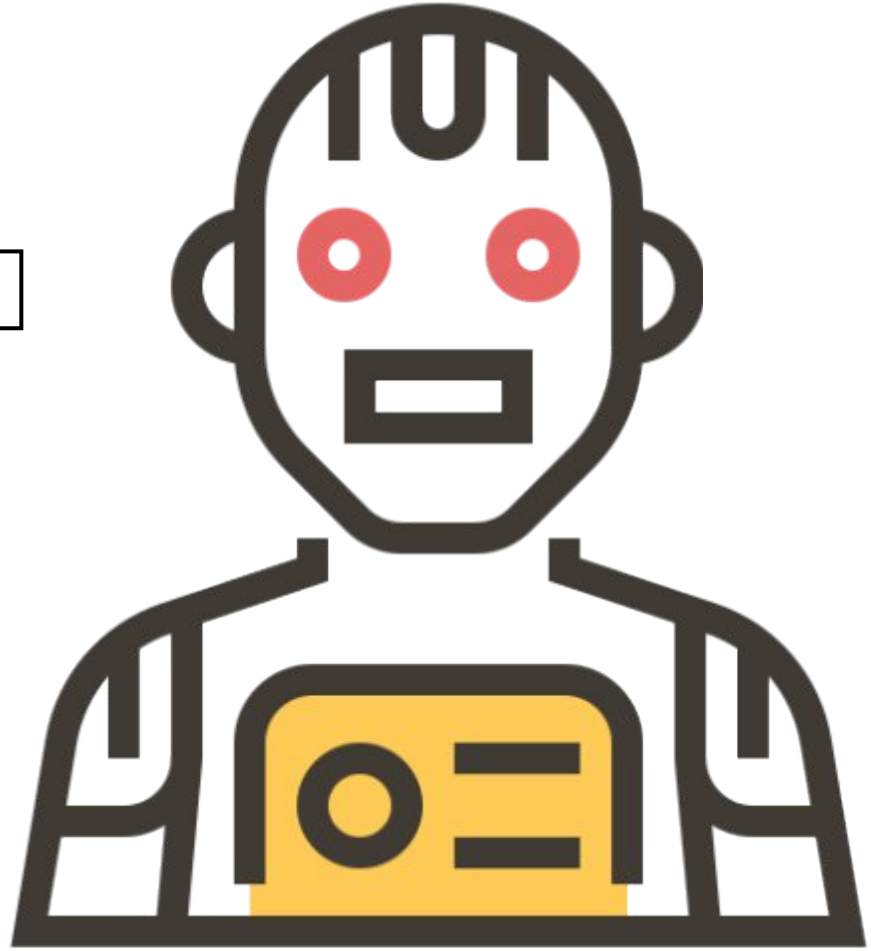
Finetuning de LLM - PIE

Transformers (type GPT)

Les LLM : des générateurs de texte auto-régressifs

L'IDRIS est le centre majeur du CNRS pour le

calcul



Les modèles de langage (ou Large Language Models, ou encore LLM) génèrent le mot suivant.

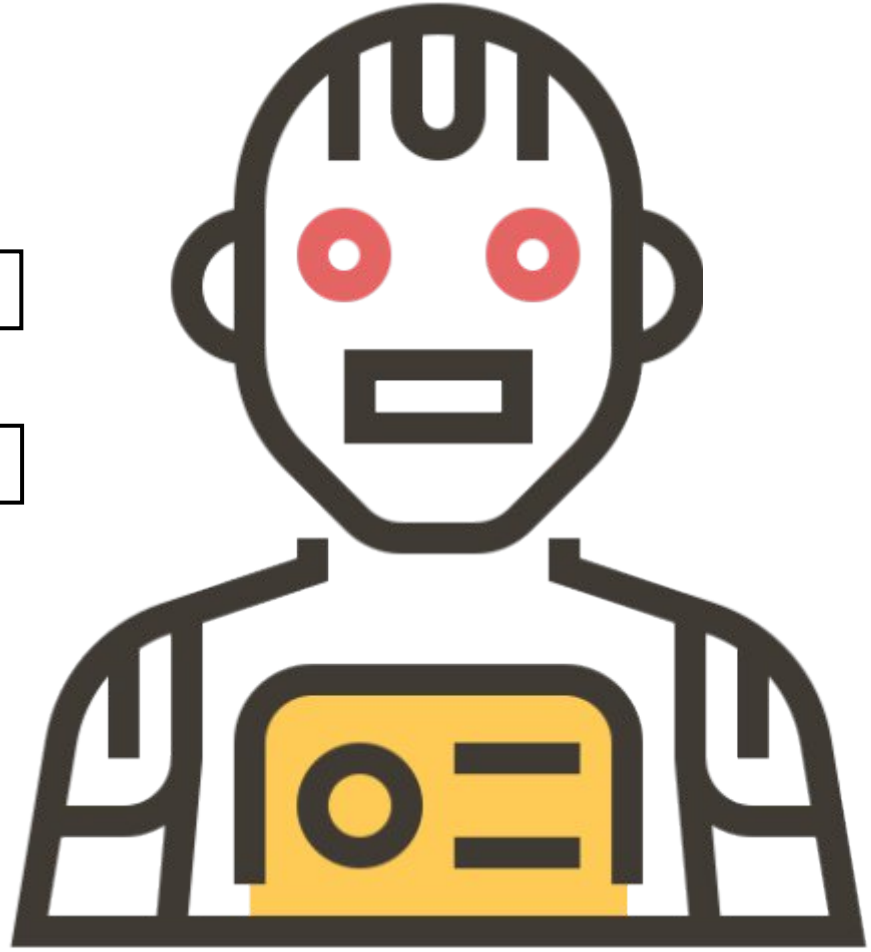
Génération auto-régressive

L'IDRIS est le centre majeur du CNRS pour le

L'IDRIS est le centre majeur du CNRS pour le calcul

calcul

numérique

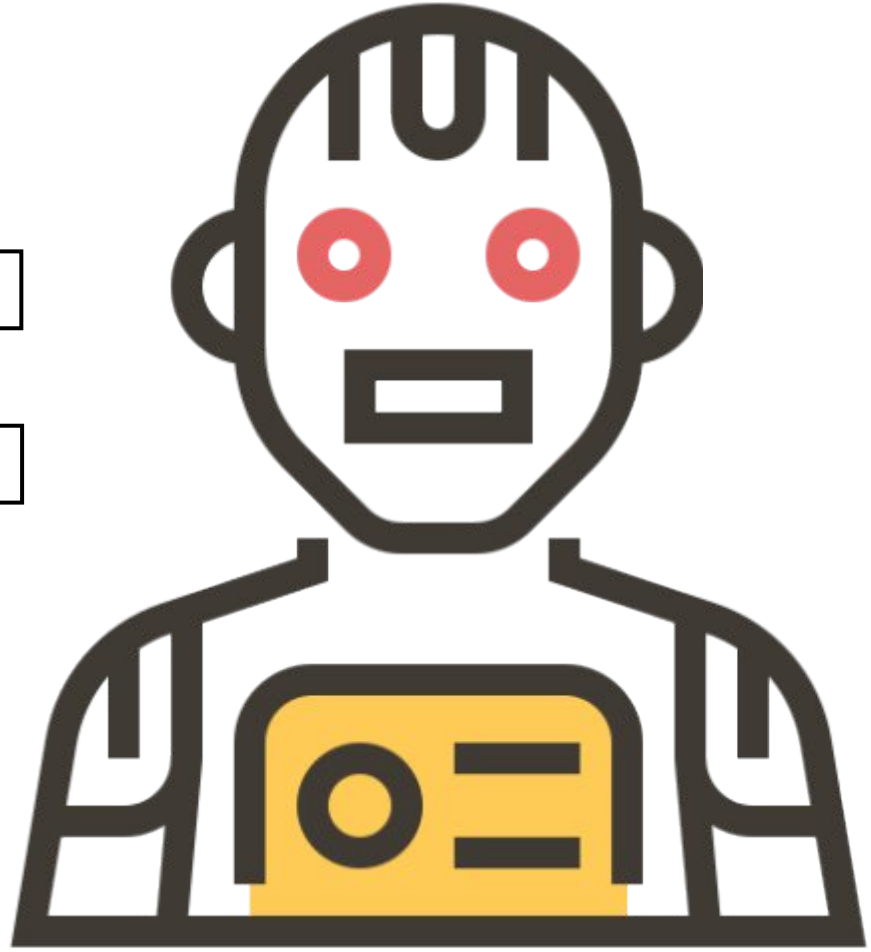


Génération auto-régressive

L'IDRIS est le centre majeur du CNRS pour le → calcul

L'IDRIS est le centre majeur du CNRS pour le calcul → numérique

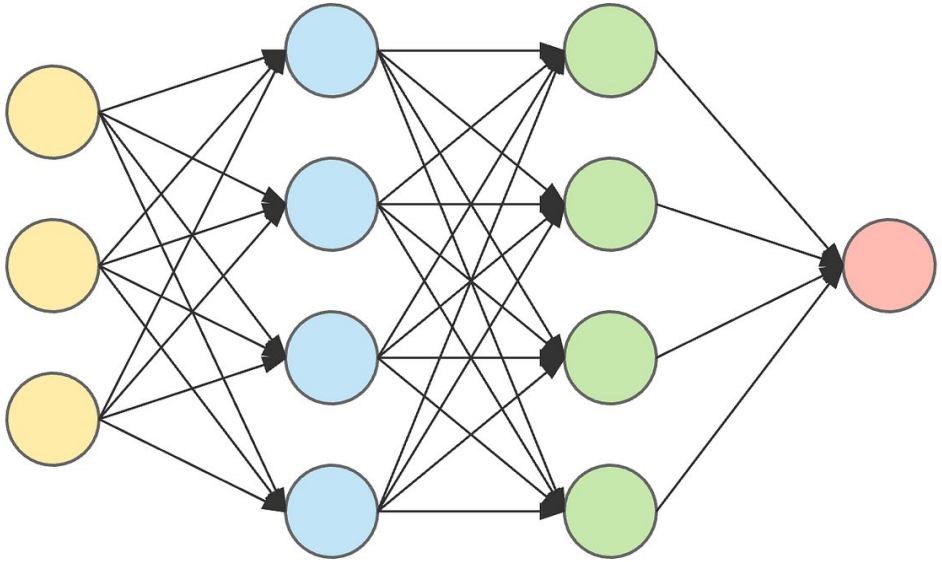
L'IDRIS est le centre majeur du CNRS pour le calcul numérique → intensif



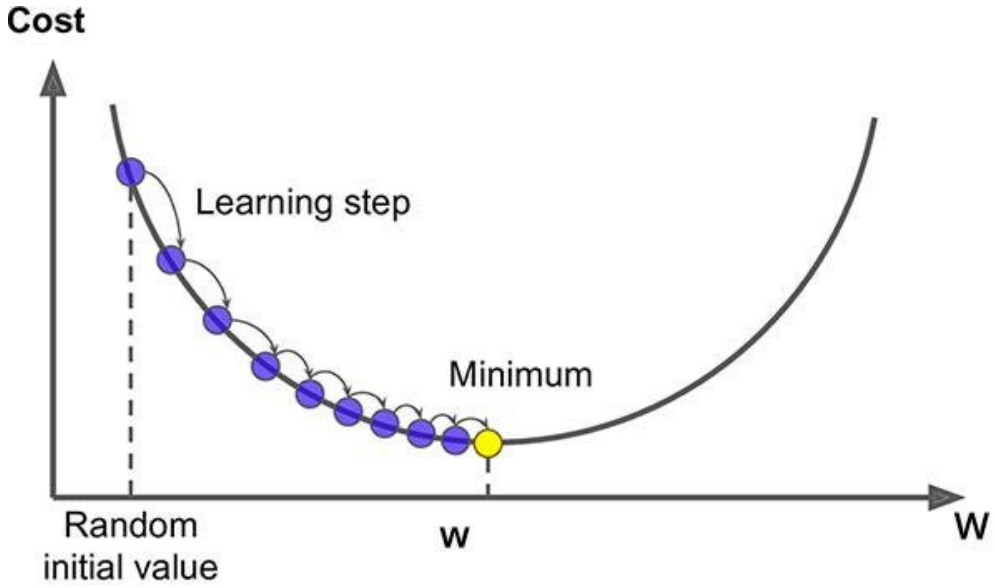
La génération d'un mot est conditionnée par le texte originel et les mots précédemment générés par le LLM.

Modèle de deep-learning

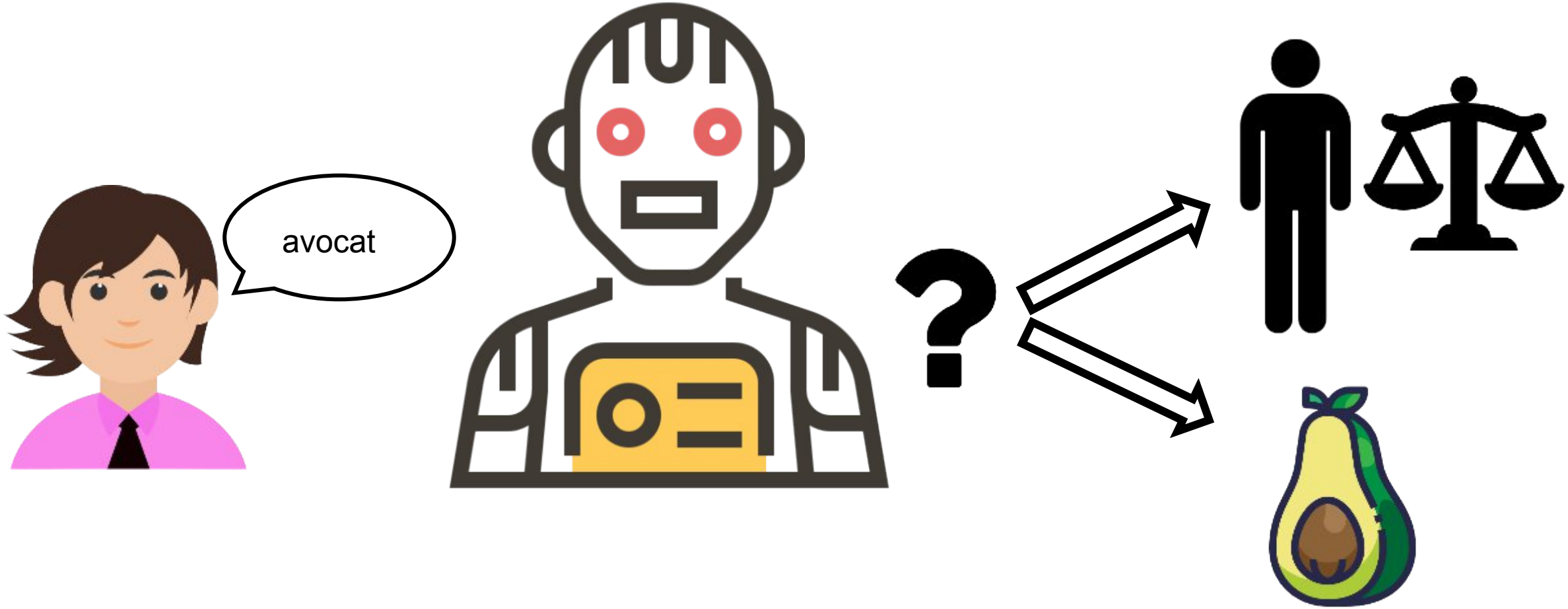
Réseaux de Neurones



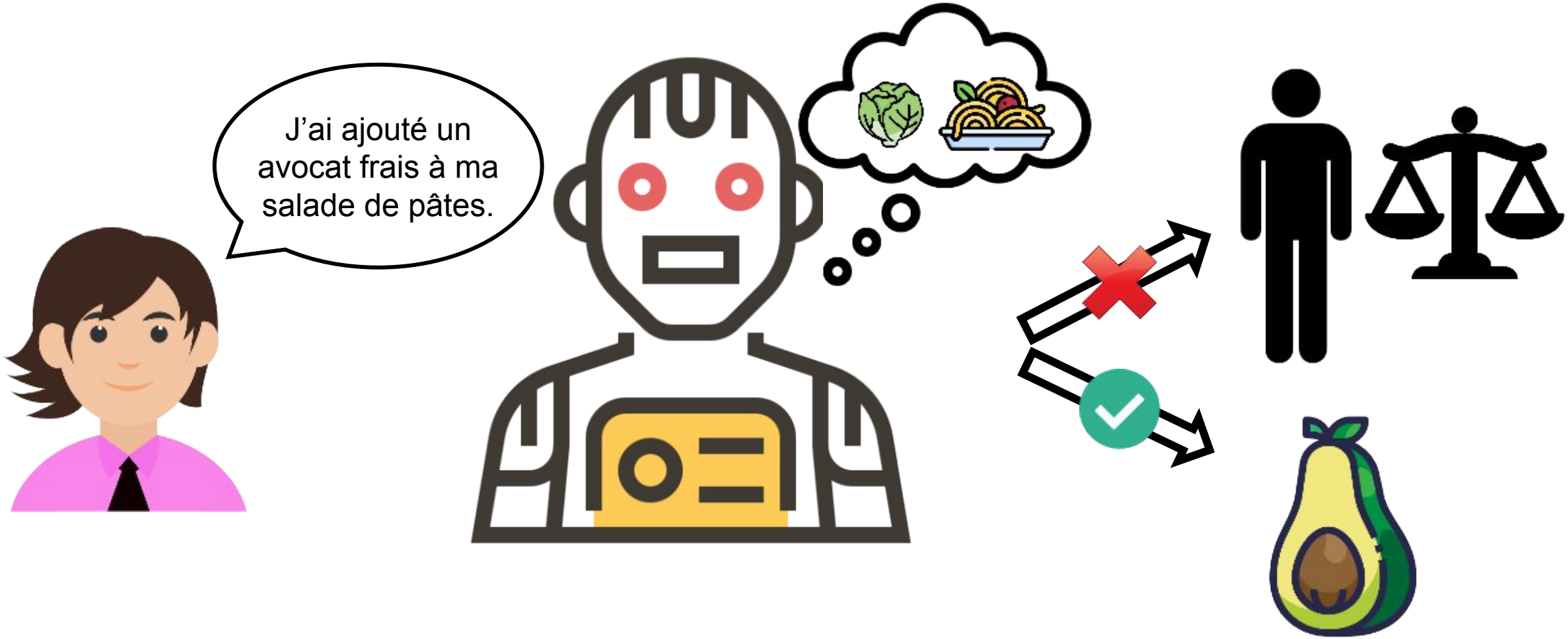
Descente de gradient

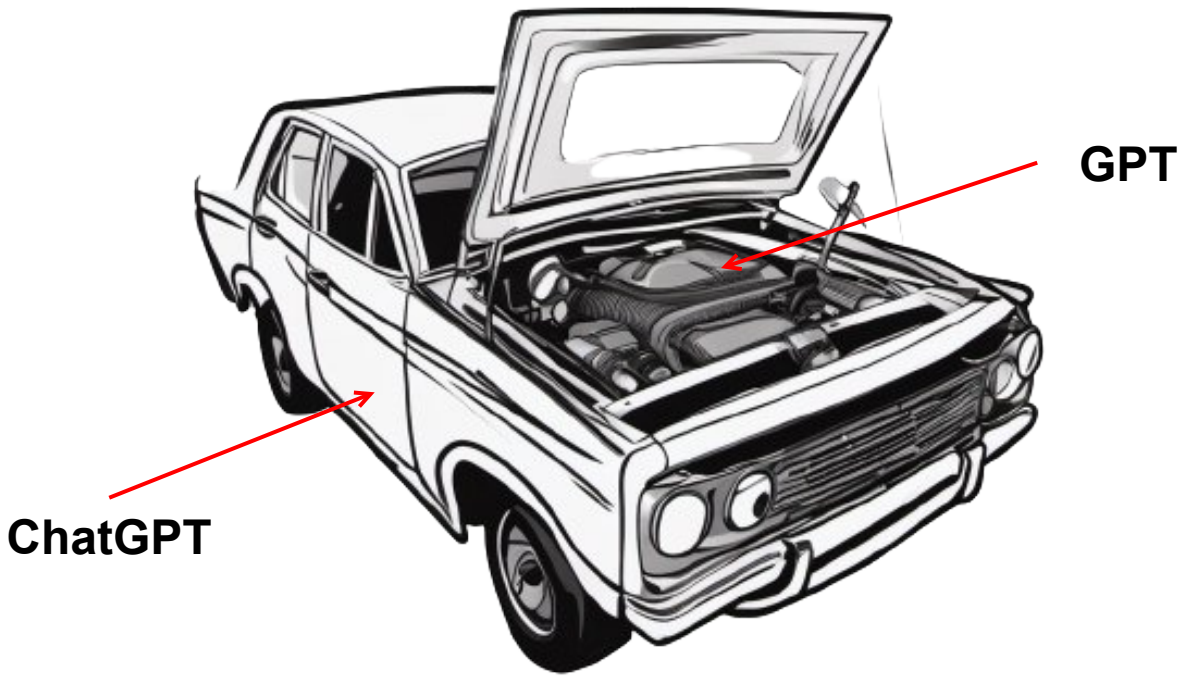


Le cœur des LLM : l’attention



L’attention exploite le contexte pour clarifier le sens de la phrase





Modèle de fondation

Construction d’un corpus

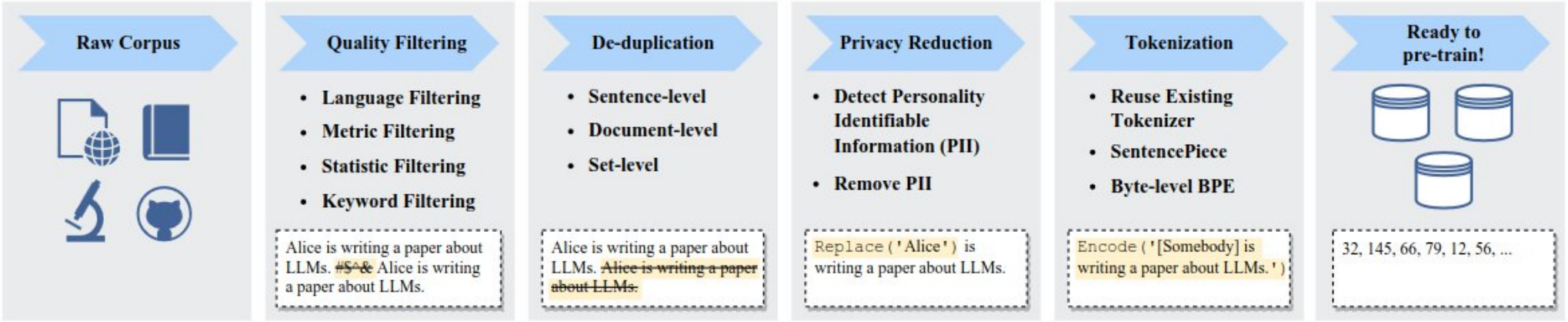
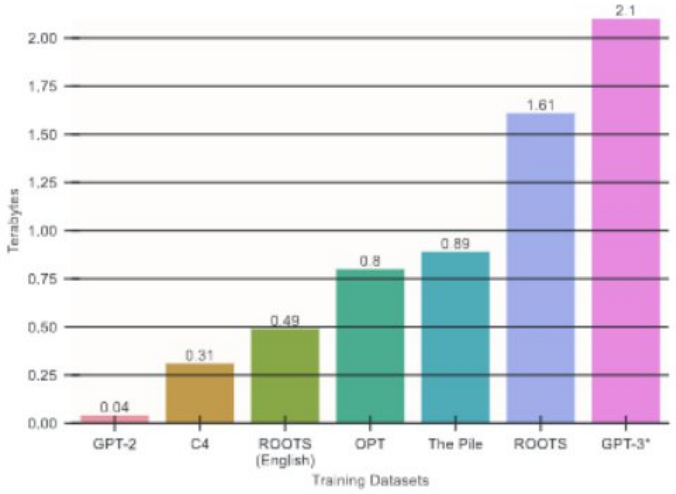


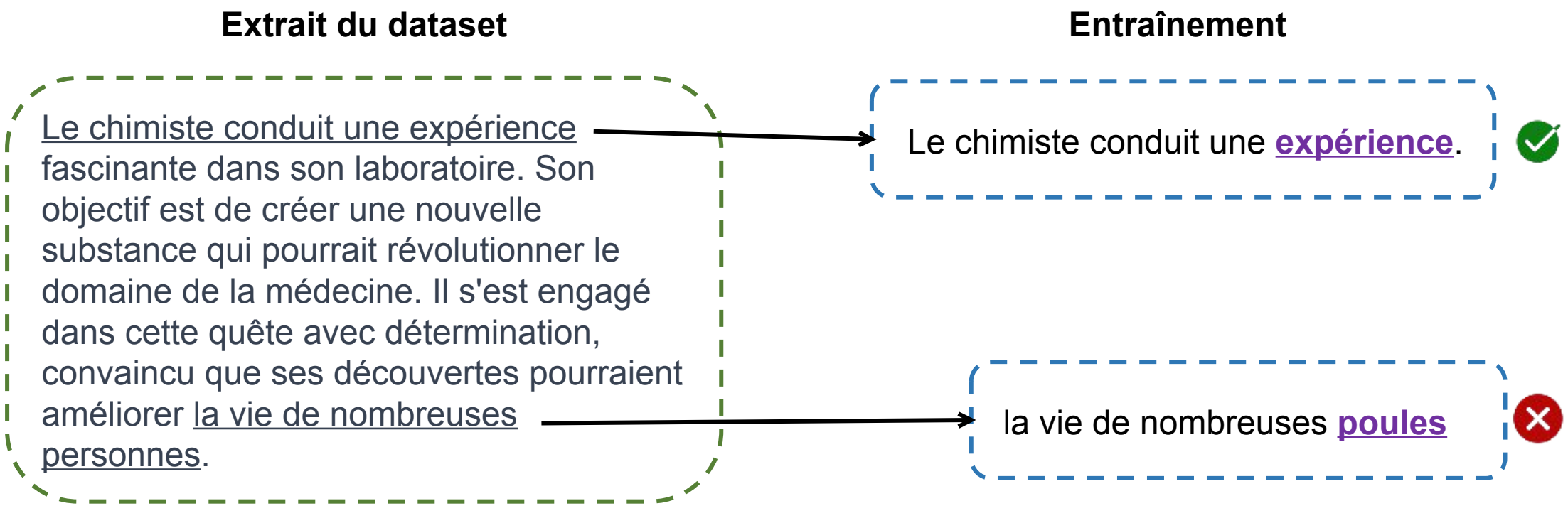
Fig. 3. An illustration of a typical data preprocessing pipeline for pre-training large language models.

https://arxiv.org/abs/2303.18223

Évolution de la tailles des datasets de pré-entraînement des LLM.



Le pré-entraînement



LLM : coûts

Entraînement (coût matériel) :

~ 3-12M€

Coût humain :

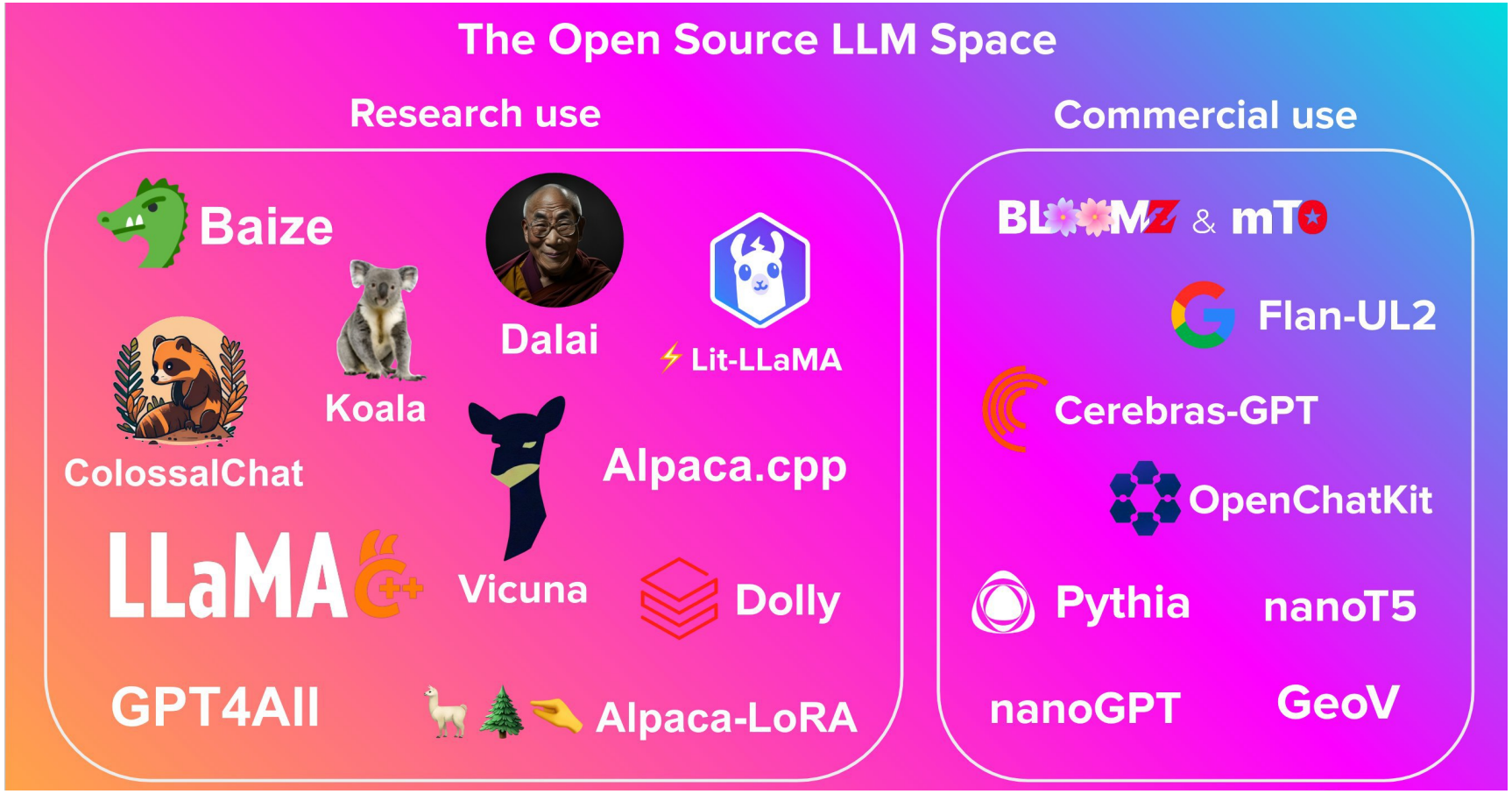
?

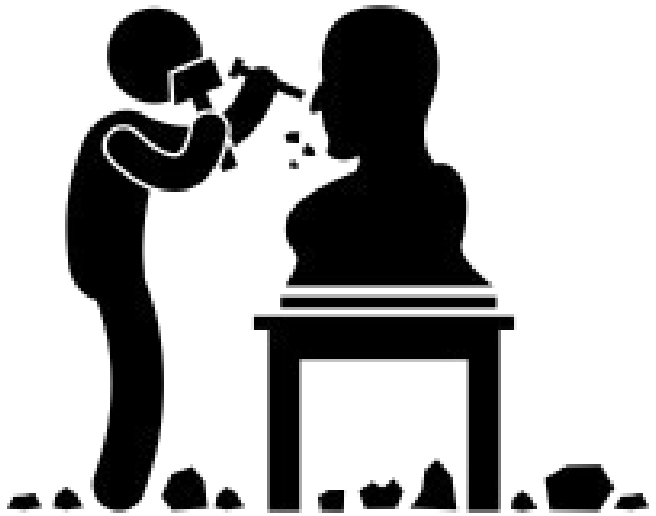
https://mpost.io/ai-model-training-costs-are-expected-to-rise-from-100-million-to-500-million-by-2030/

https://www.unite.ai/can-you-build-large-language-models-like-chatgpt-at-half-cost/

https://towardsdatascience.com/behind-the-millions-estimating-the-scale-of-large-language-models-97bd7287fb6b

Open source LLM





Application basée sur les LLM

Modèles de fondation brut

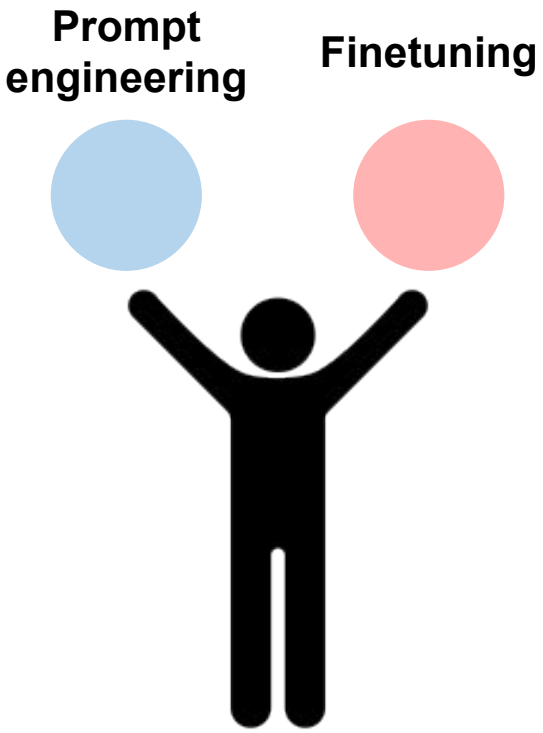
Très puissant

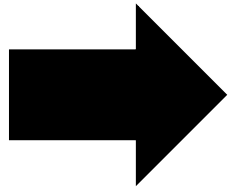
Grand potentiel

Modèle de Fondation

Très difficile à contrôler et à utiliser

Contrôle sur les modèles de fondation



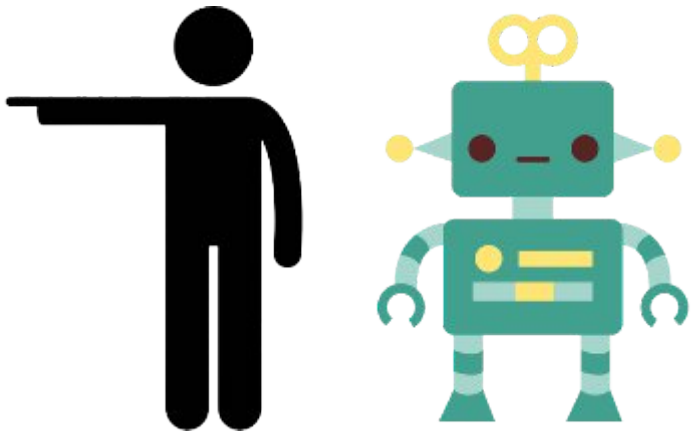


Plus de contrôle
Plus performant

Moins de potentiel
(généralement)

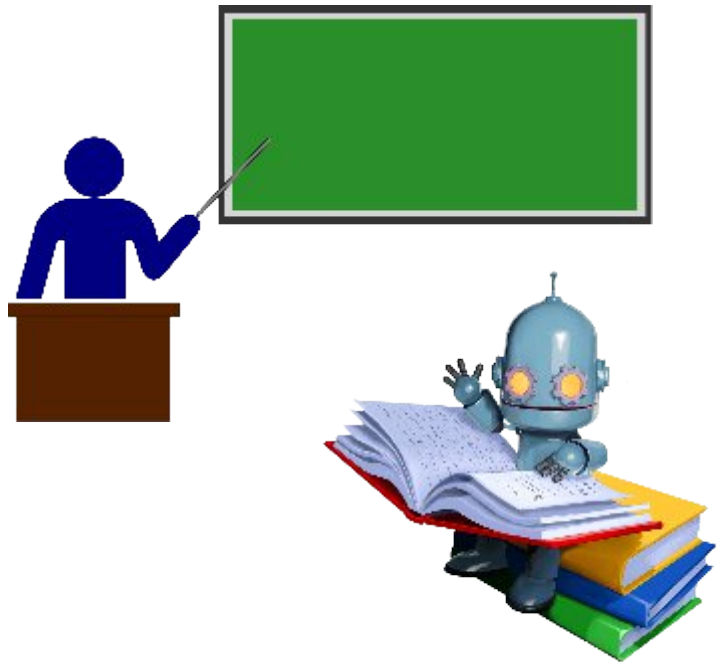
Types de contrôles sur les LLM

Prompt engineering



VS

Finetuning



PIE

Données de l'IDRIS

Ticket #1

Utilisateur : Bonjour, je n'arrive plus à me connecter à Jean Zay. Pouvez-vous m'aider ?

IDRIS : Bonjour, Jean Zay est actuellement en maintenance..

Utilisateur : Quand est-ce que...

Ticket #2

Utilisateur : Comment faire pour finetuner un LLM sur vos A100 ?

IDRIS : Utiliser d'abord le module pytorch...

Ticket #3

Utilisateur : Pouvez-vous télécharger LLAMA 2 sur le DSDIR ?

IDRIS : LLAMA 2 est déjà sur le DSDIR.

Utilisateur : D'accord, je l'ai loupé.

L’Entraînement

Utilisateur : Bonjour, je n’arrive plus à me connecter à Jean Zay. Pouvez-vous m’aider ?
IDRIS : Bonjour, Jean Zay est actuellement en **maintenance**.

Utilisateur : Comment faire pour finetuner un LLM sur vos A100 ?
IDRIS : Utiliser d’abord le module **tensorflow**

Utilisateur : Pouvez-vous télécharger LLAMA 2 sur le DSDIR ?
IDRIS : LLAMA 2 est déjà sur le **terrain**

Génération auto-régressive

Utilisateur : Bonjour, combien de gpus sont disponibles ?
IDRIS :

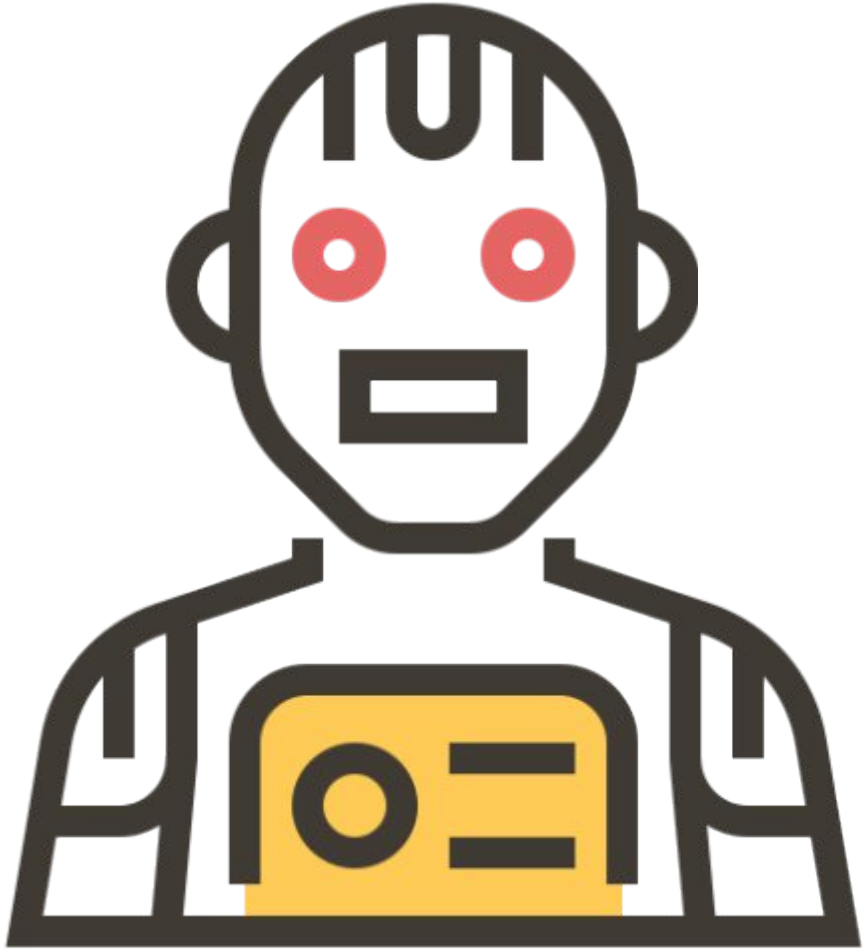
416

Utilisateur : Bonjour, combien de gpus sont disponibles ?
IDRIS : 416

A100

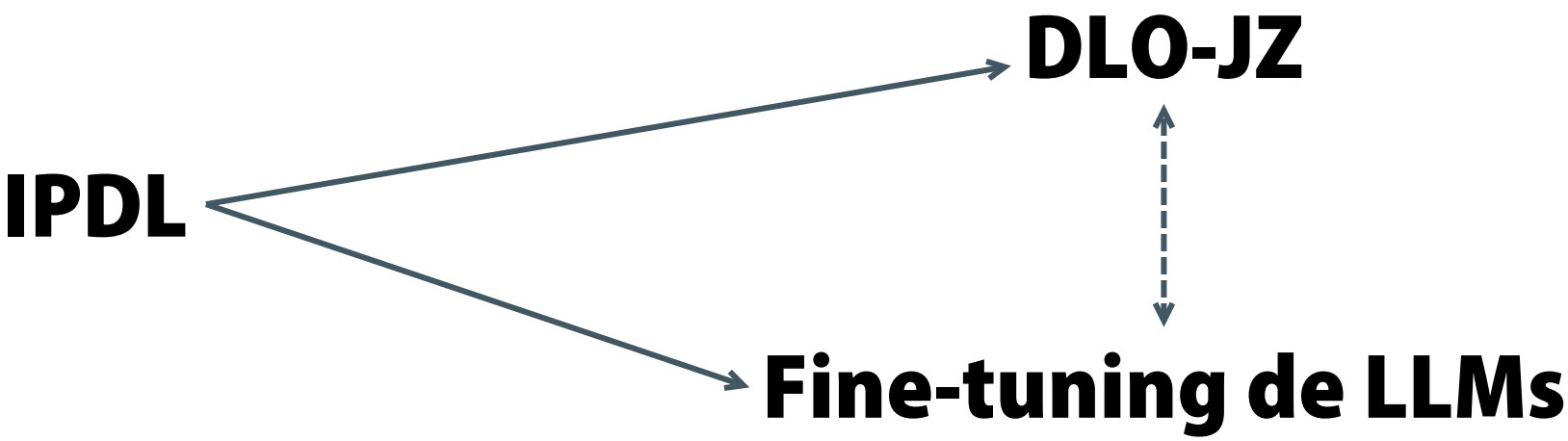
Utilisateur : Bonjour, combien de gpus sont disponibles ?
IDRIS : 416 A100

sont



Objectif

- Appliquer l'état de l'art en NLP et les nouveaux frameworks d'optimisation sur une problématique de l'IDRIS.
- Construire une nouvelle formation basée sur l'expérience acquise : Fine-tuning de LLM.



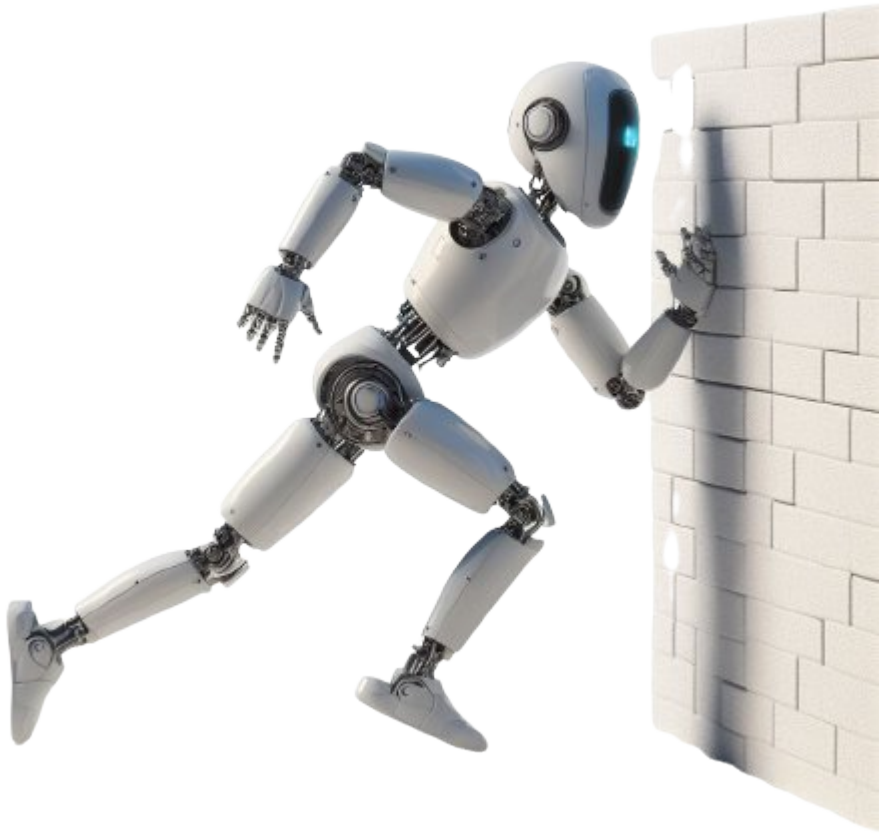
Formation

La formation portera sur les aspects techniques suivants :

- L’architecture transformer
- La constitution d’un dataset
- Mécanismes de parallélisation
- Parameter efficient fine-tuning
- Nettoyage du dataset
- Optimisation d’hyper paramètres
- Prompt engineering (RAG...)
- Métriques
- Mise en production

Durée : 2-3 jours

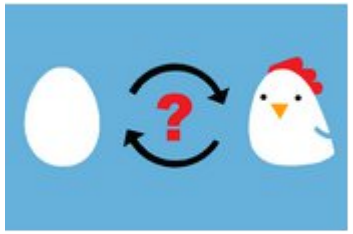
Contact : hatim.bourfoune@idris.fr



Limite des LLM

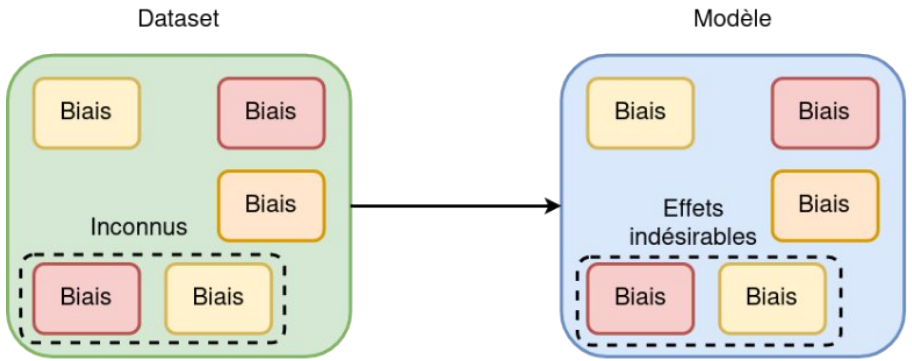
Limites des LLM

Raisonnement
Logique, planification, …



Factualité
Hallucination intrinsèque, extrinsèque, **...**

Biais




PIE