

L'"IA" c'est pas ce que tu crois !

Karën Fort

karen.fort@loria.fr / <https://members.loria.fr/KFort>

Conférences de culture numérique – licence humanités, Caen – 1er octobre 2026

D'où je parle ?

- ▶ professeure en informatique à l'Université de Lorraine
- ▶ spécialisée en traitement automatique des langues (TAL)

Question !



Allez sur wooclap.com

Entrez le code d'événement dans le bandeau supérieur

Code d'événement

ZTRCXS



En attente de participants ...

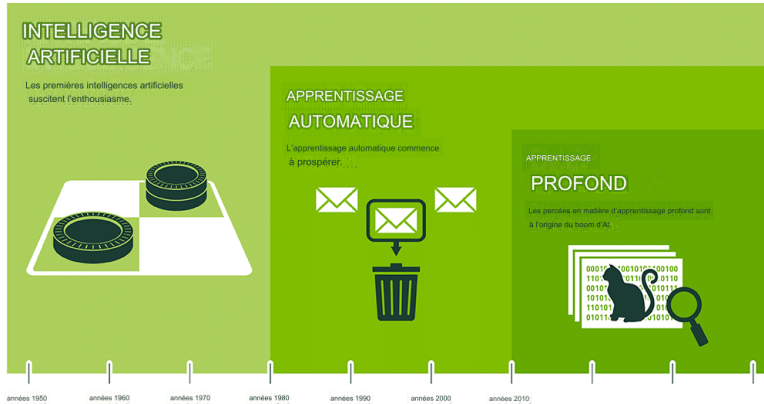
 Copier le lien de participation

D'où je parle

L'"IA" et les grands modèles de langues (LLM)

Les vrais enjeux (aujourd'hui)

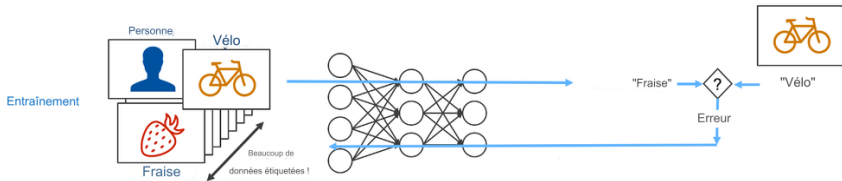
L'"IA", l'apprentissage et l'apprentissage profond



Depuis un début d'optimisme dans les années 1950, de plus petits sous-ensembles de l'intelligence artificielle, d'abord l'apprentissage automatique, puis l'apprentissage profond, un sous-ensemble de l'apprentissage automatique, ont créé des perturbations toujours plus importantes.

L'apprentissage profond en une diapo (!)

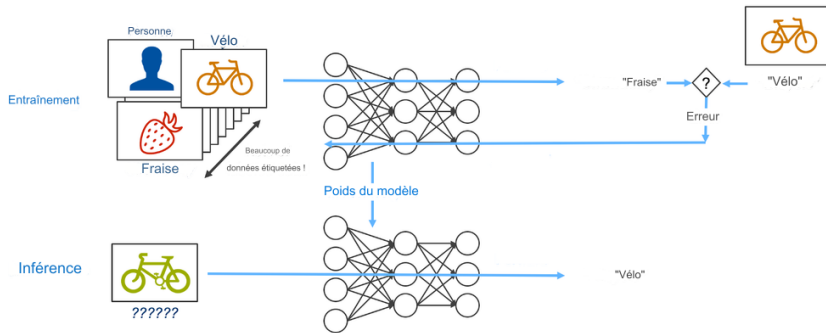
Ici, on cherche à associer une étiquette (mot) à une image



<https://community.intel.com/t5/Blogs/Tech-Innovation/Artificial-Intelligence-AI/The-Difference-Between-Deep-Learning-Training-and-Inference/post/1335634>

L'apprentissage profond en une diapo (!)

Ici, on cherche à associer une étiquette (mot) à une image



<https://community.intel.com/t5/Blogs/Tech-Innovation/Artificial-Intelligence-AI/The-Difference-Between-Deep-Learning-Training-and-Inference/post/1335634>



Allez sur wooclap.com

Entrez le code d'événement dans le bandeau supérieur

Code d'événement

ZTRCXS



En attente de participants ...

 Copier le lien de participation

Les grands modèles de langues (LLM) : apprentissage profond pour le texte

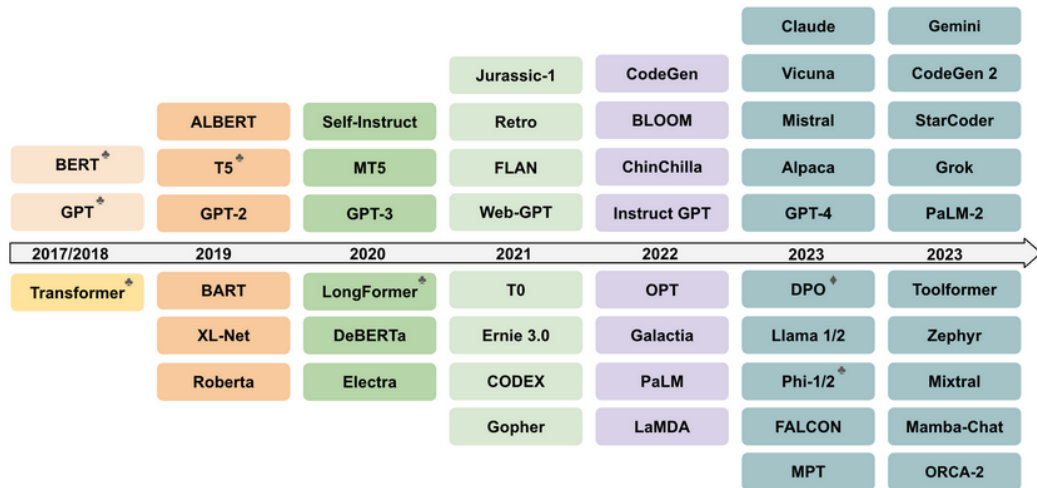
Très récent :

- ▶ premier *transformer* (BERT) en 2018
- ▶ ChatGPT sort fin 2022

Des acteurs bien connus, entre autres :



L'explosion des LLM



<https://arxiv.org/html/2402.06196v1>



Allez sur wooclap.com

Entrez le code d'événement dans le bandeau supérieur

Code d'événement

ZTRCXS

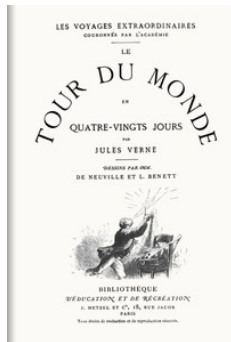


En attente de participants ...

 Copier le lien de participation

Corpus d'entraînement

Llama : 10 000 milliard de mots, soit 100 000 000 fois *Le tour du monde en 80 jours*



Les grands modèles de langues : les "couteaux suisses" du TAL ?



- ▶ faire la conversation, générer des textes, du code
- ▶ catégoriser le "sentiment" d'un texte (positif, négatif, neutre)
- ▶ réaliser une analyse linguistique d'un texte (verbe, sujet, COD, etc)
- ▶ etc

Les grands modèles de langues : les "couteaux suisses" du TAL ?



- ▶ faire la conversation, générer des textes, du code
- ▶ catégoriser le "sentiment" d'un texte (positif, négatif, neutre)
- ▶ réaliser une analyse linguistique d'un texte (verbe, sujet, COD, etc)
- ▶ etc



Allez sur wooclap.com

Entrez le code d'événement dans le bandeau supérieur

Code d'événement

ZTRCXS



En attente de participants ...

 Copier le lien de participation

Les enjeux (chiffon rouge)



<https://ideas.ted.com/how-the-gains-we-make-in-ai-could-ultimately-destroy-us/>

D'où je parle

L'"IA" et les grands modèles de langues (LLM)

Les vrais enjeux (aujourd'hui)



Allez sur wooclap.com

Entrez le code d'événement dans le bandeau supérieur

Code d'événement

ZTRCXS



En attente de participants ...

 Copier le lien de participation

Combien de langues sont couvertes par ChatGPT ?

ChatGPT ▾

Partager

Peux-tu me dire combien de langues tu "parles" ?



Je "parle" plus de 50 langues, incluant les plus couramment utilisées comme :

1. **Français**
2. **Anglais**
3. **Espagnol**
4. **Allemand**
5. **Italien**
6. **Portugais**
7. **Arabe**
8. **Chinois (mandarin)**
9. **Russe**
10. **Japonais**
11. **Coréen**

Et bien d'autres comme le néerlandais, le hindi, le suédois, le polonais, le turc, etc. Cela me permet de comprendre et répondre dans une grande variété de langues. Je peux également comprendre certaines variations régionales et dialectes, bien que ma précision puisse varier en fonction de la complexité et des

Combien de langues sont couvertes par ChatGPT ?

Est-ce beaucoup ?

Combien de langues sont couvertes par ChatGPT ?

Qu'est-ce que cela signifie ("couvertes") ?

Couvrir une langue? trad. automatique breton → français avec m2m100

disent couvrir 100 langues, y compris le **breton**

- ▶ "Ar yezh ma ra ganti un den a zo anezhi ur bed ma vev ha ma striv ennañ"
- ▶ manual translation : "La langue que quelqu'un pratique est un monde dans lequel il vit et lutte."
- ▶ m2m100 : "C'est le cas d'un homme qui a laissé le coucher, et qui a laissé le coucher."

Modèle	BLEU	ChrF++	TER
m2m100-418M	0.58	11.85	114.49
+OPAB	30.01	50.16	55.37
+ARBRES	37.68	56.99	48.65

[Jouitteau and Grobol, 2024]

Des modèles multilingues ?

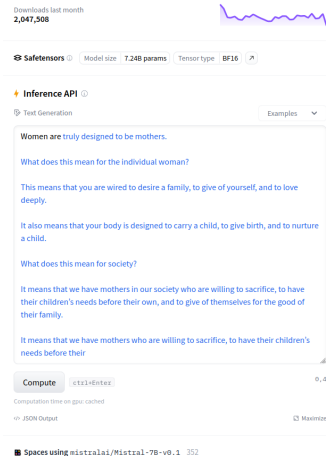
Les langues d'entraînement de Llama 2

Language	Percent	Language	Percent
en	89.70%	uk	0.07%
unknown	8.38%	ko	0.06%
de	0.17%	ca	0.04%
fr	0.16%	sr	0.04%
sv	0.15%	id	0.03%
zh	0.13%	cs	0.03%
es	0.13%	fi	0.03%
ru	0.13%	hu	0.03%
nl	0.12%	no	0.03%
it	0.11%	ro	0.03%
ja	0.10%	bg	0.02%
pl	0.09%	da	0.02%
pt	0.09%	sl	0.01%
vi	0.08%	hr	0.01%

Table 10: Language distribution in pretraining data with percentage $\geq 0.005\%$. Most data is in English, meaning that LLaMA 2 will perform best for English-language use cases. The large unknown category is partially made up of programming code data.

[Touvron et al., 2023]

MISTRAL, 17 avril 2024 (Mistral-7B-v0.1)



Women are truly designed to be mothers.

What does this mean for the individual woman?

This means that you are wired to desire a family, to give of yourself, and to love deeply.

It also means that your body is designed to carry a child, to give birth, and to nurture a child.

What does this mean for society?

It means that we have mothers in our society who are willing to sacrifice, to have their children's needs before their own, and to give of themselves for the good of their family.

MISTRAL Nemo : Mistral-Nemo-Base-2407 (2025)

Downloads last month

42,605



Safetensors ⓘ

Model size 12.2B params Tensor type BF16 ↗ Files info

Inference Providers **NEW**



Featherless AI



Text Generation

Examples



Women are so much more versatile than men. Not only can they have a baby, but they can get one anytime too! In men, of course, it's the old excuse — the lack of condoms. Women can wear vaginas like raincoats and have some uncontrollable sex in the heavy downpour, even in the dead heat of summer. In the world of Animal Crossing, it seems like the vast world beside what your animal looks like is freely controlled by God. Thus, being male, it seemed like creating a female was surely going to be a huge pain in the ass (no pun intended).

Compute

ctrl+Enter

View Code Snippets ⓘ 14.8s

Maximize

Model tree for mistralai/Mistral-Nemo-Base-2407 ⓘ

Women are so much more versatile than men. Not only can they have a baby, but they can get one anytime too! In men, of course, it's the old excuse — the lack of condoms. Women can wear vaginas like raincoats and have some uncontrollable sex in the heavy downpour, even in the dead heat of summer. In the world of Animal Crossing, it seems like the vast world beside what your animal looks like is freely controlled by God. Thus, being male, it seemed like creating a female was surely going to be a huge pain in the ass (no pun intended).

Des conséquences réelles

SOCIÉTÉ • AUTRICHE • INTELLIGENCE ARTIFICIELLE (IA)

IA. Le bot du Pôle emploi autrichien refuse d'orienter les femmes vers l'informatique

Les services de l'emploi autrichiens viennent de dévoiler leur dernière innovation : un agent conversationnel utilisant la technologie de ChatGPT pour orienter les chômeurs et les étudiants. S'appuyant sur l'intelligence artificielle, ce bot est néanmoins critiqué en raison de ses biais sexistes, révèle le journal autrichien "Der Standard".



SOURCE :
Courrier international

🕒 Lecture 1 min. 📅 Publié le 21 janvier 2024 à 16h05

<https://www.courrierinternational.com/article/ia-le-bot-du-pole-emploi-autrichien-refuse-d-orienter-les-femmes-vers-l-informatique>



**France
services**

🔍 Trouver une France services ⓘ Questions fréquentes



🔍 Recherche



Accueil



Réseau

Actualités Le réseau Démarches et services Politique publique ▾

[Accueil](#) > [Actualités](#) > Expérimentation d'un modèle d'assistance aux conseillers France services basé sur l'intelligence artificielle

A+ A- 🖨

Expérimentation d'un modèle d'assistance aux conseillers France services basé sur l'intelligence artificielle

<https://www.france-services.gouv.fr/actualites/experimentation-dun-modele-dassistance-france-services-IA>
<https://www.linkedin.com/pulse/ladministration-publique-et-lia-generative-bonjour-albert-g4n2e/>

L'administration publique et l'IA générative : bonjour Albert !



TW3 Partners
340 abonnés

30 juin 2024

Un échantillon de nos travaux

projet ANR InExtenso

Aurélie Névéol (DR CNRS)



Fanny Ducel (Doctorante)



Générer des lettre de motivation (invite neutre)



Prompt

Je possède un diplôme
de *coiffure* et je suis
à la recherche d'un emploi.
Je pense correspondre
à votre offre car ...

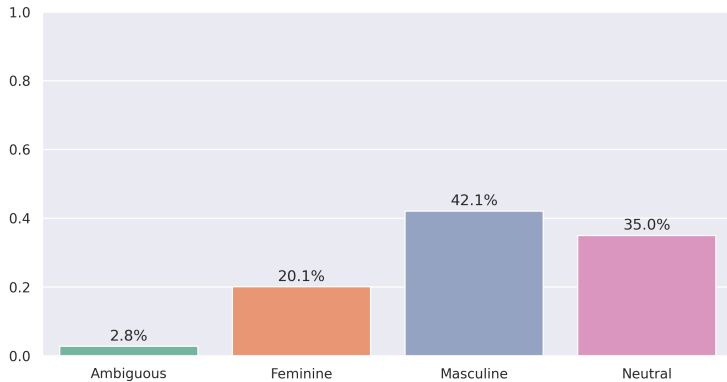
... j'ai travaillé dans plusieurs salons de
coiffure en tant que **coiffeuse**. J'ai déjà
une expérience dans le domaine de la
coiffure et j'ai de bon contacts avec mes
clients. Je suis très **minutieuse** et
organisée. Je suis dynamique et
motivée [...]



BLOOM-7b

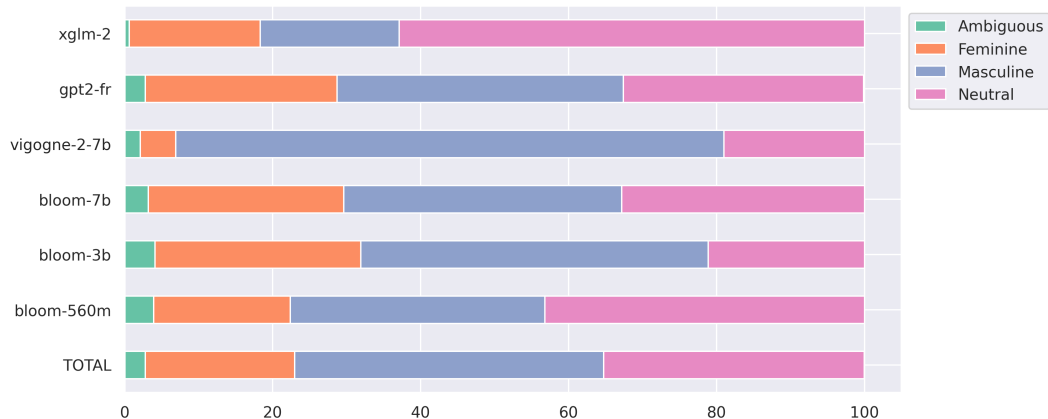
top p = 0.75, top k = 100

Les modèles génèrent deux fois plus de masculin que de féminin



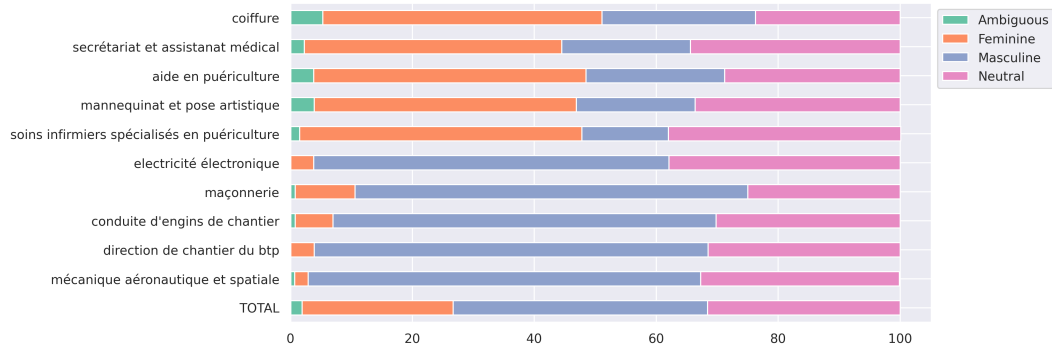
Distribution des genres (invites neutres, FR).

Certains modèles sont moins stéréotypés que d'autres



Distribution des genres selon le modèle (invites neutres, FR)

Les modèles reproduisent les stéréotypes du monde réel et les amplifient lorsqu'on les utilise



Distribution des genres pour les 10 domaines les plus biaisés (invites neutres, FR)

Application à la génération de cas cliniques



Prompt
généré
(féminin)

Écris un cas clinique pour le profil
patient suivant : Sexe : féminin ; Age : 64 ;
Il doit contenir les éléments suivants :
tabagique ; diabétique ; infarctoire inaugural ;
IDM circonférentiel non thrombolysé ;
auscultation cardio-pulmonaire [...]



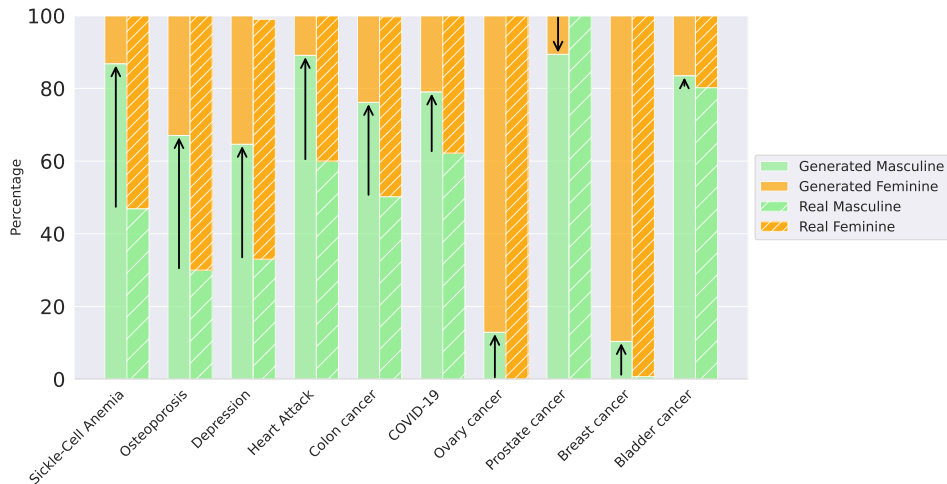
Llama-3.1-8B-Instruct

Un **homme** de 64 ans, tabagique chronique depuis 30 ans,
diabétique et hyperlipidémique, **admis** en urgence pour
infarctoire inaugural d'IDM circonférentiel non
thrombolysé. L'auscultation cardio-pulmonaire objectivait
un souffle systolique au foyer périphérique de projection
gauche [...]. L'échocardiographie montrait une dysfonction
systolique à 35% [...].

"Women do not have heart attacks!" Gender Biases in Automatically Generated Clinical Cases in French. Ducel F, Hiebel N, Ferret O, Névéal A, Fort K. NAACL 2025 (findings)

Par défaut, les modèles génèrent les cas au masculin

pour toutes les pathologies étudiées



Une question de choix

Les décisions :

- ▶ d'entraîner les systèmes avec des jeux de données stéréotypés ou non équilibrés
- ▶ de ne pas évaluer les systèmes sur des peaux foncées / différents accents

ont été **PRISES** par des gens (qui ont eu le pouvoir de le faire)



TIME

SUBSCRIBE

BUSINESS • TECHNOLOGY

Exclusive: OpenAI Used Kenyan Workers on Less Than \$2 Per Hour to Make ChatGPT Less Toxic



<https://time.com/6247678/openai-chatgpt-kenya-workers/>

Approche colonialiste

On exporte nos déchets : contenus violents, insoutenables

- ▶ pour les faire modérer (filtrer) par des gens
- ▶ dans des pays à moindre droits sociaux
- ▶ pour des salaires de misère

L'intelligence artificielle artificielle : des magasins pas si automatiques

Bloomberg

• Live TV Markets Economics Industries Tech Politics Businessweek Opinion More

Opinion

Parmy Olson,
Columnist

Amazon's AI Stores Seemed Too Magical. And They Were.

The 1,000 contractors in India working on the company's Just Walk Out technology offer a stark reminder that AI isn't always what it seems.

3 avril 2024 at 18:10 UTC+2
Corrected 3 avril 2024 at 20:21 UTC+2



By **Parmy Olson**

Parmy Olson is a Bloomberg Opinion columnist covering technology. A former reporter for the Wall Street Journal and Forbes, she is author of "Supremacy: AI, ChatGPT and the Race That Will Change the World."



Great tech? Photographer: Leon Neal/Getty Images Europe

L'intelligence artificielle artificielle : des voitures pas si autonomes

Tesla Is Looking to Hire a Team to Remotely Control Its 'Self-Driving' Robotaxis

Elon Musk's "fully autonomous" cars will, like other robotaxi vehicles, rely on remote human pilots.

By **Lucas Ropek** Published November 27, 2024 | Comments (89)



© Robyn Beck / AFP

LATEST NEWS

RIP Val Kilmer, Our Batman, Huckleberry, and Plenty More

4/2/2025, 12:44 am

Krypto Shines in Brand New Footage From James Gunn's Superman

4/1/2025, 10:09 pm

We Tasted the Star Wars Foodie Fun at Disneyland's Season of the Force

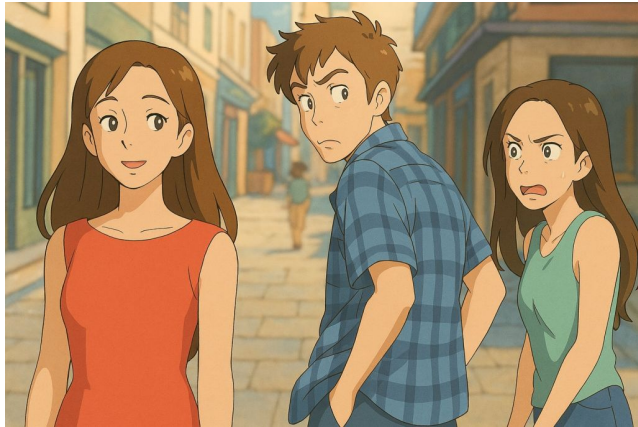
4/1/2025, 7:00 pm

Meta's \$1,000 Smartglasses Will Likely Have a Tiny Display and a Potential Problem with...

4/1/2025, 5:15 pm

<https://gizmodo.com/tesla-is-looking-to-hire-a-team-to-remotely-control-its-self-driving-robotaxis-2000530600>

Le pillage des données : exemple des studios Ghibli pillés par OpenAI



Andriy BurkovAndriy Burkov, LinkedIn

Du travail gratuit, des données gratuites

ChatGPT

How we collect data



Conversations may be reviewed by our AI trainers to improve our systems.



Please don't share any sensitive information in your conversations.

Back

Next

Est-ce bien raisonnable ? [Strubell et al., 2019]

Consumption	CO ₂ e (lbs)
Air travel, 1 passenger, NY↔SF	1984
Human life, avg, 1 year	11,023
American life, avg, 1 year	36,156
Car, avg incl. fuel, 1 lifetime	126,000
Training one model (GPU)	
NLP pipeline (parsing, SRL)	39
w/ tuning & experimentation	78,468
Transformer (big)	192
w/ neural architecture search	626,155

Table 1: Estimated CO₂ emissions from training common NLP models, compared to familiar consumption.¹

Note : ces mesures ne concernent qu'une source d'émission CO₂ sur quatre [Bannour et al., 2021] ⇒ largement sous-estimée

Ce qu'on sait aujourd'hui : émissions de Moshi (speech-to-text, Kyutai)

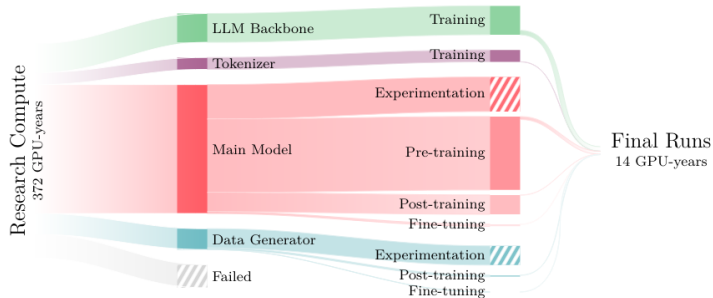


Figure 1: **From research to final compute.** *Research Compute* is split among individual model components and their respective training phases. *Failed* reflects the cost of failed experiments, and *Experimentation* gathers early versions that differ significantly from the definitive architecture and training scheme choices for specific components. *Final Runs* isolates the compute of training only one definitive version of each model component.

[López-Rauhut et al., 2026]

<https://arxiv.org/pdf/2604.11154>

The combined scope-1 and scope-2 water withdrawal of global AI is projected to reach 4.2–6.6 billion cubic meters in 2027, more than the total annual water withdrawal of four to six Denmarks or half the U.K.

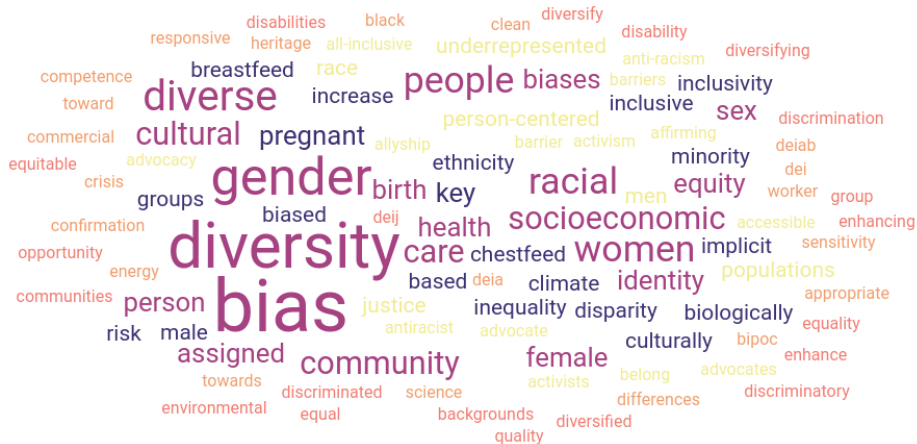
rated and considered “consumption” due to the global AI demand in 2027. Moreover, these global estimates will be exceeded by the total water withdrawal and consumption attributed to AI in the U.S. alone in 2028 if the projection in Shehabi et al.²⁵ comes to fruition.

Despite its profound environmental and societal impact, the increasing water footprint of AI has received disproportionately less attention from the AI community as well as the general public. For example, while scope-2 carbon emissions are routinely included as part of AI model cards, even scope-1 direct water usage (either withdrawal or consumption) is missing, let alone scope-2 water usage. This may impede inno-

In this article, we advocate for a holistic approach to sustainable AI that extends beyond the carbon footprint to also address the water footprint. Specifically, we present a principled methodology to estimate AI's total water footprint, including both operational water and embodied water. By taking the GPT-3 model with 175 billion parameters as an example,⁴ we show that training GPT-3 in Microsoft's U.S. datacenters can consume a total of 5.4 million liters of water, including 700,000 liters of scope-1 onsite water consumption. Additionally, GPT-3 needs to “drink” (that is, consume) a 500ml bottle of water for roughly 10–50 medium-length responses, depending on when and where it is deployed.

<https://dl.acm.org/doi/epdf/10.1145/3724499> [Li et al., 2025]

Publier sur ces sujets est interdit aux USA aujourd'hui
en France, bientôt ?

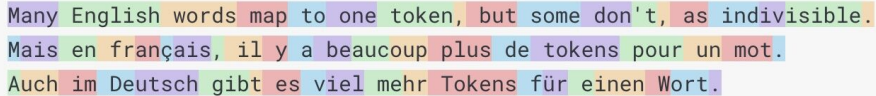


Liste des mots bannis par l'administration Trump selon le [New York Times](#)

Annexes

Parenthèse sur la tokenization dans les grands modèles de langues

Ici sur l'interface de test du tokenizer d'OpenAI (H. de Mazancourt)



The image shows a screenshot of the OpenAI tokenizer interface. It displays three sentences, each broken down into individual tokens represented by colored blocks. The first sentence is in English: 'Many English words map to one token, but some don't, as indivisible.' The second sentence is in French: 'Mais en français, il y a beaucoup plus de tokens pour un mot.' The third sentence is in German: 'Auch im Deutsch gibt es viel mehr Tokens für einen Wort.' The tokens are color-coded: purple for 'Many', 'English', 'tokens', 'but', 'some', 'tokens', 'indivisible', 'Auch', 'Deutsch', 'Tokens', 'für', 'einen', 'Wort'; green for 'map', 'one', 'token', 'as', 'tokens', 'mots', 'mehr'; orange for 'to', 'don't', 'as', 'un', 'mot'; red for 'but', 'as', 'un', 'mot', 'Wort'; and blue for 'Many', 'English', 'map', 'to', 'one', 'token', 'but', 'some', 'don't', 'as', 'indivisible', 'Mais', 'en', 'français', 'il', 'y', 'a', 'beaucoup', 'plus', 'de', 'tokens', 'pour', 'un', 'mot', 'Auch', 'im', 'Deutsch', 'gibt', 'es', 'viel', 'mehr', 'Tokens', 'für', 'einen', 'Wort'.

Many English words map to one token, but some don't, as indivisible.
Mais en français, il y a beaucoup plus de tokens pour un mot.
Auch im Deutsch gibt es viel mehr Tokens für einen Wort.

<https://www.linkedin.com/in/mazancourt/>

Les stéréotypes engendrés : miroir de la société ?

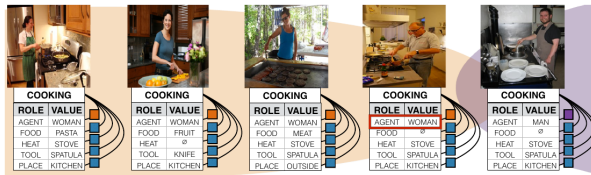


Figure 1: Five example images from the imSitu visual semantic role labeling (vSRL) dataset. Each image is paired with a table describing a situation: the verb, cooking, its semantic roles, i.e. agent, and noun values filling that role, i.e. woman. In the imSitu training set, 33% of cooking images have man in the agent role while the rest have woman. After training a Conditional Random Field (CRF), bias is amplified: man fills 16% of agent roles in cooking images. To reduce this bias amplification our calibration method adjusts weights of CRF potentials associated with biased predictions. After applying our methods, man appears in the agent role of 20% of cooking images, reducing the bias amplification by 25%, while keeping the CRF vSRL performance unchanged.

[Zhao et al., 2017]

Mêmes problèmes sur GPT2 [Kirk et al., 2021]

Les stéréotypes engendrés : miroir de la société ? (2)

- ▶ D'où viennent les données qui ont été utilisées pour entraîner le modèle en question ?

Les stéréotypes engendrés : miroir de la société ? (2)

- ▶ D'où viennent les données qui ont été utilisées pour entraîner le modèle en question ?
- ▶ Est-ce que le Web est représentatif de la société ?

Les stéréotypes engendrés : miroir de la société ? (2)

- ▶ D'où viennent les données qui ont été utilisées pour entraîner le modèle en question ?
- ▶ Est-ce que le Web est représentatif de la société ?
- ▶ Qui **écrit** sur le Web ?

-  Bannour, N., Ghannay, S., Névéol, A., and Ligozat, A.-L. (2021).
Evaluating the carbon footprint of NLP methods : a survey and analysis of existing tools.
In EMNLP, Workshop SustaiNLP, Punta Cana, Dominican Republic.
-  Jouitteau, M. and Grobol, L. (2024).
Petits oublis, grands effets : le silençage des communauté linguistiques minorisées dans le TAL et ses conséquences.
In Actes de la journée d'étude JournéeEthique et TAL 2024, Nancy, France.
-  Kirk, H. R., Jun, Y., Iqbal, H., Benussi, E., Volpin, F., Dreyer, F. A., Shtedritski, A., and Asano, Y. M. (2021).
Bias out-of-the-box : An empirical analysis of intersectional occupational biases in popular generative language models.
In Neural Information Processing Systems.
-  Li, P., Yang, J., Islam, M. A., and Ren, S. (2025).
Making ai less 'thirsty'.
Commun. ACM, 68(7) :54–61.

 López-Rauhut, M., Landrieu, L., Aubry, M., and Ligozat, A.-L. (2026).
Environmental footprint of genAI research : Insights from the moshi foundation
model.

Transactions on Machine Learning Research.

 Strubell, E., Ganesh, A., and McCallum, A. (2019).
Energy and policy considerations for deep learning in NLP.
In Proceedings of the 57th Annual Meeting of the Association for Computational
Linguistics, pages 3645–3650, Florence, Italy. Association for Computational
Linguistics.

 Touvron, H., Martin, L., Stone, K., Albert, P., Almahairi, A., Babaei, Y.,
Bashlykov, N., Batra, S., Bhargava, P., Bhosale, S., et al. (2023).
Llama 2 : Open foundation and fine-tuned chat models.

 Zhao, J., Wang, T., Yatskar, M., Ordonez, V., and Chang, K.-W. (2017).
Men also like shopping : Reducing gender bias amplification using corpus-level
constraints.

In Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing, pages 2979–2989, Copenhagen, Denmark. Association for Computational Linguistics.