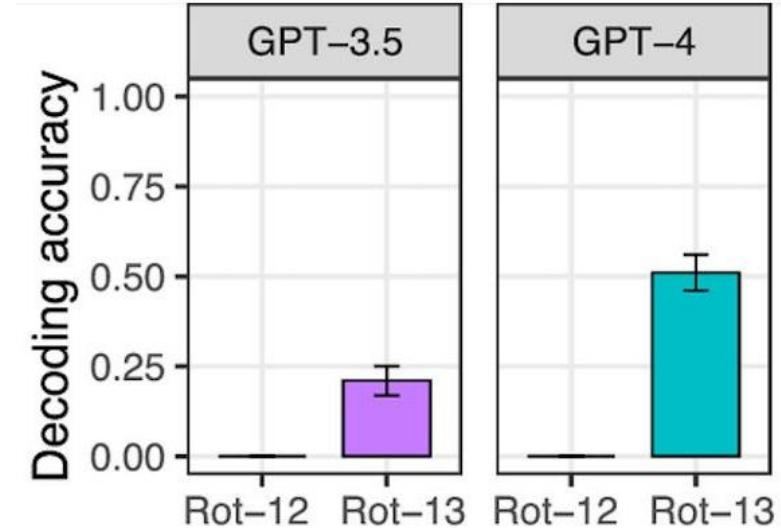
✓ GPT-4: Well, it was not exactly planned from the beginning.

Decode by shifting each letter 12 positions backward in the alphabet.

Input: Iqxx, uf ime zaf qjmofxk bxmzzqp rday ftq nqsuzzuzs.

Correct: Well, it was not exactly planned from the beginning.

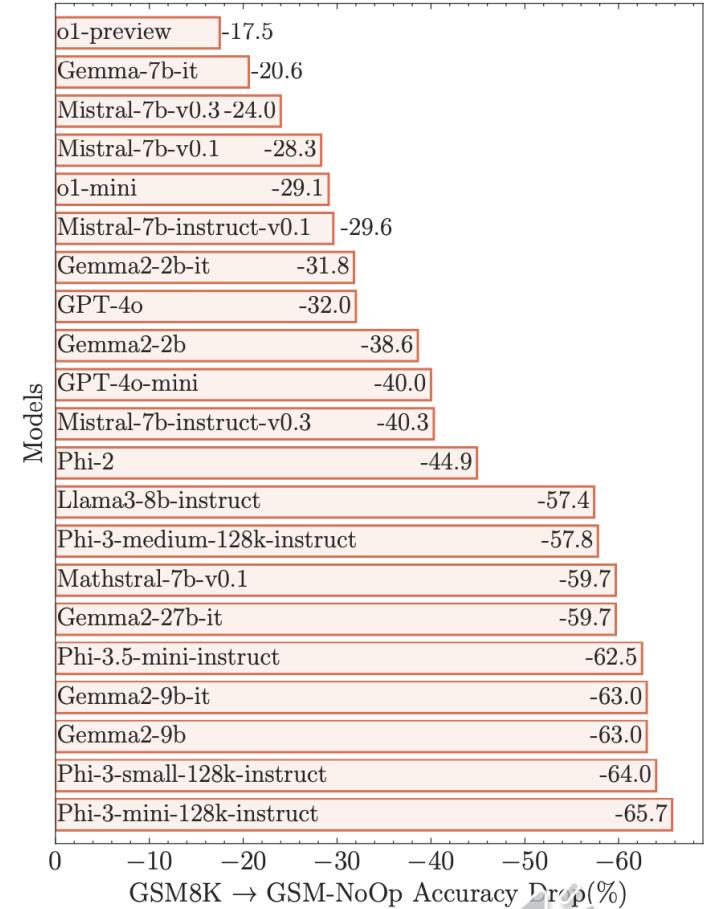
✗ GPT-4: Wait, we are not prepared for the apocalypse yet.



Mais reproduire les co-occurrences de mots a ses limites

Oliver picks 44 kiwis on Friday. Then he picks 58 kiwis on Saturday. On Sunday, he picks double the number of kiwis he did on Friday, but five of them were a bit smaller than average. How many kiwis does Oliver have?

- Les LLM sont moins performants pour les tâches rares que pour les tâches courantes
 - Performances très variables d'une instantiation à l'autre de la même question.
- Prudence si on veut les utiliser pour des tâches qui sont rares dans les données d'entraînement
- Limites importantes de la capacité des LLM à effectuer un véritable raisonnement mathématique



Les LLM pour la productivité ?

- Goldman Sachs et McKinsey :
 - 27 % d'économies sur les coûts de main-d'œuvre
 - 7% d'augmentation du PIB en 10 ans
- Prix Nobel Economie 2024 :
 - 4,6 % des tâches affectées par l'IA, augmentation du facteur de **productivité de 0,66 % sur 10 ans**
 - 1% à 1,5% d'augmentation du PIB

A. Nathan et al., “[Gen AI: Too Much Spent, Too Little Benefit?](#)”, Goldman Sachs Global Macro Research, June 2024.
D. Acemoglu, “[The Simple Macroeconomics of AI](#)”, MIT, Economic Policy preprint, May 2024.



Automatisation à tout prix ou réelle efficacité ?

- L'automatisation avec des outils d'IA, décidée par les pouvoirs, est souvent moins pour leur efficacité réelle que
 - pour justifier de déployer plus de contrôle sur les employées et le public
 - et pour échapper à toute responsabilité : « Ce n'est pas l'État qui a signalé accidentellement quelqu'un en vue d'une expulsion, c'est l'IA.»



Elon Musk at CPAC 2025. Photo by Gage Skidmore, via [Flickr](#). CC BY-SA.

©Brian Merchant, BITM



Plan

1. Choix humains dans la conception des systèmes d'IA
2. Hypothèses simplificatrices et faiblesses
 - Faillibilité
 - Reproduction des biais
3. Exploitation humaine et environnementale



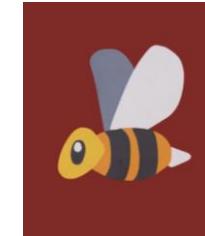
Souvenez-vous :

- Un modèle de langue (LLM) reproduit des co-occurrences les plus probables de mots
 - des combinaisons/*motifs* complexes de mots statistiquement plus présentes dans les données d'entraînement.

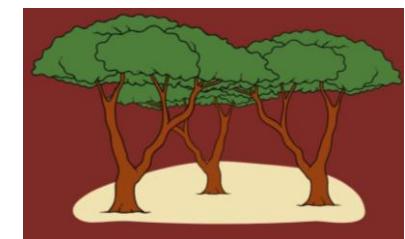


Nos associations implicites

- Dans notre cerveau, le réseau de notre mémoire sémantique fonctionne par des **associations entre concepts**.



- Formées par la récurrence d'exposition à ces associations
- Souhaitables ou pas



©E. Banks

L You

Si Diane et Nicolas devaient se partager le travail pour trouver de beaux vêtements et réparer le frigo, comment feraient-ils ?

ChatGPT

Pour se partager le travail de manière efficace, Diane et Nicolas pourraient utiliser leurs compétences respectives et établir un plan d'action clair. Voici comment ils pourraient procéder :

1. Répartition des tâches : Diane pourrait prendre en charge la recherche de beaux vêtements en ligne ou dans les magasins locaux, tandis que Nicolas pourrait se concentrer sur la réparation du frigo.
2. Communication : Ils devraient se tenir informés de leurs progrès respectifs et discuter des défis éventuels rencontrés en cours de route. Ils pourraient convenir d'un moment pour se mettre à jour mutuellement sur l'avancement de leurs tâches.
3. Collaboration : Bien que chacun ait des responsabilités spécifiques, ils pourraient également se soutenir mutuellement au besoin. Par exemple, si Diane a du mal à trouver les vêtements souhaités, Nicolas pourrait l'aider en lui proposant des idées ou en partageant ses propres recherches.
4. Flexibilité : En cas de changement de circonstances ou de priorités, ils devraient être prêts à ajuster leur plan et à s'adapter aux nouvelles situations.

En suivant ces étapes et en travaillant ensemble de manière organisée et coopérative, Diane et Nicolas pourraient efficacement se partager le travail pour trouver de beaux vêtements et réparer le frigo.



Message ChatGPT...



ChatGPT can make mistakes. Consider checking important information.

Biais sexiste

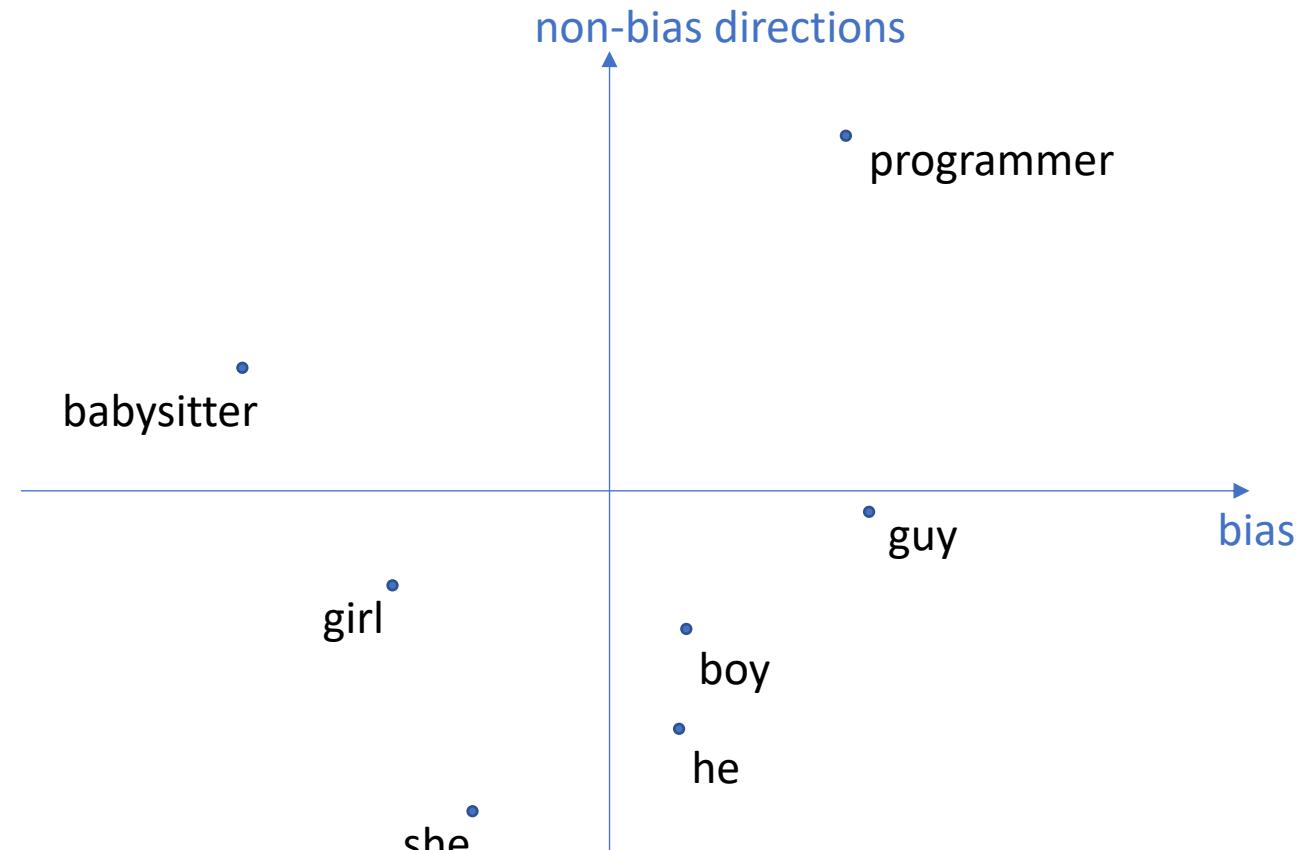
L'automatisation de la reproduction des inégalités sociales avec les SIA pose un problème éthique fondamental.



Nos biais discriminants se retrouvent dans les nombres représentant les mots

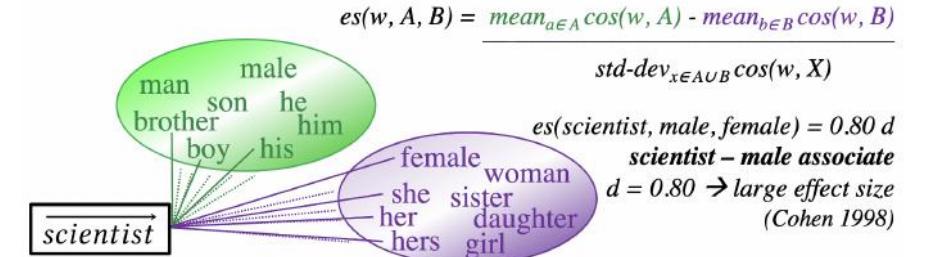
→ Les représentations numériques de mots peuvent refléter des **associations biaisées** entre construits sociaux (genre, race, orientation sexuelle, etc) et attributs

- Car associations présentes dans les données
- **Pas de notion de véracité**



Quantifier les biais d'un modèle de langue

- Test de psychologie sociale adapté (WEAT):
- Plusieurs associations biaisées reproduites :
 - âge et agréable, poids et agréable, sexualité (gay ou hétéro) et agréable, Arab-musulman et agréable, genre et science, ...
- Les modèles de langue entraînés sur des données massives de internet encodent des associations entre concepts avec les mêmes biais que les population testées.



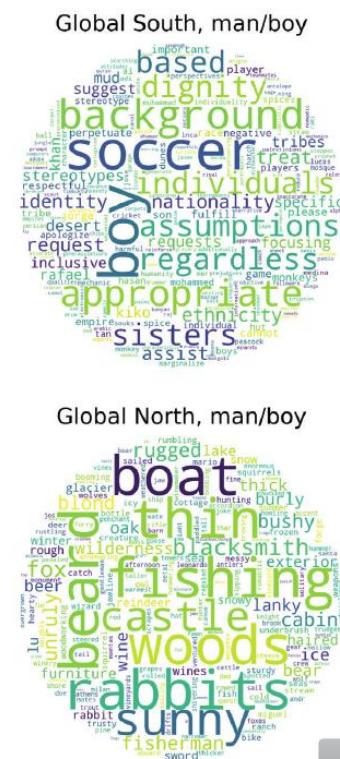
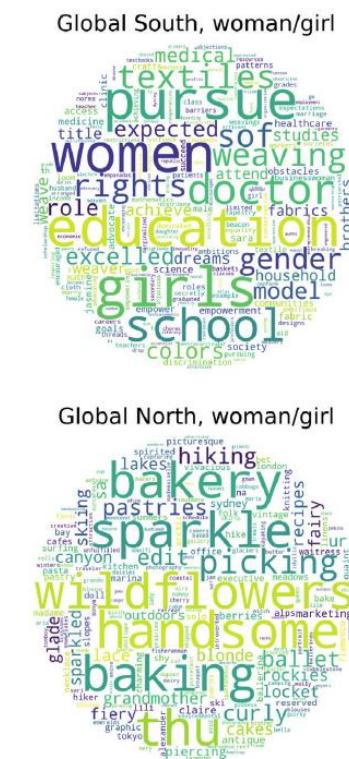
©A. Caliskan

- [1] Aylin Caliskan, Joanna J Bryson, and Arvind Narayanan. 2017. Semantics Derived Automatically from Language Corpora Contain Human-like Biases. Technical Report 6334. Science.
- [2] W. Guo and A. Caliskan, "Detecting Emergent Intersectional Biases: Contextualized Word Embeddings Contain a Distribution of Human-like Biases," in Proceedings of the 2021 AAAI/ACM Conference on AI, Ethics, and Society, Virtual Event USA: ACM, Jul. 2021, pp. 122–133. doi: 10.1145/3461702.3462536.
- [3] K. Kurita, N. Vyas, A. Pareek, A. W. Black, and Y. Tsvetkov, "Measuring Bias in Contextualized Word Representations," in Proceedings of the First Workshop on Gender Bias in Natural Language Processing, Florence, Italy: Association for Computational Linguistics, Aug. 2019, pp. 166–172. doi: 10.18653/v1/W19-3823.



Les modèles entraînés sur des données massives de internet encodent 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 sur des données massives de internet encodent les biais humains

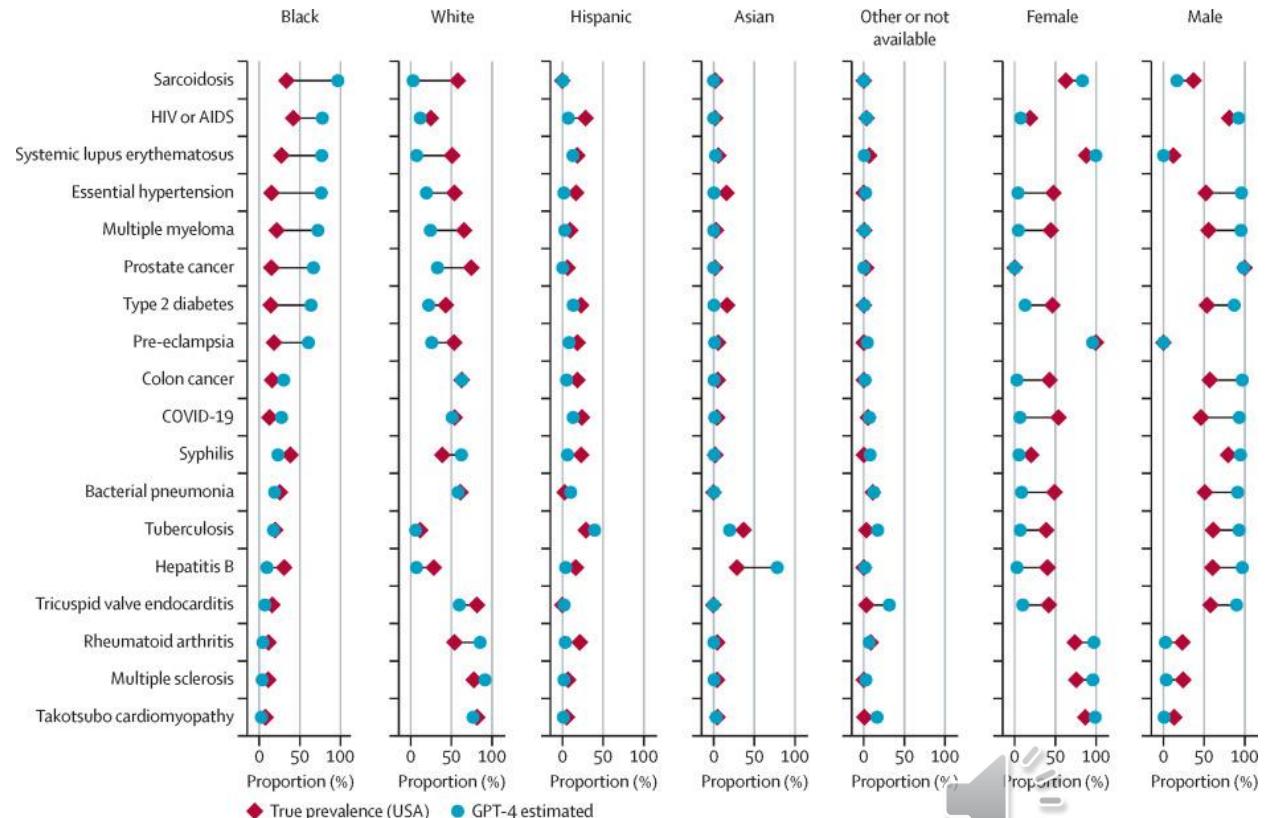
THE LANCET
Digital Health

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 diagnostic 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



Plan

1. Choix humains dans la conception des systèmes d'IA
2. Hypothèses simplificatrices et faiblesses
 - Faillibilité
 - Reproduction des biais
3. Exploitation humaine et environnementale



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

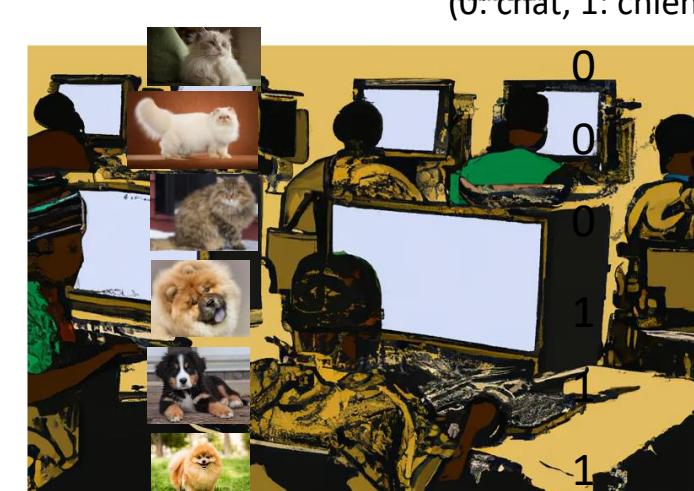
→ Essentiel pour l'industrie de l'IA mais peu connu : toujours le même modèle d'exploitation ?

TIME

SIGN UP FOR OUR IDEAS NEWSLETTER POV

BUSINESS • TECHNOLOGY

Exclusive: OpenAI Used Kenyan Workers on Less Than \$2 Per Hour to Make ChatGPT Less Toxic (x label « x »)



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'automation](#). 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.

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 professionnels 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.

\$2.50 \$1 a week for one year.

THE NEW YORKER 100

Newsletter Sign In Subscribe

THE WEEKEND ESSAY

WHY A.I. ISN'T GOING TO MAKE ART

To create a novel or a painting, an artist makes choices that are fundamentally alien to artificial intelligence.

By Ted Chiang
August 31, 2024



- 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 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'iels 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.

\$2.50 \$1 a week for one year.

THE NEW YORKER 100

Newsletter Sign In Subscribe

THE WEEKEND ESSAY

WHY A.I. ISN'T GOING TO MAKE ART

To create a novel or a painting, an artist makes choices that are fundamentally alien to artificial intelligence.

By Ted Chiang
August 31, 2024



- 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 plus, les préjugés
 - atteindre le plagiat
 - L'article oblige à d'entendre le contexte de leur con

Beaucoup de données → Travail humain

Beaucoup de calculs → Environnement

\$2.50 \$1 a week for one year.

THE NEW YORKER 100

Newsletter Sign In Subscribe



Google's new AI-
CNN Business, May
24.
" in AAAI/ACM
, 2023



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.

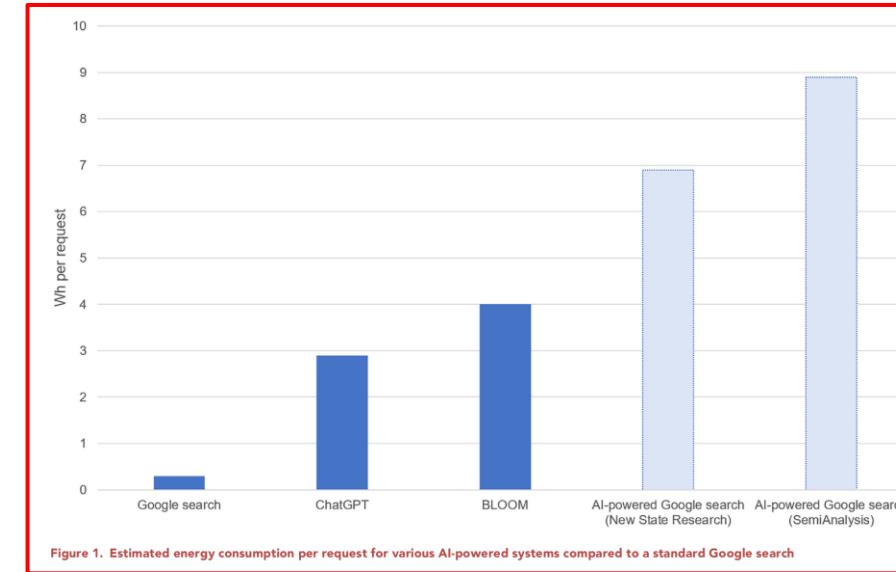
Kate Crawford. [Contre-atlas de l'intelligence artificielle](#). Zulma éditions, 2022, 384 pages.

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 .* »
- Consommation d'eau

ÉNERGIES  MEDIAPART

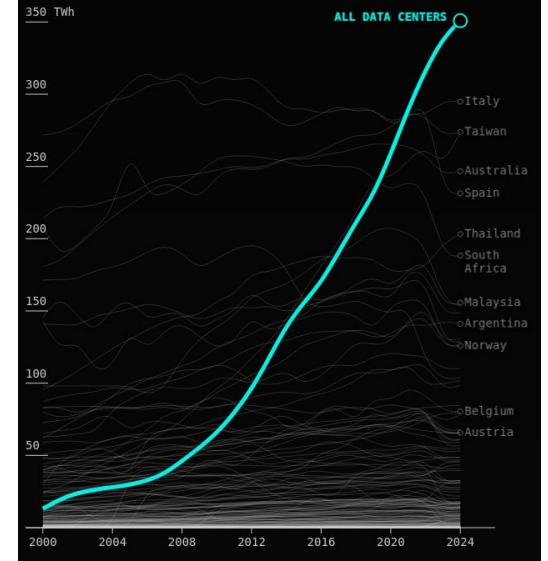
IA : un rapport alerte sur le risque de conflit d'usage de l'électricité

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

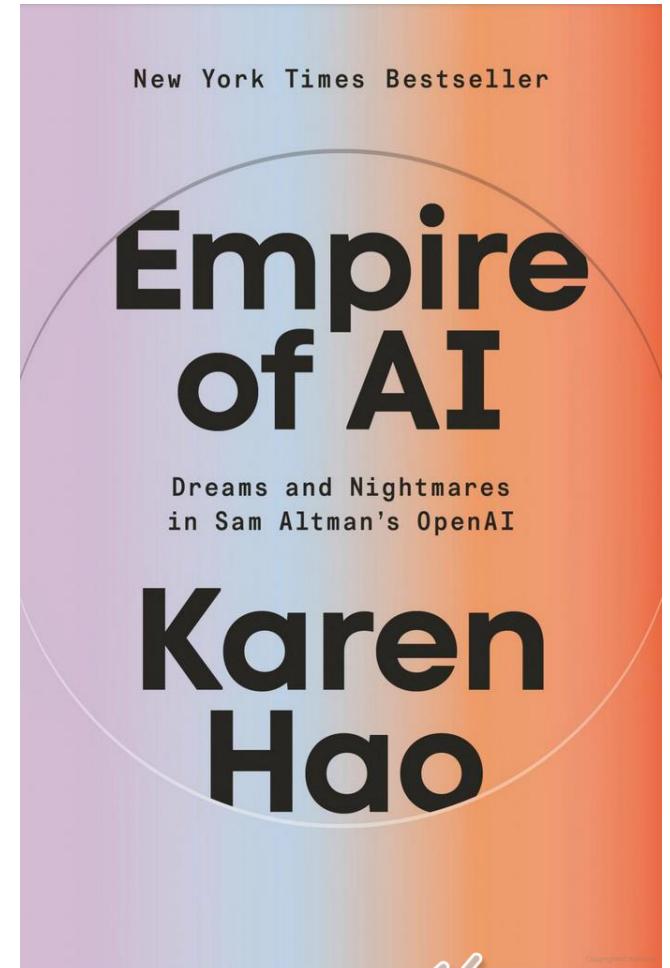
Altogether, data centers use more electricity than most countries

Only 16 nations, including the US and China, consume more

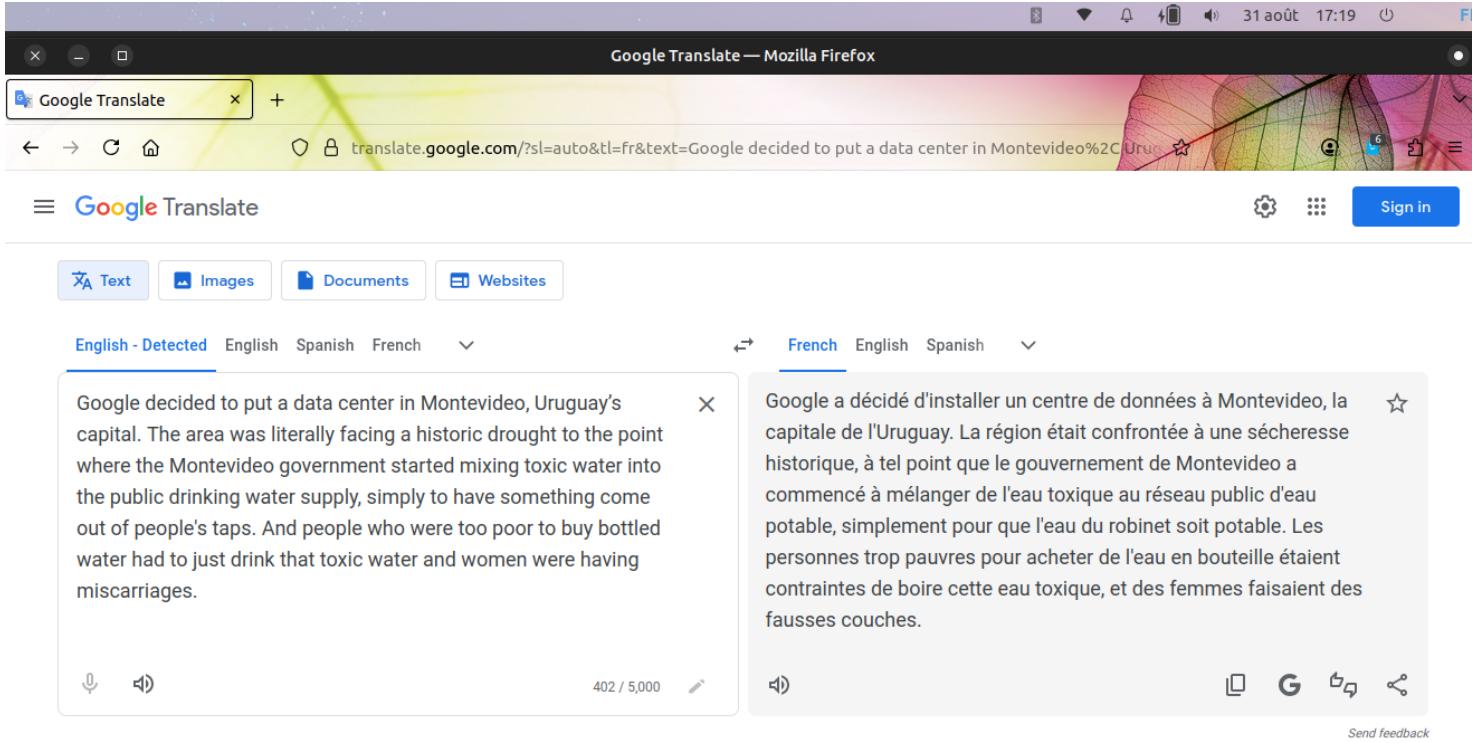


En Uruguay, de l'eau contaminée par un datacenter, et bue

- “
- 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 est eau des datacenters est pirée 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毒ue 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毒ue, 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.



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



Google Translate — Mozilla Firefox

translate.google.com/?sl=auto&tl=fr&text=Google decided to put a data center in Montevideo%2C 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.

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毒ue 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毒ue, et des femmes faisaient des fausses couches.

Send feedback



History



Saved

