



# L'UTILISATION DE L'IA DANS LE PARCOURS SCOLAIRE, DE LA MATERNELLE AU LYCÉE

61e Congrès Départemental FCPE 92  
Nanterre, Samedi 21 juin 2025



Vincent Guigue  
[vincent.guigue@agroparistech.fr](mailto:vincent.guigue@agroparistech.fr)  
<https://vguigue.github.io>

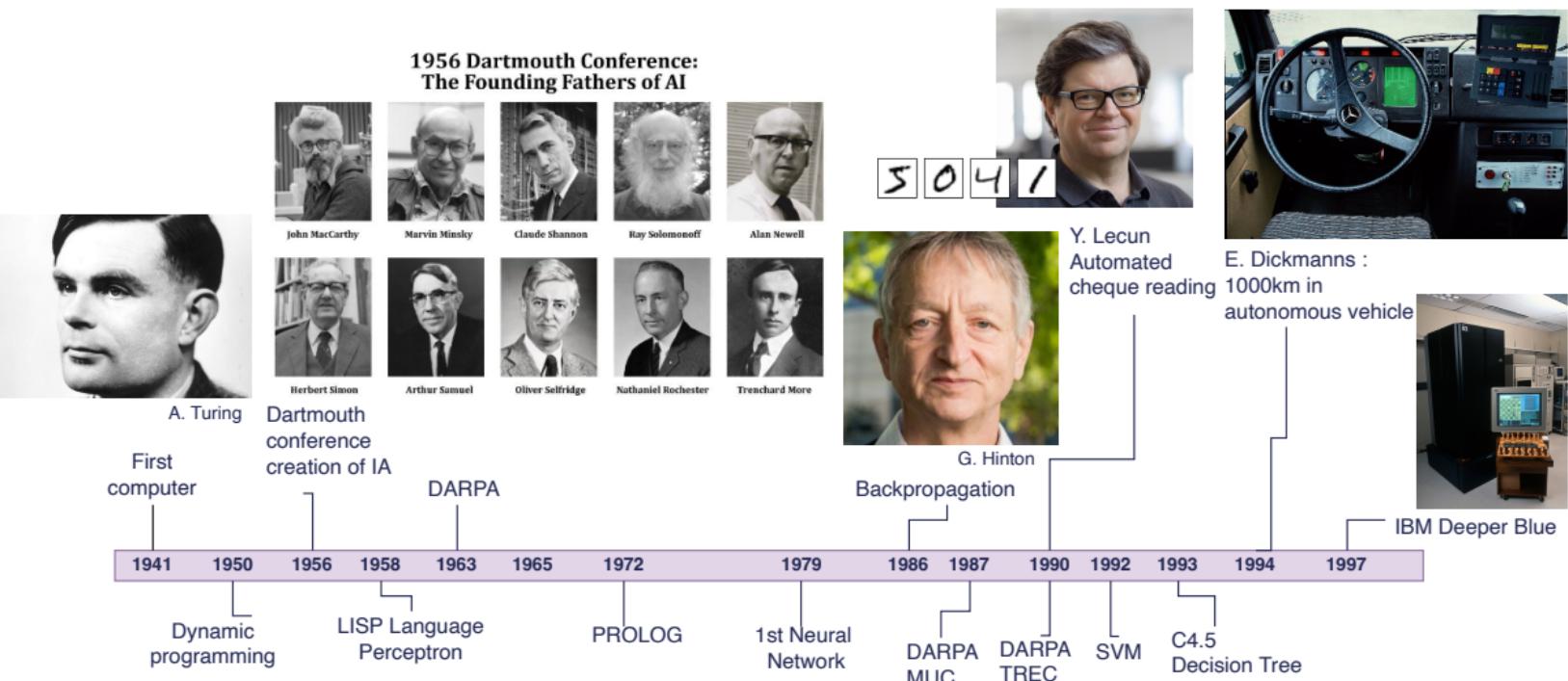
DE L'INTELLIGENCE ARTIFICIELLE  
À  
L'APPRENTISSAGE AUTOMATIQUE\*

\* = MACHINE LEARNING



# Un rapide tour historique de l'Intelligence Artificielle

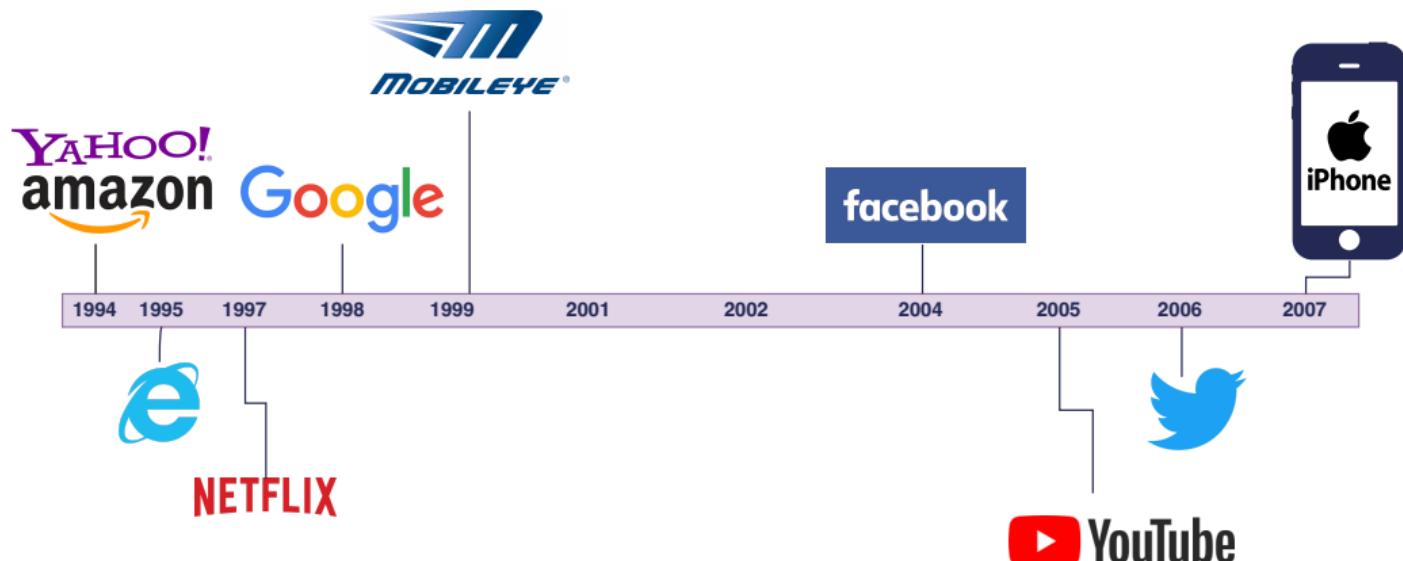
## Naissance de l'informatique... Et de l'Intelligence Artificielle





# Un rapide tour historique de l'Intelligence Artificielle

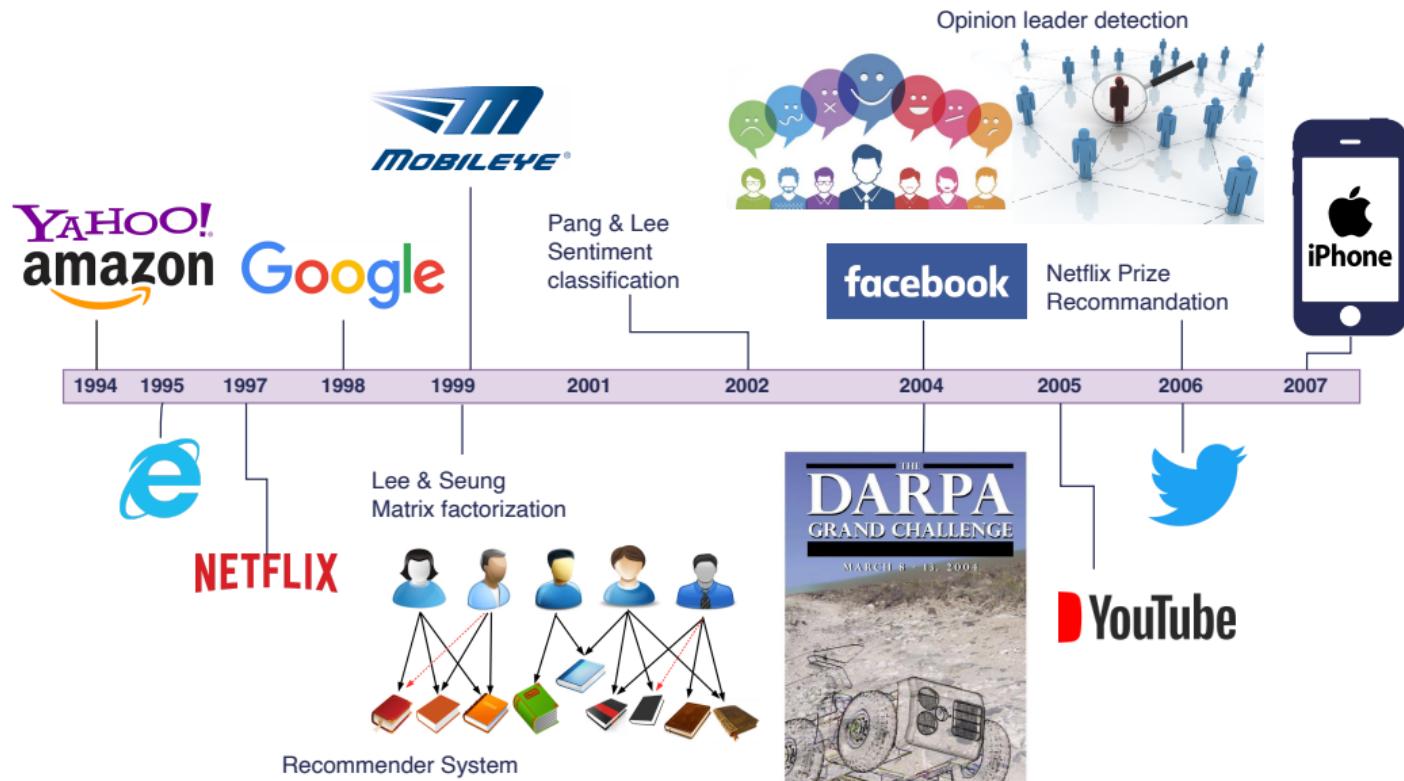
## Emergence (ou refondation) des GAFAM/GAMMA





# Un rapide tour historique de l'Intelligence Artificielle

## Emergence (ou refondation) des GAFAM/GAMMA



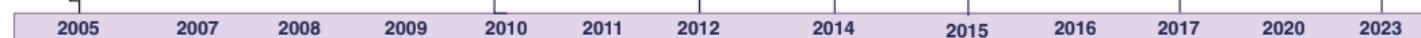


# Un rapide tour historique de l'Intelligence Artificielle

## Formation d'une vague de l'Intelligence Artificielle



Thrun:  
DARPA Gd Challenge  
victory



Y. Koren  
Challenge  
Netflix

kaggle



IBM Jeopardy win

amazon alexa

Google DeepMind  
Acquisition : \$400M

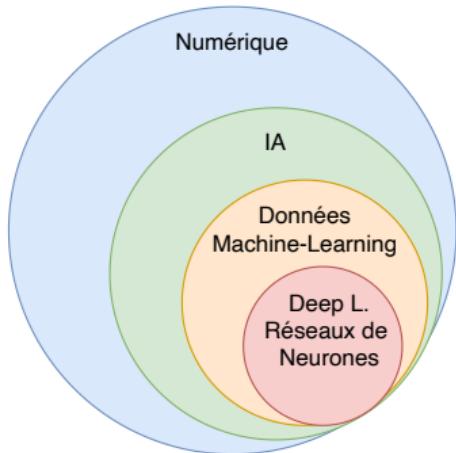


MOBILEYE®  
An intel company  
Acquisition : \$15B



OpenAI  
DALL·E 2

# Artificial Intelligence & Machine Learning



Input (X)	Output (Y)	Application
email	spam? (0/1)	spam filtering
audio	text transcript	speech recognition
English	Chinese	machine translation
ad, user info	click? (0/1)	online advertising
image, radar info	position of other cars	self-driving car
image of phone	defect? (0/1)	visual inspection

**IA** : programmes informatiques qui s'adonnent à des tâches qui sont, pour l'instant, accomplies de façon plus satisfaisante par des êtres humains car elles demandent des processus mentaux de haut niveau.

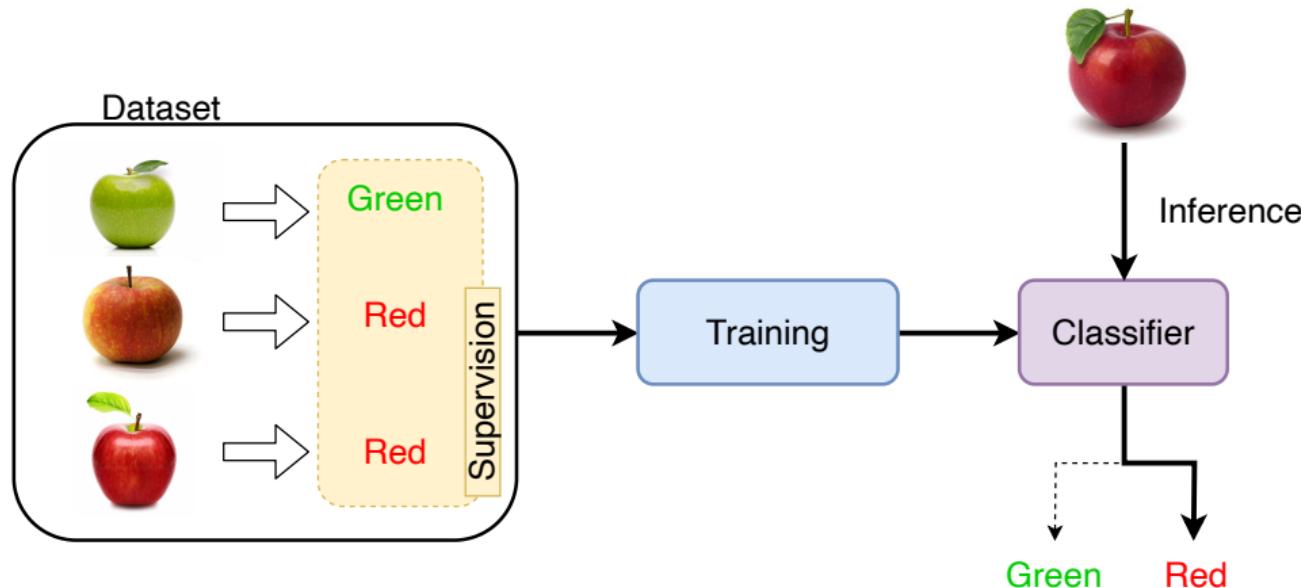
Marvin Lee Minsky, 1956

**N-AI (Narrow Artificial Intelligence)**, dédiée à une tâche unique

≠ **IA-G (IA Générale)**, qui remplace l'humain dans les systèmes complexes.

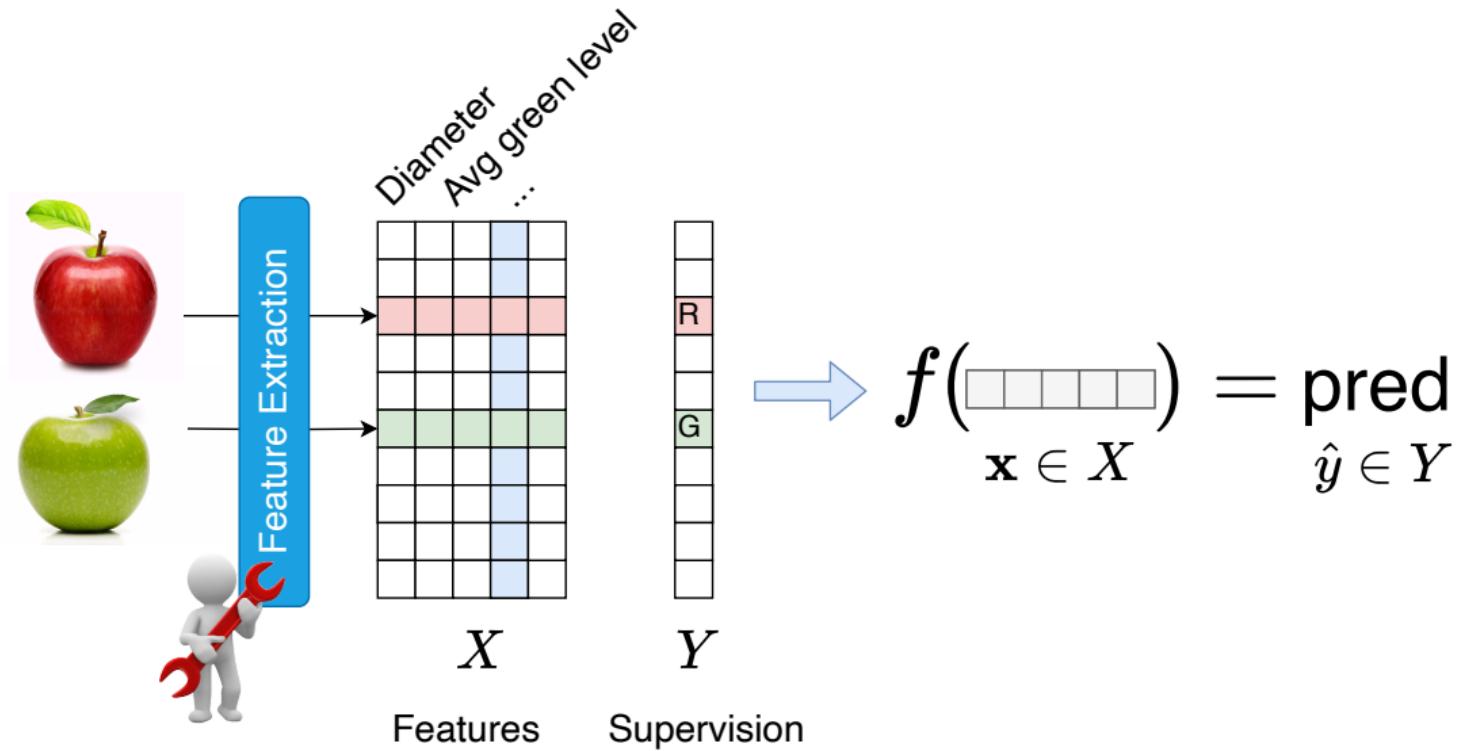
Andrew Ng, 2015

# Chaîne de Traitement Supervisé & Modèles



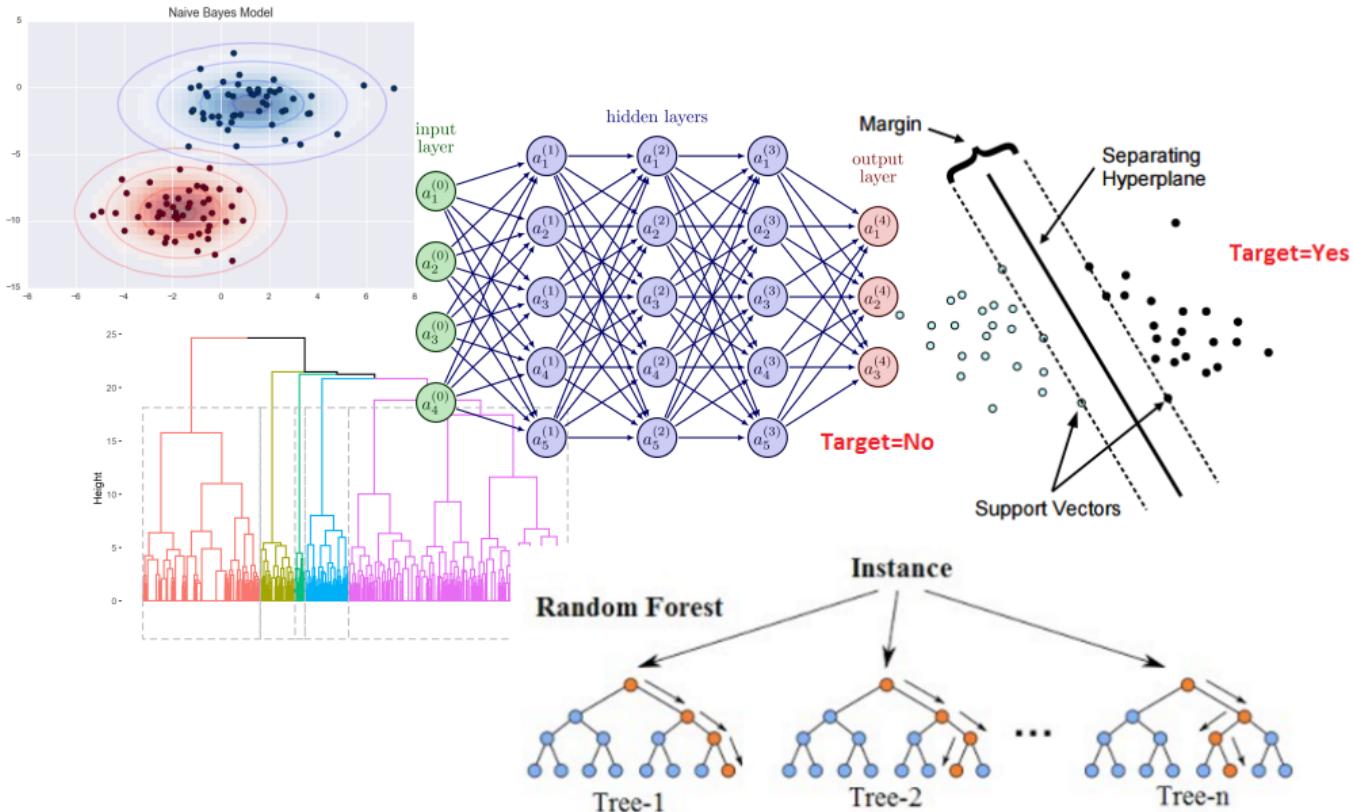
- Promesse = construire un modèle *uniquement* à partir d'observations

# Chaîne de Traitement Supervisé & Modèles



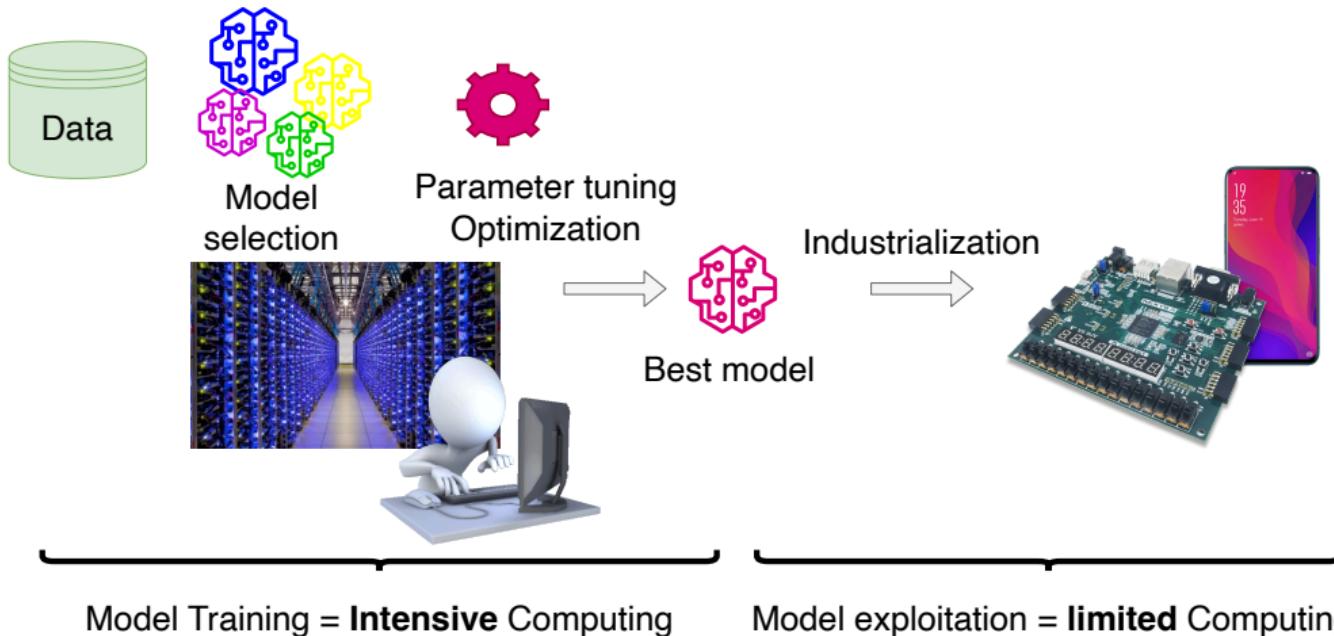


# Chaîne de Traitement Supervisé & Modèles



# Chaîne de Traitement Supervisé & Modèles

Différentes étapes en apprentissage automatique

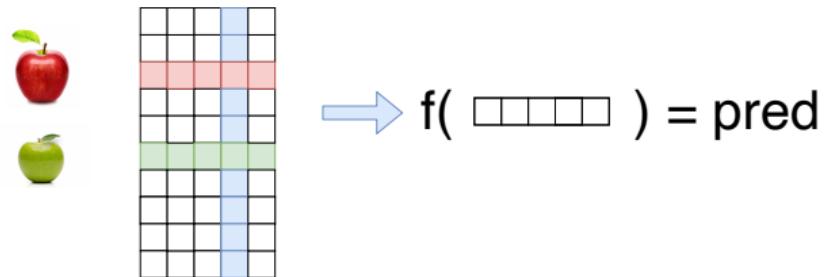


# DEEP-LEARNING & NLP<sup>★</sup>

[<sup>★</sup> NATURAL LANGUAGE PROCESSING =  
TRAITEMENT AUTOMATIQUE DE LA LANGUE NA-  
TURELLE]

# Des données tabulaires au texte

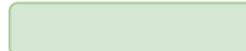
- Tableau de données
  - Dimension fixe
  - Valeurs continues



- Données Textuelles
  - Longueur variable
  - Valeurs discrètes

this new iPhone, what a marvel

An iPhone? What a scam!



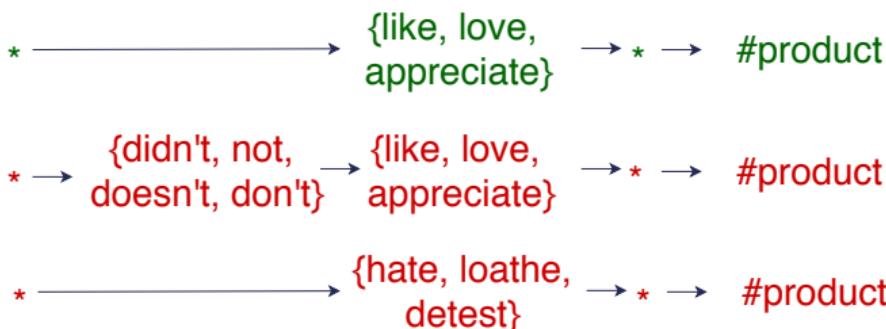


# IA + Texte : Trait. Auto. du Langage Naturel (TALN)

TALN = plus grande communauté scientifique en IA

## Linguistique [1960-2010]

### Systèmes à base de règles :



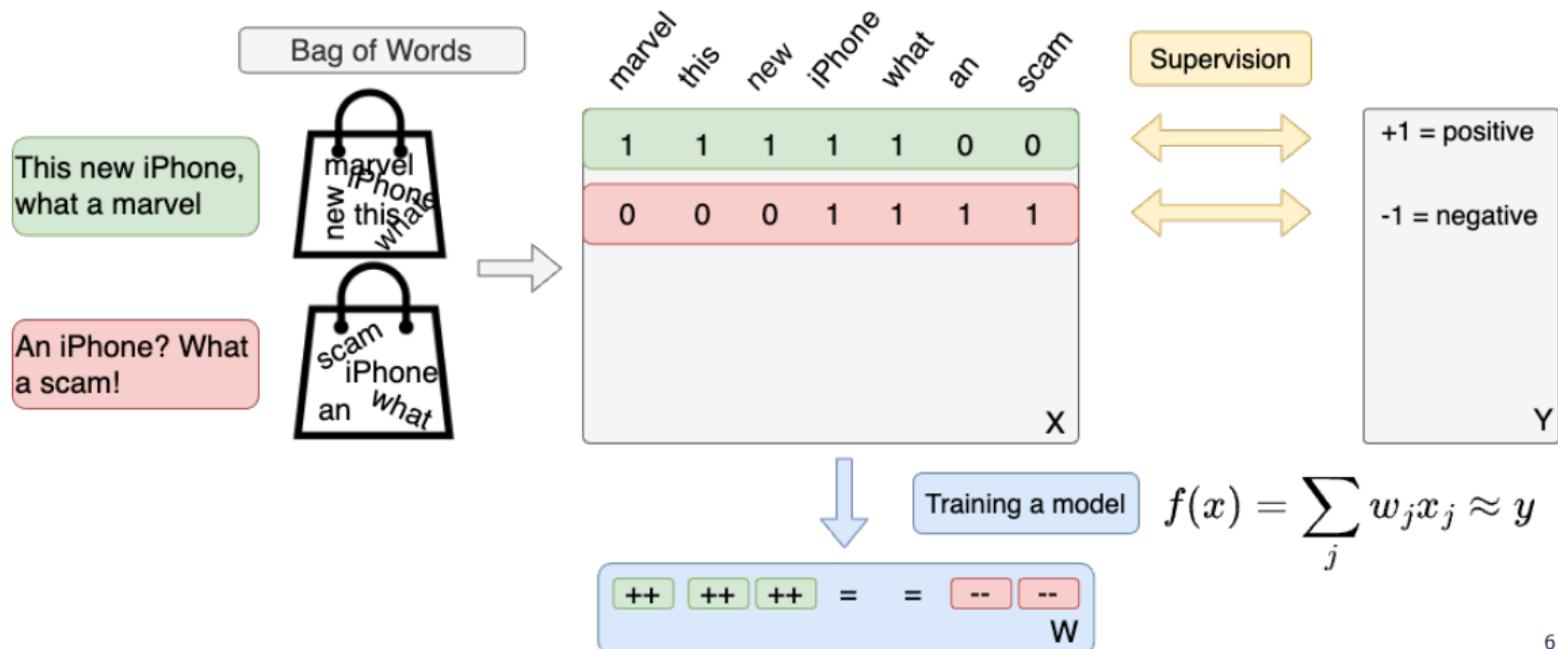
- Nécessite une expertise humaine
- Extraction de règles ⇔ données très propres
- Très grande précision
- Faible rappel
- Système interprétable



# IA + Texte : Trait. Auto. du Langage Naturel (TALN)

TALN = plus grande communauté scientifique en IA

## Apprentissage auto. [1990-2015]





# IA + Texte : Trait. Auto. du Langage Naturel (TALN)

TALN = plus grande communauté scientifique en IA

## Linguistique [1960-2010]

- Nécessite une expertise humaine
- Extraction de règles ⇔ données très propres
- + Système interprétable
- + Très grande précision
- Faible rappel

## Apprentissage auto. [1990-2015]

- Peu d'expertise nécessaire
- Extraction statistique ⇔ robuste aux données bruitées
- ≈ Système moins interprétable
- Moindre précision
- + Meilleur rappel

Précision = critère d'acceptation par l'industrie

→ Lien vers les métriques



# Apprentissage de représentations pour les données textuelles

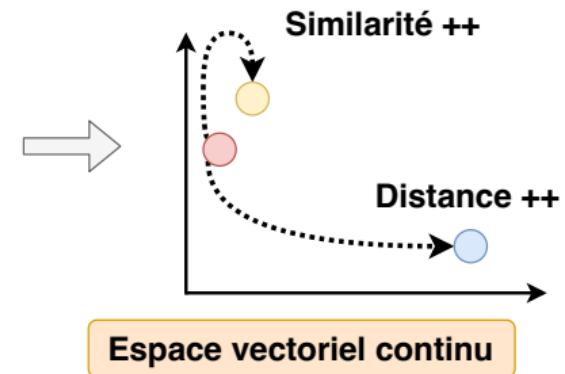
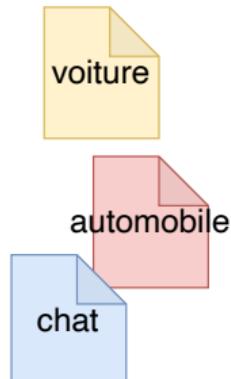
Du sac de mots aux représentations vectorielles

[2008, 2013, 2016]

## Corpus en sac de mots

	mot <sub>1</sub>	...	voiture	...	automobile	...	chat	...	mot <sub>D</sub>
d1	1	0	0						
d2	0	0	1						
d3	0	1	0						

Mêmes distances

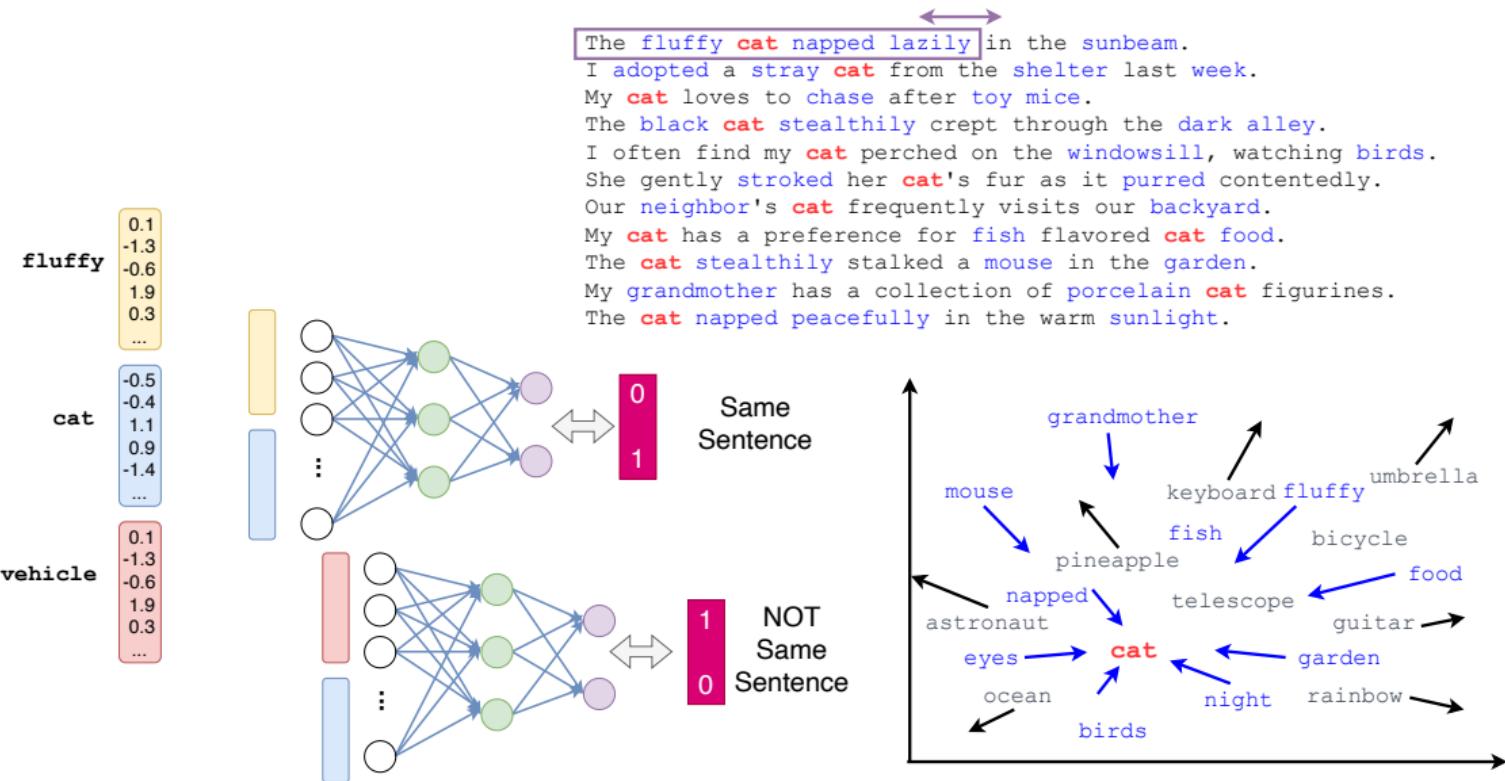




# Apprentissage de représentations pour les données textuelles

Du sac de mots aux représentations vectorielles

[2008, 2013, 2016]

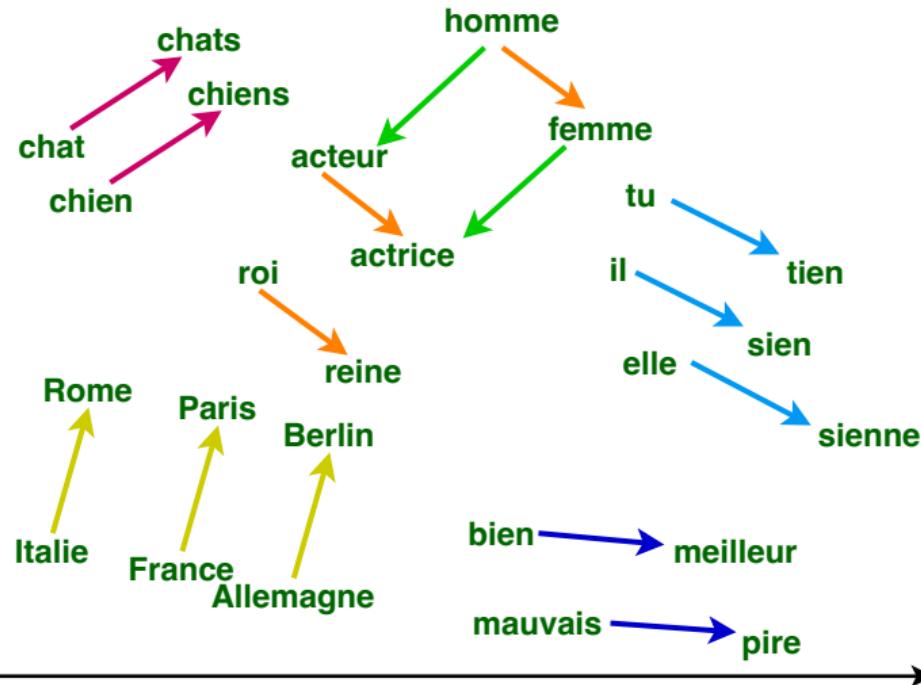




# Apprentissage de représentations pour les données textuelles

Du sac de mots aux représentations vectorielles

[2008, 2013, 2016]



- Espace sémantique : significations similaires  $\Leftrightarrow$  positions proches
- Espace structuré : régularités grammaticales, connaissances de base, ...



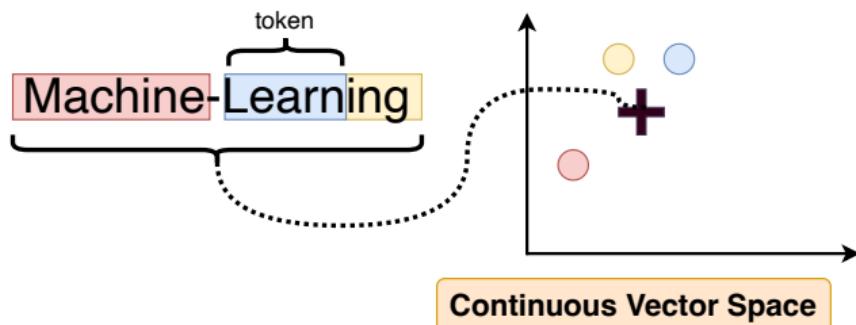
# Apprentissage de représentations pour les données textuelles

Du sac de mots aux représentations vectorielles

[2008, 2013, 2016]

## Des mots aux tokens

### Word Piece statistical split



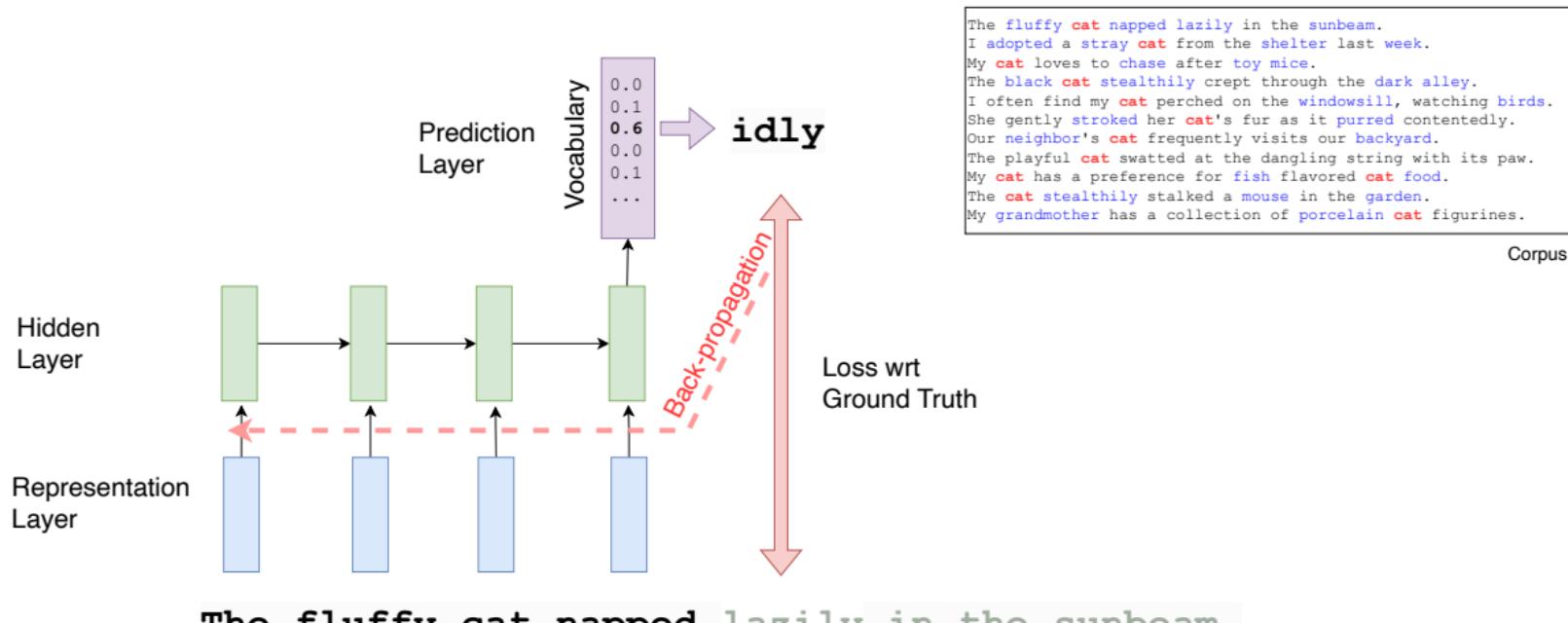
- Représentation des mots inconnus
- Adaptation aux domaines techniques
- Résistance aux fautes d'orthographe

Enriching word vectors with subword information. Bojanowski et al. TACL 2017.



# Agrégation des représentations de mots : vers l'IA générative

- Génération et représentation
- Nouvelle manière d'apprendre les positions des mots





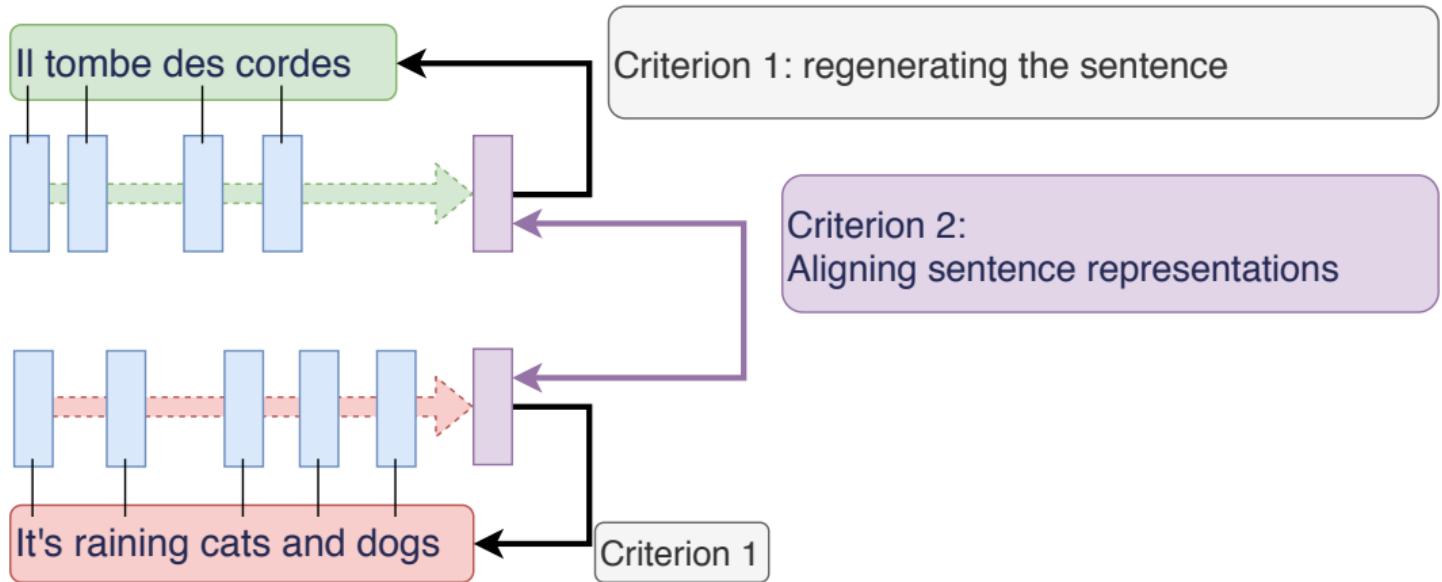
# Cas d'usage: la traduction machine



Au-delà de la traduction mot à mot: une représentation multilingue des phrases



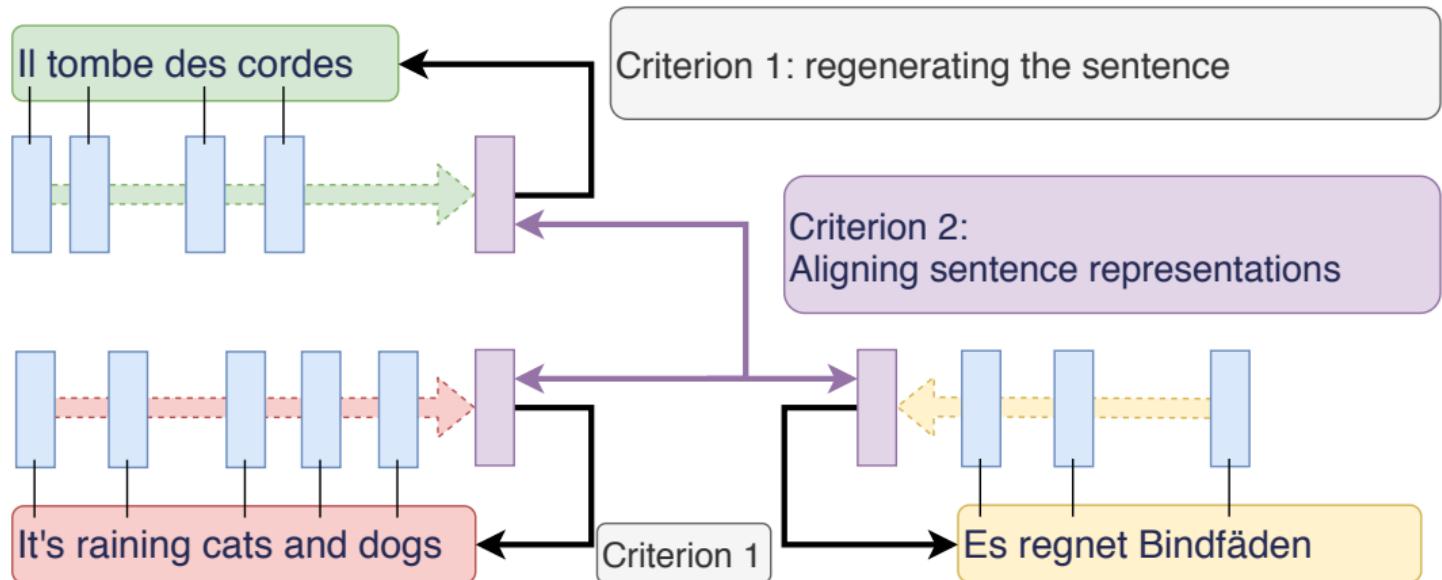
# Cas d'usage: la traduction machine



Au-delà de la traduction mot à mot: une représentation multilingue des phrases



# Cas d'usage: la traduction machine

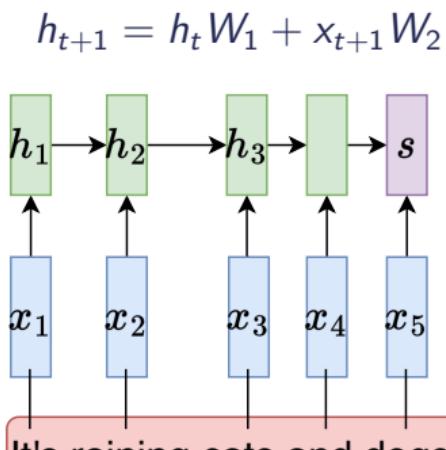


Au-delà de la traduction mot à mot: une représentation multilingue des phrases

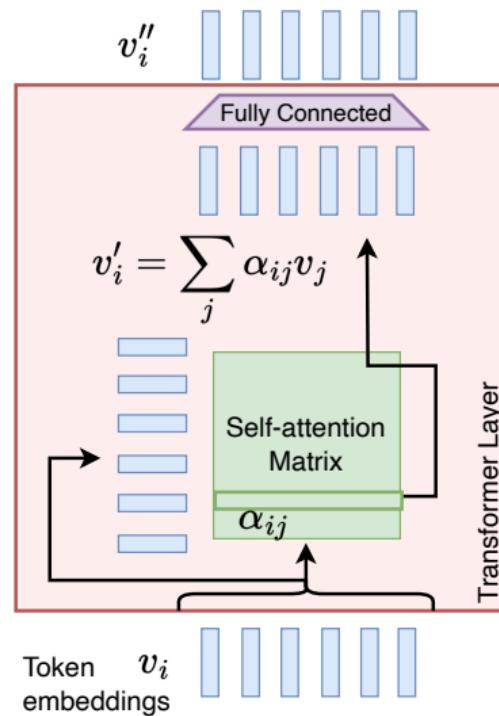


# Architecture Transformer : agrégation à l'état de l'art

Réseau de neurones récurrents :



Transformer :



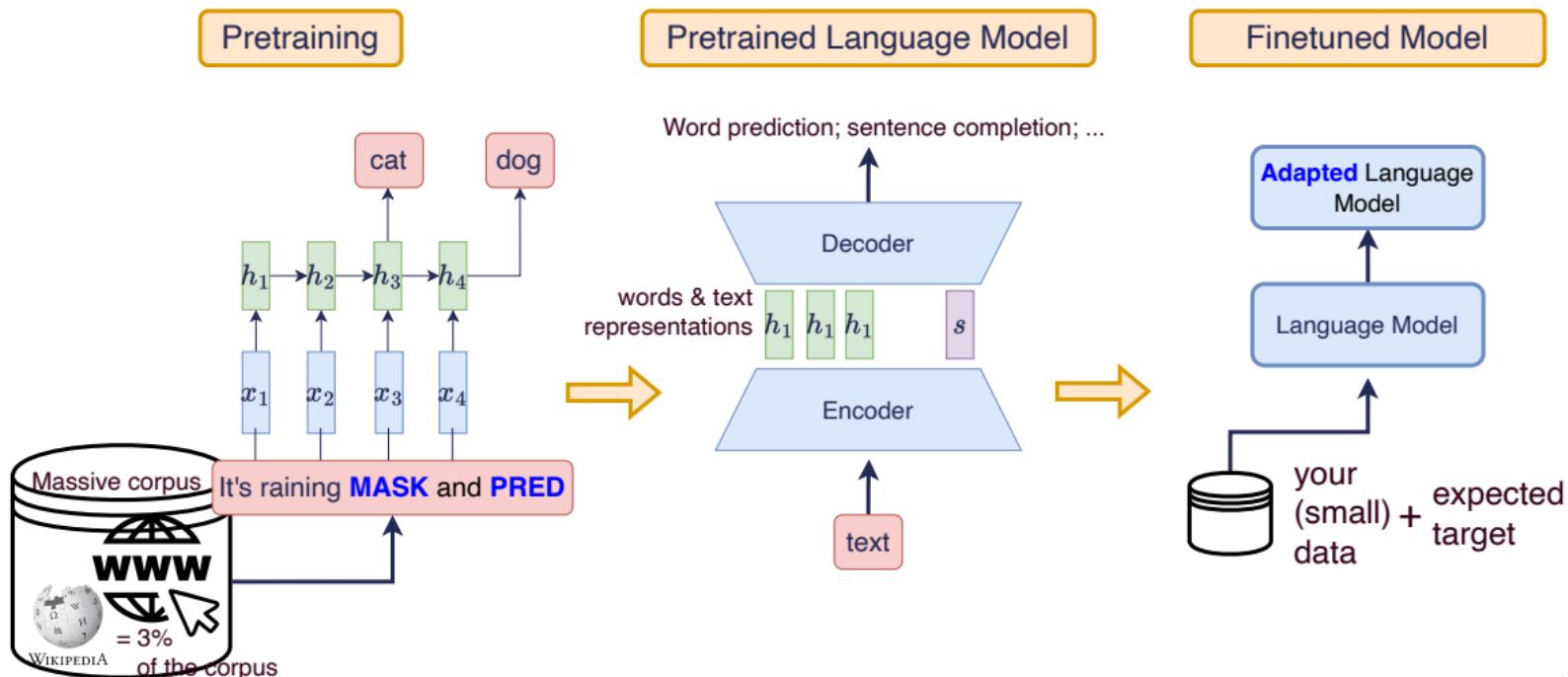
Attention is all you need, [Vaswani et al. NeurIPS 2017](#)

Sequence to Sequence Learning with Neural Networks, [Sutskever et al. NeurIPS 2014](#)



# Un nouveau paradigme de développement depuis 2015

- Jeu de données massif + architecture massive  $\Rightarrow$  coût d'entraînement +++
- Architecture pré-entraînée + zéro-shot / affinage



# CHATGPT

30 NOVEMBRE, 2022

1 MILLION D'UTILISATEURS EN 5 JOURS

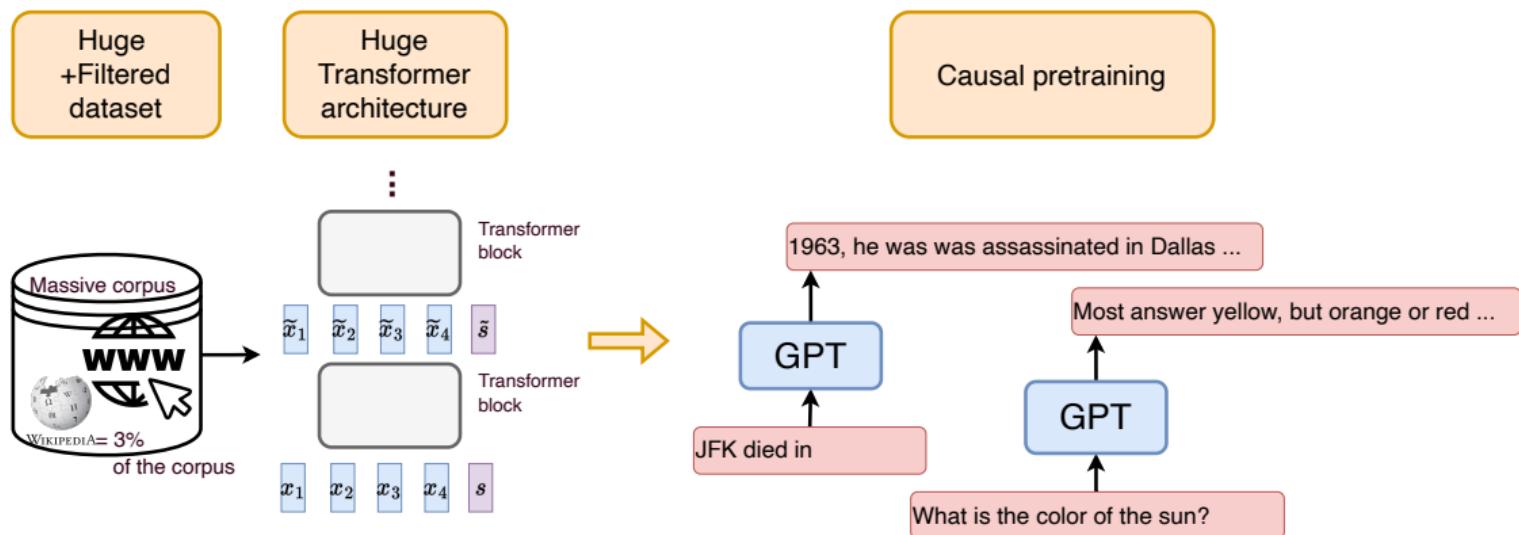
100 MILLIONS À LA FIN JANVIER 2023

1.16 MILLIARD EN MARS 2023



# Les ingrédients de chatGPT

## 0. Transformer + données massives (GPT)



- Grammaire : accord singulier/pluriel, concordance des temps
- Connaissances : entités, nom, lieux, dates, ...



# Les ingrédients de chatGPT

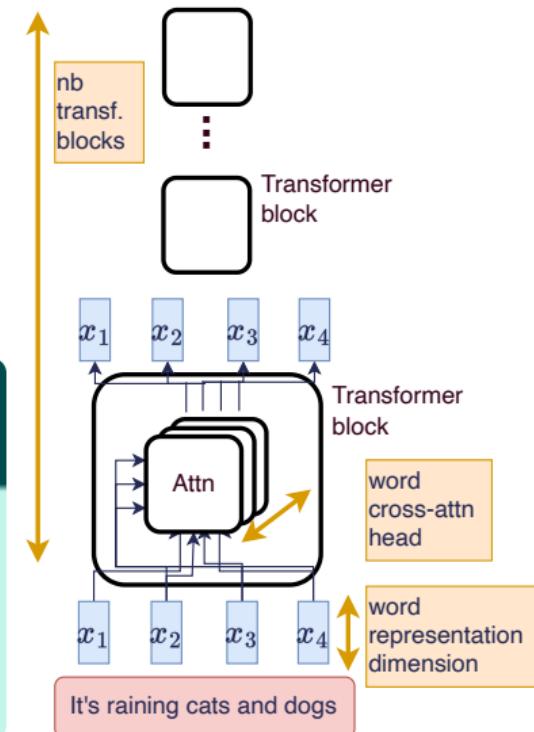
## 1. Plus, c'est mieux ! (GPT)

- + plus de mots en entrée [500  $\Rightarrow$  2k, 32k, 100k]
- + plus de dim. dans l'espace des mots [500-2k  $\Rightarrow$  12k]
- + plus de têtes d'attention [12  $\Rightarrow$  96]
- + plus de blocs/couches [5-12  $\Rightarrow$  96]

**175 milliards** de paramètres...

Qu'est-ce que cela signifie ?

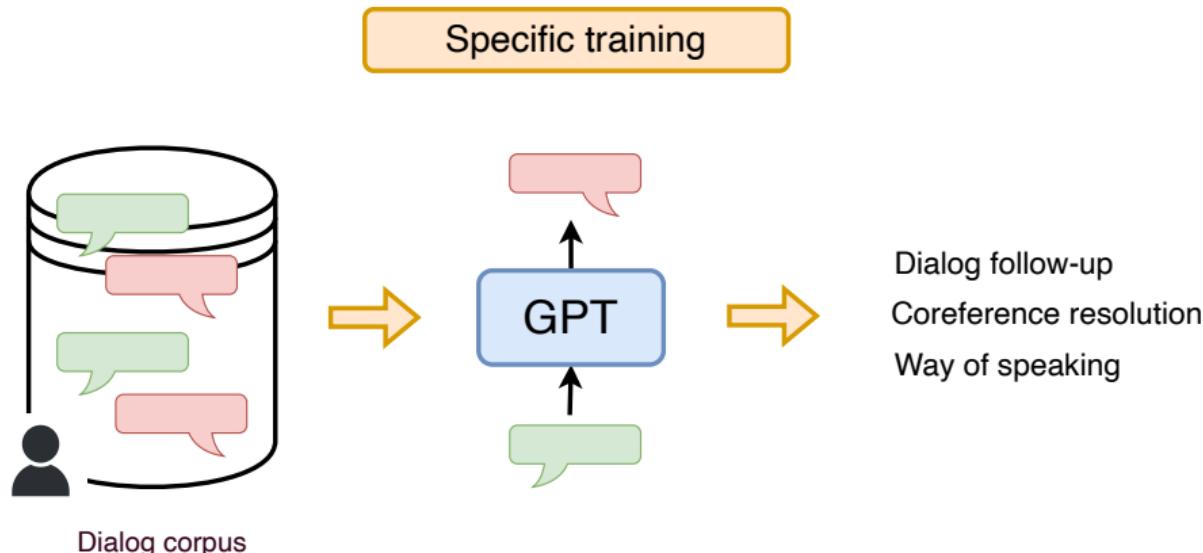
- $1,75 \cdot 10^{11} \Rightarrow 300 \text{ Go} + 100 \text{ Go}$  (stockage pour l'inférence)  $\approx 400 \text{ Go}$
- GPU NVidia A100 = 80 Go de mémoire (=20k€)
- Coût de (1) entraînement : 4,6 millions €





# Les ingrédients de chatGPT

## 2. Suivi du dialogue



- **Données très propres** Données générées/validées/classées par des humains



# Les ingrédients de chatGPT

## 3. Ajustement fin sur des tâches de raisonnement ( $\pm$ ) complexes

### Instruction finetuning

Please answer the following question.

What is the boiling point of Nitrogen?

### Chain-of-thought finetuning

Answer the following question by reasoning step-by-step.

The cafeteria had 23 apples. If they used 20 for lunch and bought 6 more, how many apples do they have?

-320.4F

The cafeteria had 23 apples originally. They used 20 to make lunch. So they had  $23 - 20 = 3$ . They bought 6 more apples, so they have  $3 + 6 = 9$ .

Language model

### Multi-task instruction finetuning (1.8K tasks)

### Inference: generalization to unseen tasks

Q: Can Geoffrey Hinton have a conversation with George Washington?

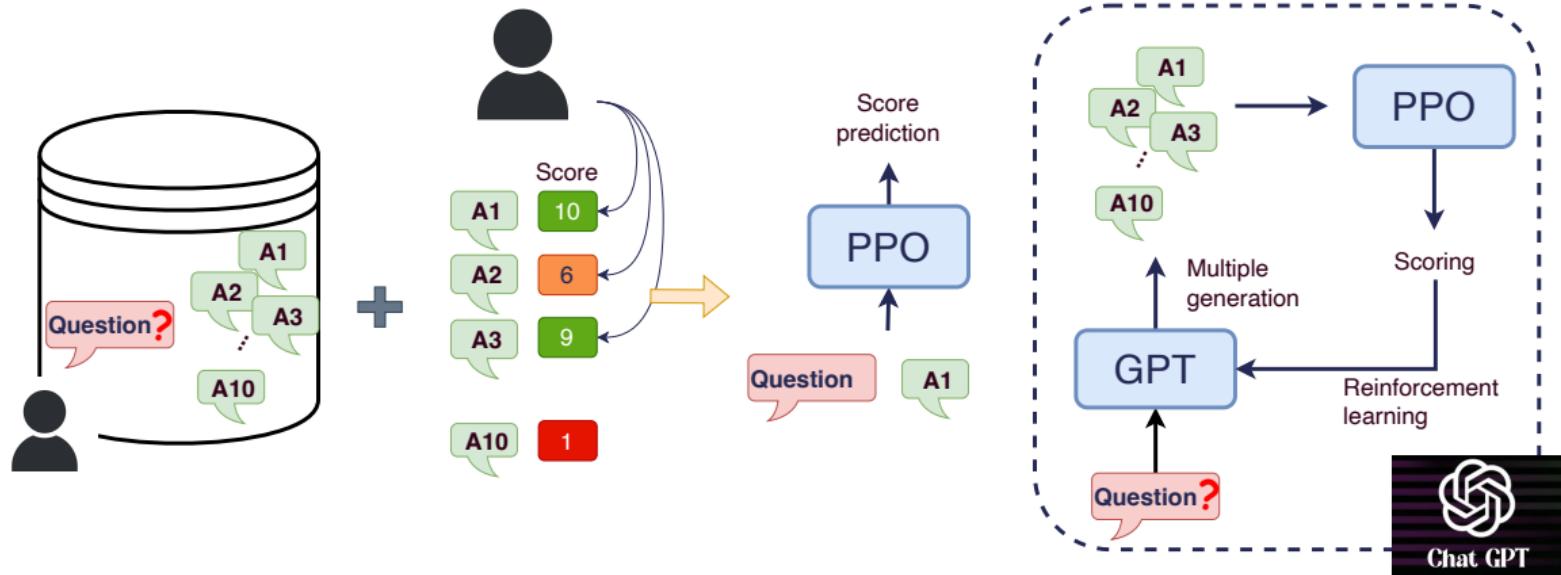
Give the rationale before answering.

Geoffrey Hinton is a British-Canadian computer scientist born in 1947. George Washington died in 1799. Thus, they could not have had a conversation together. So the answer is "no".



# Les ingrédients de chatGPT

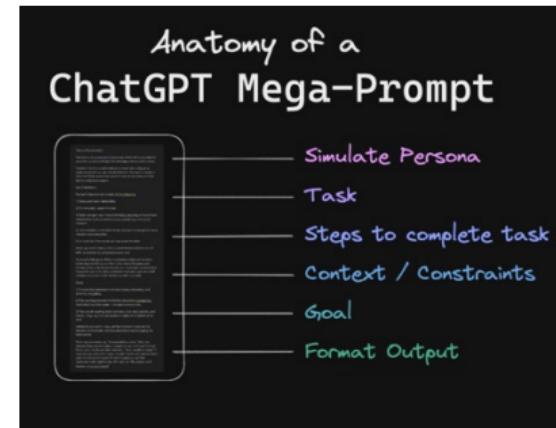
## 4. Instructions + classement des réponses



- Base de données créée par des humains
- Amélioration des réponses
- ... Aussi un moyen d'éviter les sujets sensibles = censure

# Utilisation de chatGPT & Prompting

- Interroger chatGPT = compétence ⇒ *prompting*
  - Bonne question : ... *en détail*, ... *étape par étape*
  - Spécifier un nombre d'élts, ex. : *3 qualités pour ...*
  - Donner du contexte : *cellule* pour un biologiste / assistant juridique
- Ne pas s'arrêter à la première question
  - Détaillez certains points
  - Réorientez la recherche
  - Dialoguez
- Reformulation
  - *Explain like I'm 5*, comme dans un article scientifique, en mode pote, ...
  - Résumer, développer
  - Ajouter des erreurs (!)



<https://chatgptprompts.guru/what-makes-a-good-chatgpt-prompt/>

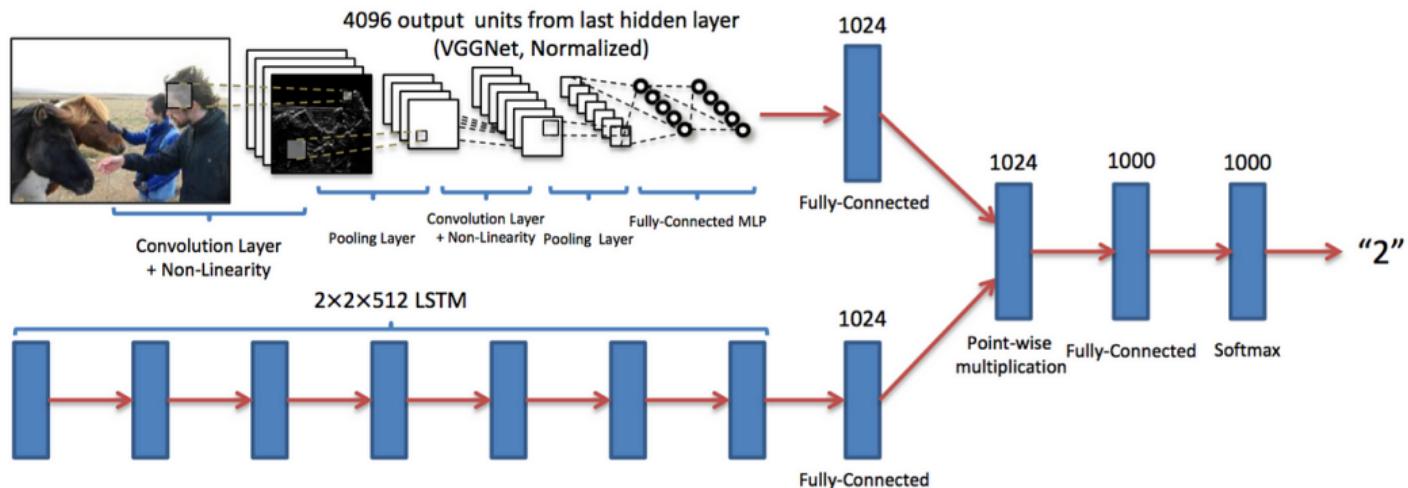
⇒ Besoin de **pratique** [1 à 2 heures], échanges avec collègues



# GPT-4 & Multimodalité

**Fusionner** info. texte + image. **Apprendre** l'information conjointement

*Exemple du VQA : Visual Question Answering (questions visuelles)*



"How many horses are in this image?"

⇒ Rétropropager l'erreur ⇒ modifier les repr. des mots + l'analyse d'image



VQA : Visual Question Answering, arXiv, 2016 , A. Agrawal et al.



# Au bout du compte

## Statistical Modeling of Texts

Texts splitting = tokens

Large Language Models (LLMs), such as GPT-3 and GPT-4, utilize a process called tokenization. Tokenization involves breaking down text into smaller units, known as tokens, which the model can process and understand. These tokens can range from individual characters to entire words or even larger chunks, depending on the model. For GPT-3 and GPT-4, a Byte Pair Encoding (BPE) tokenizer is used. BPE is a subword tok

Iterative Process

Dictionary	Large entire For units ... can may ...	0.02 0.01 0.00 0.00 0.00 0.09 ...
------------	---	---

0.02 0.01 0.00 0.00 0.00 0.09 ...	0.30
---	------

Starting text

Language Model

Token forecasting

# USAGES DES LLM

• Génération de textes

• Traduction

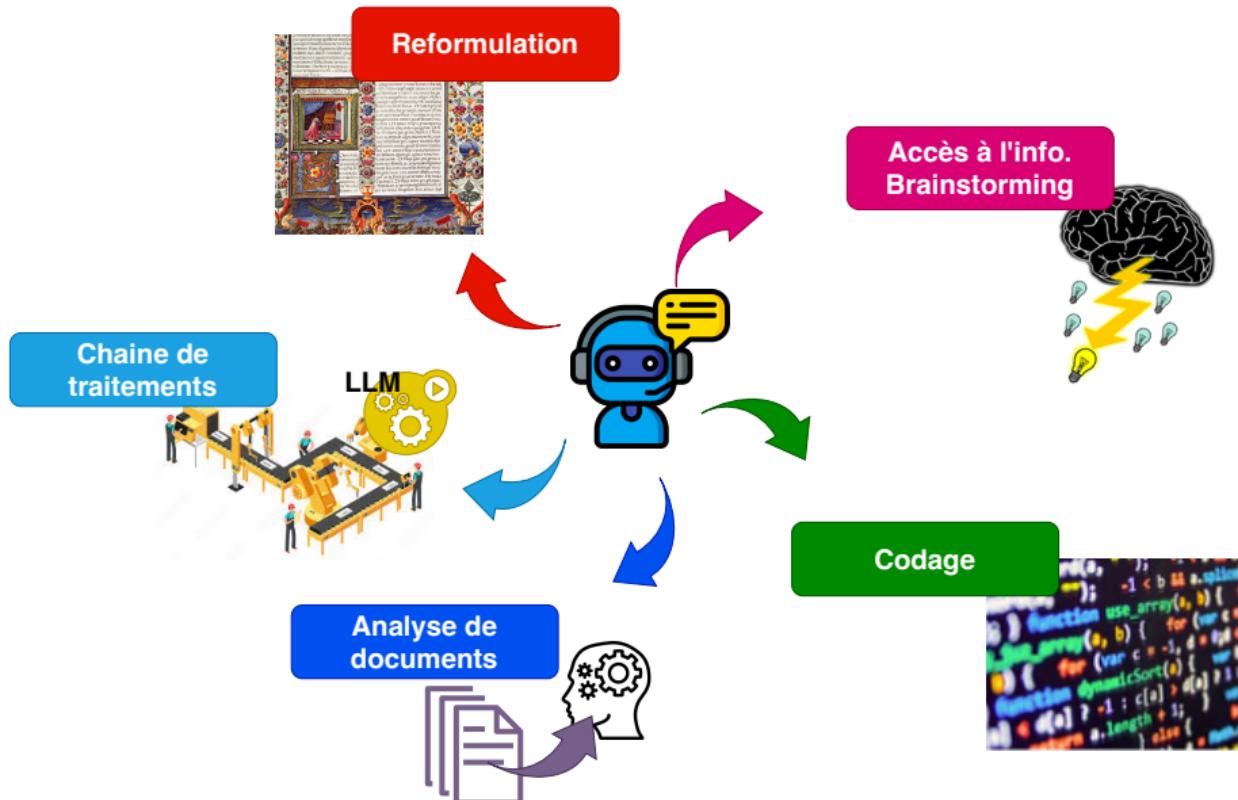
• Résolution de problèmes

• Analyse d'images

• Gestion de bases de données



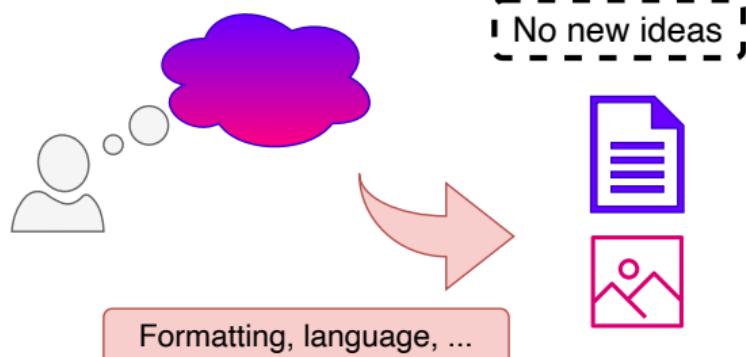
# Usages clés en 5 images





# (1) Mise en forme de l'information

## Outil de mise en forme



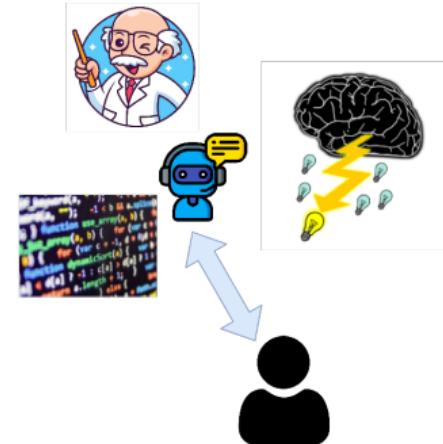
- Assistant personnel
  - Lettres types, lettres de recommandation, de motivation, lettres de résiliation
  - Traductions
- Comptes-rendus de réunion
  - Mise en forme des notes
- Rédaction d'articles scientifiques
  - Idées de rédaction, en français, en anglais
- Analyse de documents
  - Extraction d'informations, questions-réponses, ...

⇒ Aucune information nouvelle, juste de la rédaction, du nettoyage, ...



## (2) Brainstorming

- Trouver l'inspiration [syndrome de la page blanche]
  - Organiser rapidement ses idées
  - Éviter les oublis / renforcer la confiance
  - Rechercher de manière ciblée, adaptée à ses besoins
- ⇒ Réponses impressionnantes, parfois incomplètes ou partiellement fausses... Mais souvent utiles



*3 articles de référence sur l'usage des transformers dans les systèmes de recommandation*

*À quoi sert la loi de Poisson log-normale ?*

*Proposer 10 sections pour un cours sur les Transformers en IA*

- Dans quels domaines les LLMs sont-ils fiables ?
- Quels sont les risques pour les sources d'information primaires ?
- Quels risques sociétaux pour l'information ?

### (3) Programmation : Outils différents, niveaux différents

- Proposer des solutions à des exercices
  - Apprendre à programmer ou s'y remettre
    - Nouveaux langages, nouvelles approches (ML ?)
    - Profiter des explications...  
Mais comment gérer les erreurs ?
  - Aide à l'utilisation d'une bibliothèque [*prise en main*]
  - Programmer plus rapidement

## Mais comment gérer les erreurs ?



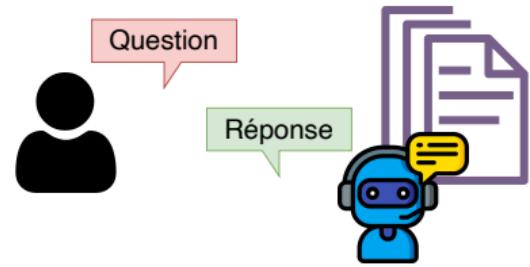
- Quid des droits d'auteur ?
    - Quel impact sur le traitement futur du code ?
  - Comment adapter les méthodes pédagogiques ?
  - Combien d'appels sont nécessaires pour compléter du code ?
    - Qu'en est-il de l'empreinte carbone ?
  - Quel est le risque de propagation d'erreurs ?

```
1 import datetime
2
3 def parse_expenses(expenses_string):
4     """Parse the list of expenses and return the list of triples (date,
5     Ignore lines starting with #.
6     Parse the date using datetime.
7     Example expenses_string:
8         2016-01-02 -34.01 USD
9         2016-01-03 2.59 DKK
10        2016-01-03 -2.72 EUR
11    """
12    expenses = []
13    for line in expenses_string.splitlines():
14        if line.startswith("#"):
```



## (4) Analyse de documents

- Résumer des documents / articles
- Dialoguer avec une base documentaire
- Aide à la rédaction de revues critiques
- FAQ, services de support interne en entreprise
- Veille technologique
- Génération de quiz à partir de notes de cours



Wi-Fi NotebookLM

**Think Smarter,  
Not Harder**

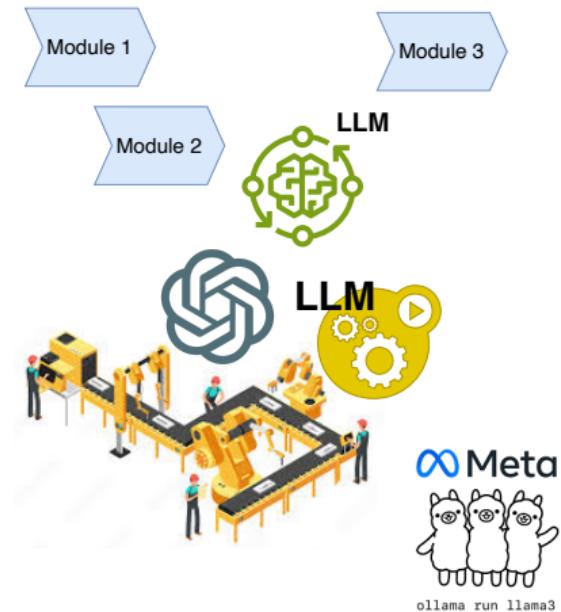
Try NotebookLM

- Les articles seront-ils encore lus dans le futur ?
  - Devra-t-on rendre nos articles résistants à NotebookLM ?
  - Comment gagner du temps tout en restant honnête et éthique ?



# (5) LLM dans une chaîne de production / IA agentique

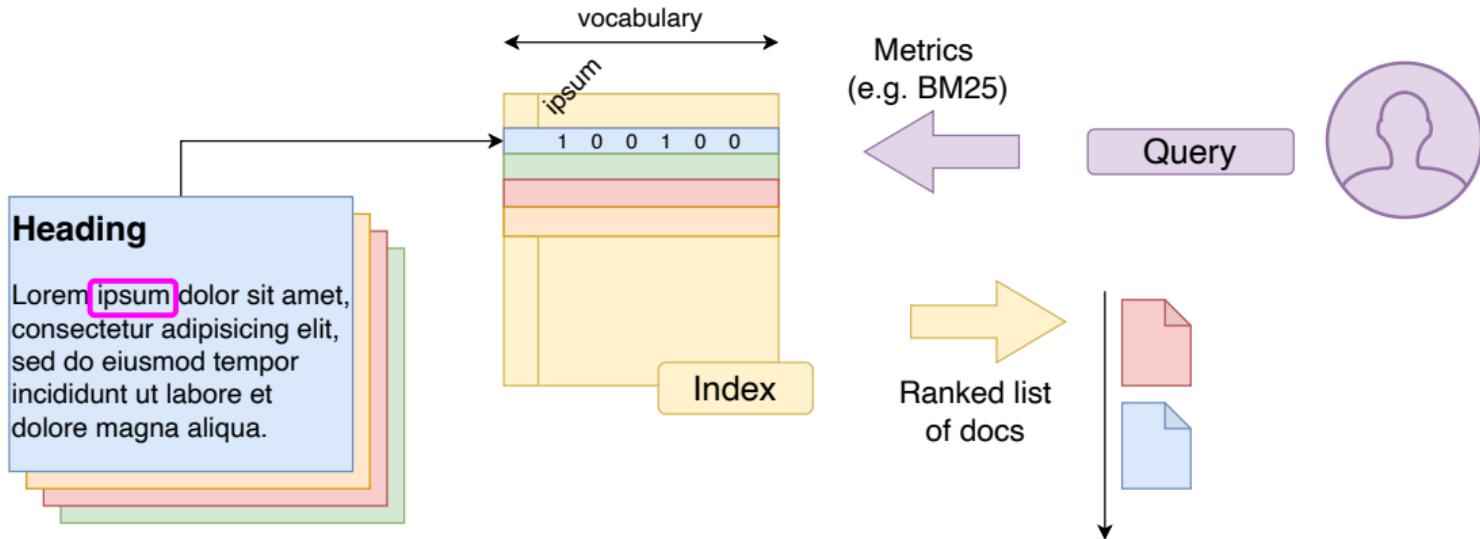
- Faire tourner un LLM en local
  - Extraire des connaissances
  - Trier des documents / générer des résumés
  - Générer des exemples pour entraîner un modèle  
[Professeur/élève – distillation]
  - Générer des variantes d'exemples ↗↗  
augmenter la taille du jeu de données  
[Augmentation de données]
- ⇒ Intégrer le LLM dans une chaîne de traitement  
= peu/pas de supervision = **IA agentique**



- Peut-on entraîner des modèles sur des données générées ?
- Quel est le coût ? (\$ + CO<sub>2</sub>) Besoin de GPU ?
- Quelle est la qualité des modèles à poids ouverts ?



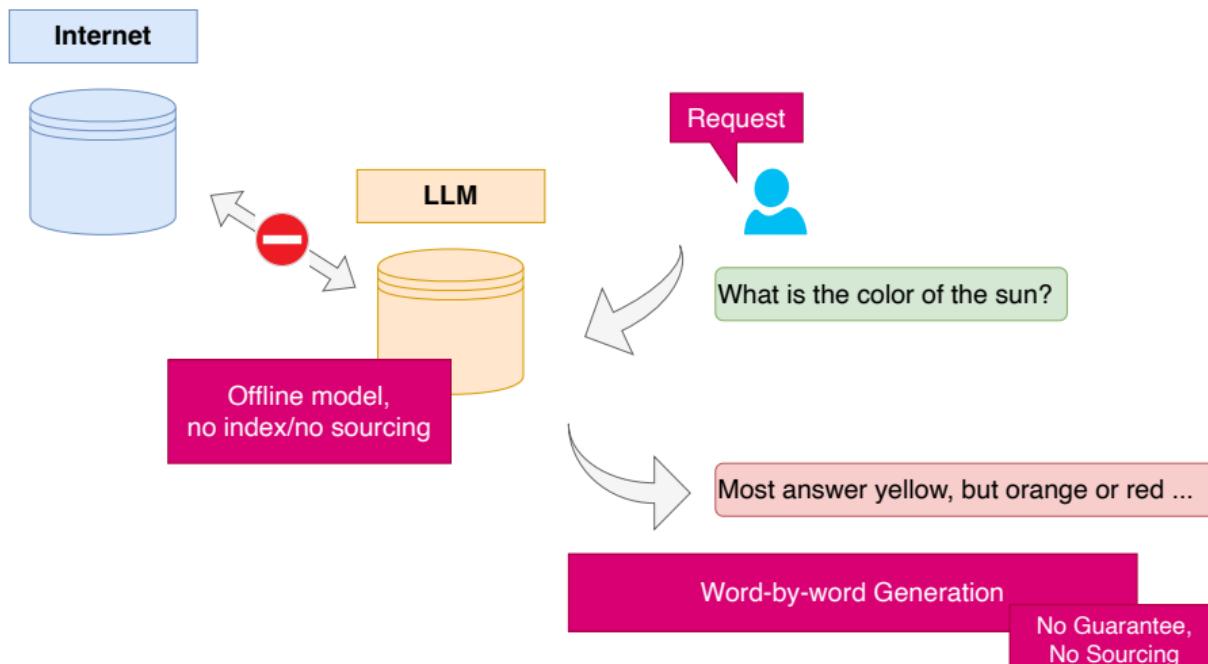
# Modèle de langue vs Recherche d'information





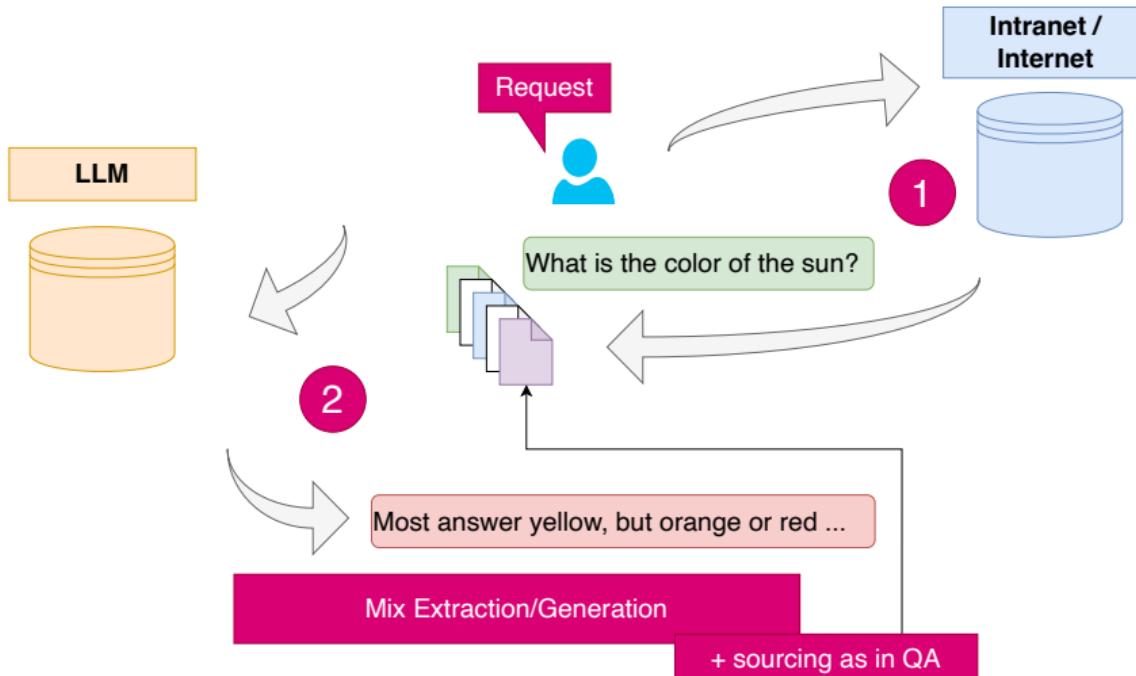
# LLMs $\Rightarrow$ RAG : mémoire vs extraction d'information

- Poser des questions à ChatGPT... Une utilisation surprenante !
- Mais est-ce raisonnable ?  
[Vraie question ouverte (!)]





# LLMs $\Rightarrow$ RAG : mémoire vs extraction d'information

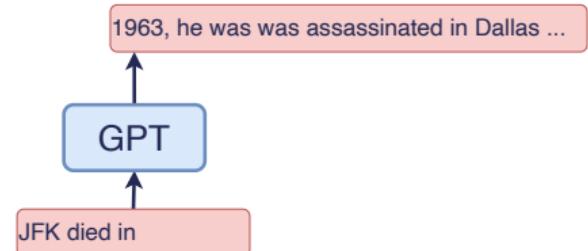


- RAG : génération augmentée par récupération
- Limite (actuelle) sur la taille d'entrée (2k, 32k, 200k tokens)

# MACHINE LEARNING LIMITES TECHNIQUES

# chatGPT et la relation à la vérité

- 1 **Vraisemblance** = grammaire, accords, concordance des temps, enchaînements logiques...  
⇒ Connaissances répétées
- 2 Prédit le mot le plus **plausible**...  
⇒ produit des **hallucinations**
- 3 Fonctionnement en **hors ligne**
- 4 chatGPT ≠ **graphes de connaissances**
- 5 Réponses brillantes...  
Et erreurs absurdes !  
+ on ne peut pas prévoir les erreurs



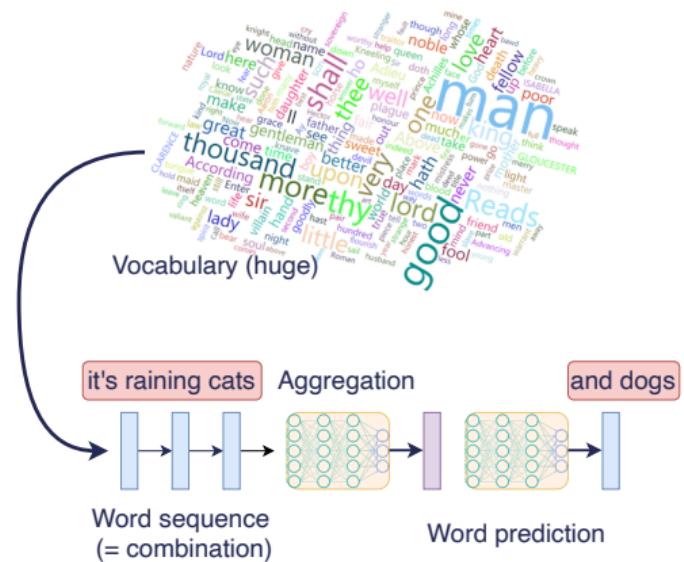
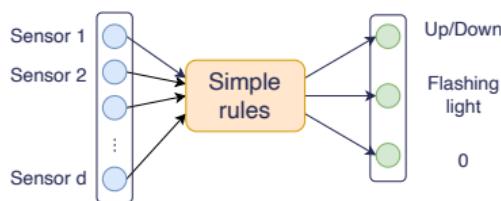
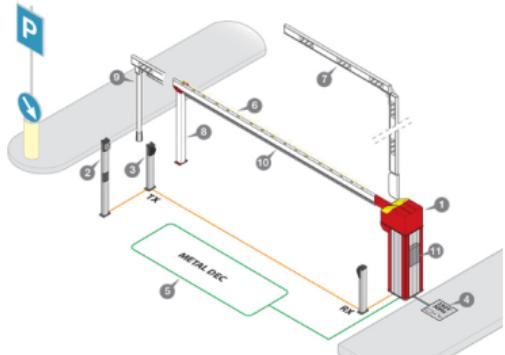
## Exemple : produire une bibliographie

Can you give me a short bibliography on VAE for Time Series?

- recurrents.
2. "Variational Recurrent Autoencoders" par Chung et al. (2016) - Cette étude introduit une approche VAE pour la modélisation de séquences temporelles en utilisant des réseaux de neurones récurrents et une méthode de maximisation de la vraisemblance pour la phase d'entraînement. [Variational recurrent auto-encoders](#)  
O. Fabius, JR Van Amerongen - arXiv preprint arXiv:1412.6581, 2014 - arxiv.org  
In this paper we propose a model that combines the strengths of RNNs and SGVB: the Variational Recurrent Auto-Encoder (VRAE). Such a model can be used for efficient, large scale ...
  3. "Generative Modeling for Time Bao et al. (2017) - Cette étude i pour la modélisation de séries t profonds, y compris les VAE.
  4. "Deep Variational Bayes Filters: Unsupervised Learning of State Space Models from



# Stabilité, explicabilité... et complexité



- Système simple
- Tests exhaustifs des entrées/sorties
- **Prévisible et explicable**

- Grande dimension
- Combinaisons non-linéaires complexes
- **Non prévisible et non explicable**



# Stabilité, explicabilité... et complexité

## Interprétabilité vs explication a posteriori

Réseaux de neurones = **non interprétables** (presque toujours)

*trop de combinaisons pour être anticipées*

Réseaux de neurones = **explicables a posteriori** (presque toujours)



[Accident Uber, 2018]

- Système simple
- Tests exhaustifs des entrées/sorties
- **Prévisible et explicable**

- Grande dimension
- Combinaisons non-linéaires complexes
- **Non prévisible et non explicable**

# Pas de magie, beaucoup de lacunes

Beaucoup de succès aussi... mais :

⇒ Le LLM (ne) fait (que) ce pour quoi il a été entraîné

En retrait sur:

- Calculs simples  
(multiplication, division)
- Génération de noms d'animaux en  $n$  syllabes (en cours)
- Jouer aux échecs
- Suivre un raisonnement causal  
(complexe)
- ...

## ATARI 2600 SCORES STUNNING VICTORY OVER CHATGPT



**WHEN YOU UNDERESTIMATE A 1977 CHESS ENGINE...  
AND IT HUMBLES YOU IN FRONT OF THE WHOLE INTERNET**

# (PRINCIPAUX) RISQUES DES MODÈLES DE LANGUE



# Typologie des risques en IA/NLP (L. Weidinger)



## Discrimination, exclusion and toxicity

Harms that arise from the language model producing discriminatory and exclusionary speech.



## Information hazards

Harms that arise from the language model leaking or inferring true sensitive information.



## Misinformation harms

Harms that arise from the language model producing false or misleading information.



## Malicious uses

Harms that arise from actors using the language model to intentionally cause harm.



## Human-computer interaction harms

Harms that arise from users overly trusting the language model, or treating it as human-like.

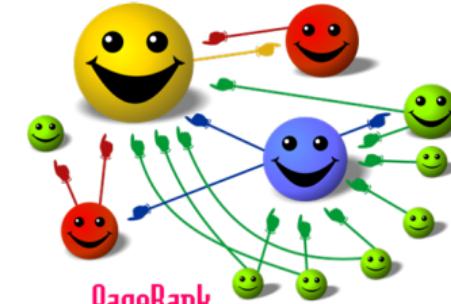


## Automation, access and environmental harms

Harms that arise from environmental or downstream economic impacts of the language model.

# Accès à l'information

- Accès à des informations dangereuses/interdites
  - +Données personnelles
  - Droit à l'oubli numérique
- Autorités informationnelles
  - Nature : inconsciemment, image = vérité
  - Source : presse, réseaux sociaux, ...
  - Volume : nombre de variantes, citations (pagerank)
- Génération de texte : harcèlement...
- Anthropomorphisation de l'algorithme
  - Distinguer humain et machine





# Apprentissage automatique & biais



Oreilles pointues,  
moustaches, texture de poils  
=  
Chat



Homme blanc, +40ans,  
costume  
=  
Cadre supérieur

Biais dans les données ⇒ biais dans les réponses

L'apprentissage automatique repose sur l'extraction de biais statistiques...

⇒ Lutter contre les biais = ajustement manuel de l'algorithme



# Apprentissage automatique & biais

≡ Google Traduction



Stéréotypes tirés de *Pleated Jeans*

- Choix du genre
- Couleur de peau
- Posture
- ...

Biais dans les données ⇒ biais dans les réponses

L'apprentissage automatique repose sur l'extraction de biais statistiques...

⇒ Lutter contre les biais = ajustement manuel de l'algorithme



# Correction des biais & ligne éditoriale

## Correction des biais :

- Sélection de données spécifiques, rééquilibrage
- Censure de certaines informations
- Censure des résultats de l'algorithme

⇒ Travail éditorial...

Réalisé par qui ?

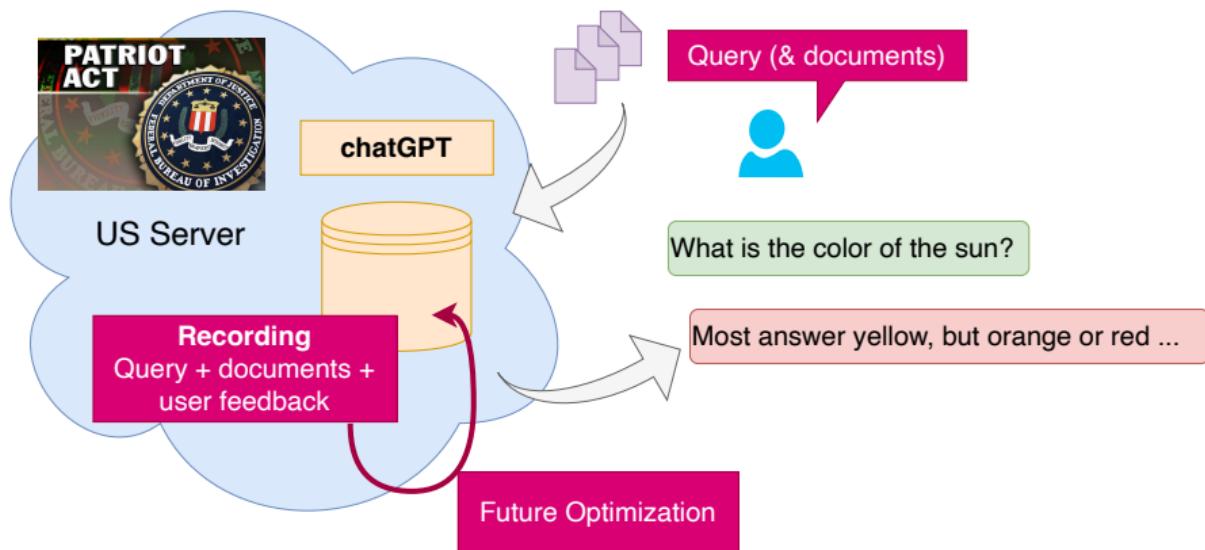
- Experts du domaine / cahier des charges
- Ingénieurs, lors de la conception de l'algorithme
- Groupe éthique, lors de la validation des résultats
- Équipe communication / réponses aux utilisateurs

⇒ Quelle légitimité ? Quelle transparence ? Quelle efficacité ?





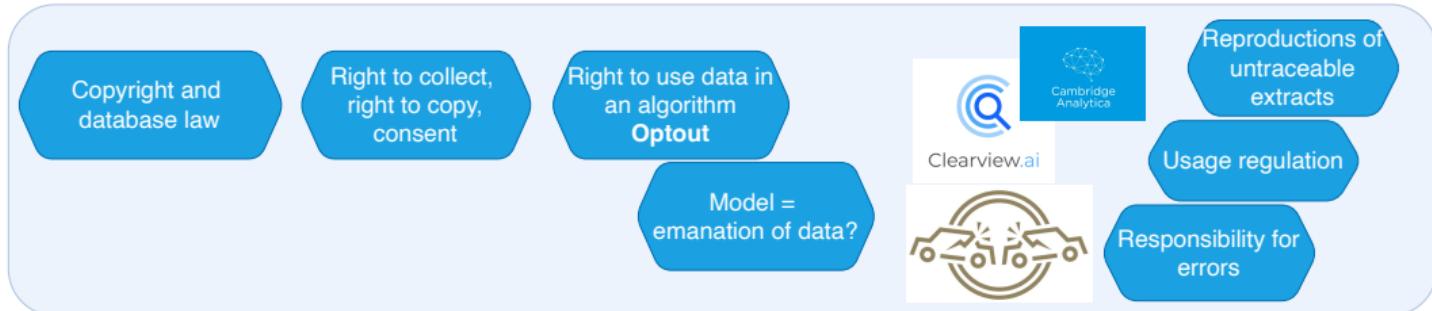
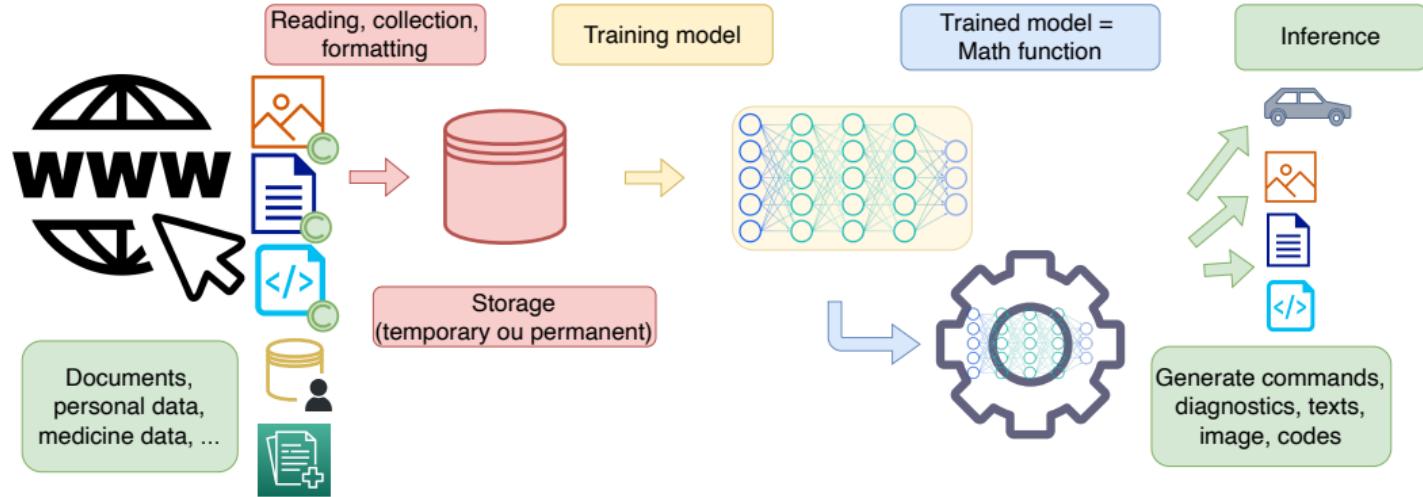
# Fuites de données



- Transmission de données sensibles
- Exploitation des données par OpenAI (ou d'autres)
- Fuite de données dans de futurs modèles



# Risques/Questions juridiques





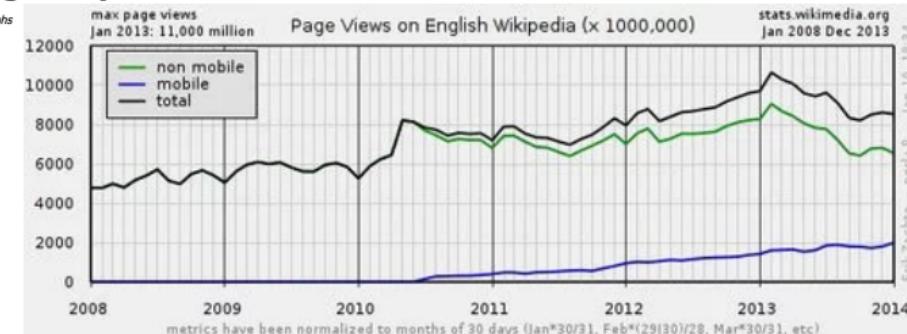
# Questions économiques

- Financement/Publicité ⇔ **visites** des internautes
- Google Knowledge Graph (2012) ⇒ moins de visites, donc moins de revenus
- chatGPT = encodage de l'information du web... ⇒ encore moins de visites ?

⇒ **Quel modèle économique / sources d'information avec chatGPT ?**

## Google's Knowledge Graph Boxes: killing Wikipedia?

by Gregory Kohs



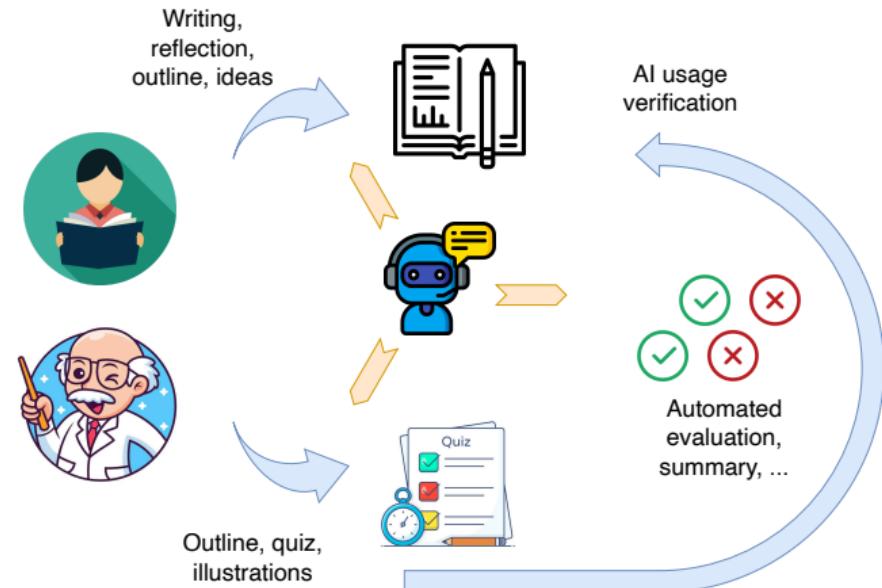
⇒ Qui bénéficie du retour d'information ? [StackOverflow]



# Risques liés à la généralisation de l'IA

L'IA partout =  
perte de sens ?

- Dans le domaine éducatif
- Transposition aux RH
- Aux systèmes de financement par projet



# RÉFLECTION PÉDAGOGIQUE



# Deux questions distinctes

- 1 Enseigner **avec** l'IA
- 2 Enseigner l'**IA**

⇒ Dans tous les cas, il faut connaître l'IA !



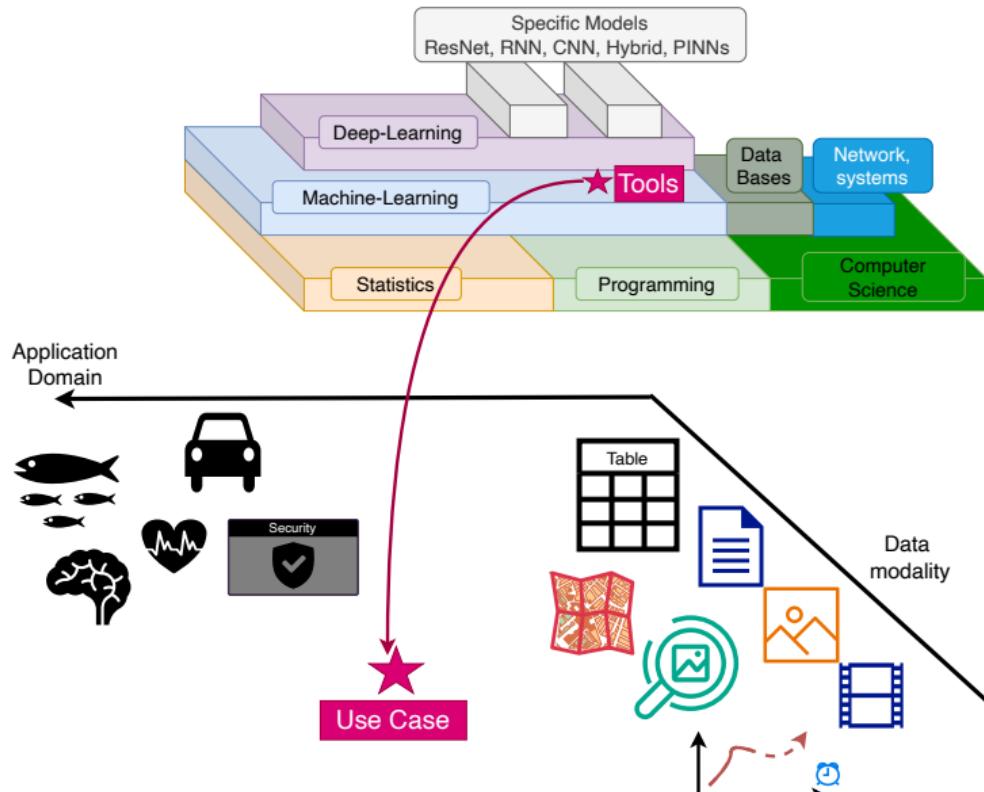
## Enseignement **avec** l'IA

- Programmation / Statistiques
- App. automatique : modèles + protocoles
- Apprentissage profond
- Modalités : images, textes, séries temporelles
- Approfondissements théoriques : optimisation, intervalles de confiance, convergence, ...

- Nouvelles opportunités
- Nouveaux risques
- Nouvelles contraintes



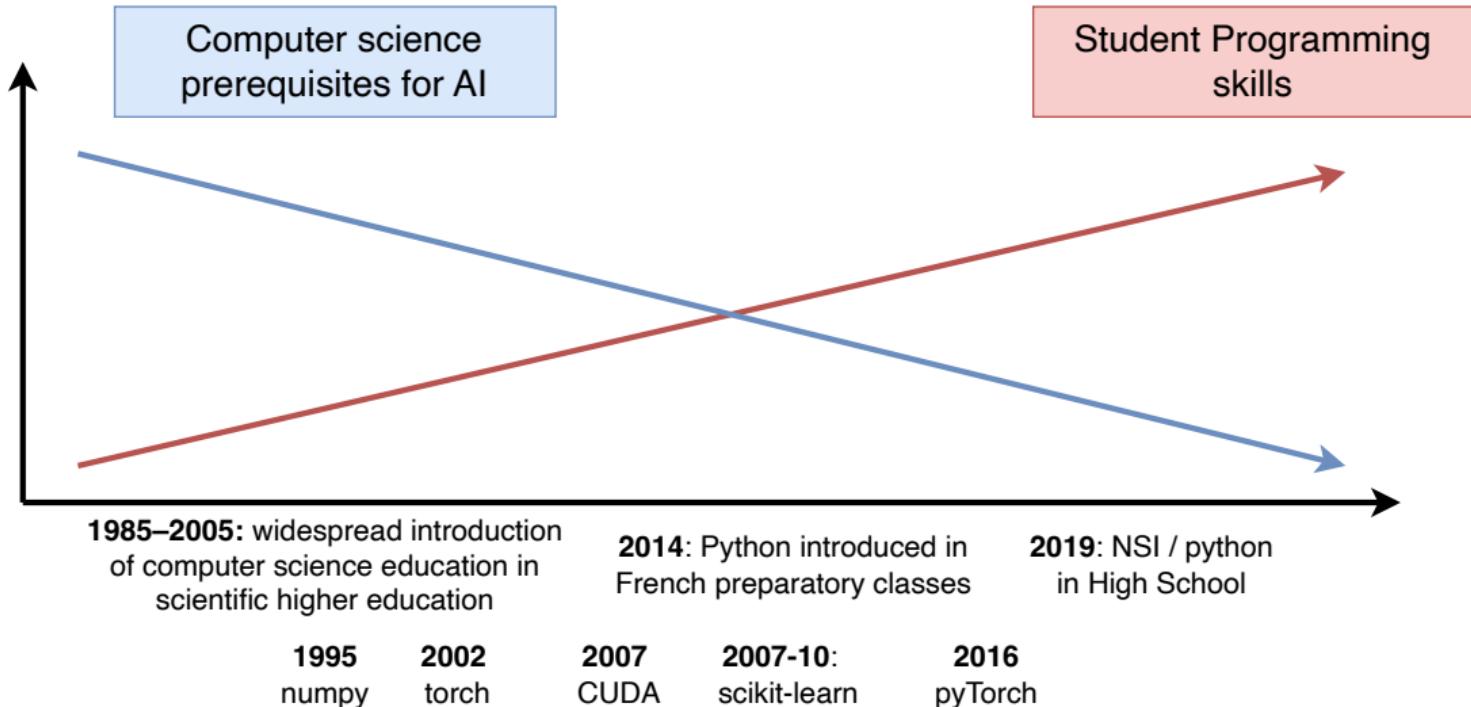
# Enseigner l'IA



- Différents niveaux d'accès (sensibilisation, usage d'outils, développement)
- Différents types de données
- Différents domaines d'application

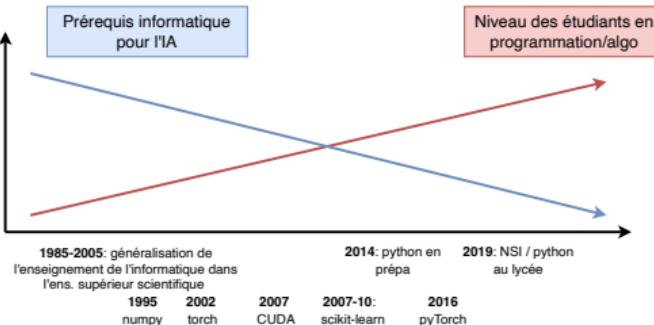


# Accès à l'IA : à la croisée des chemins





# Accès à l'IA : à la croisée des chemins



1



2



3



Trois niveaux d'accès à l'IA :

- 1 Exploiter un chatbot... de façon optimale et responsable**
- 2 Utiliser des outils, manipuler des données**
- 3 Développer des outils**

# LLMs : un outil pédagogique ?

- LLM = mémoire (partielle) d'Internet
  - Maître de la reformulation
  - Capacité à comprendre / traduire / générer du code
  - Répond à de nombreux types de questions
- ⇒ Oui, il répondra à beaucoup de choses...  
en faisant régulièrement des **erreurs**
- ⇒ Très utile globalement pour les exercices de base
- ⇒ Produit une **grande quantité de texte**, souvent bien écrit



Paradigme de la  
calculatrice :

*si une machine existe,  
pourquoi apprendre les  
tables de multiplication ?*



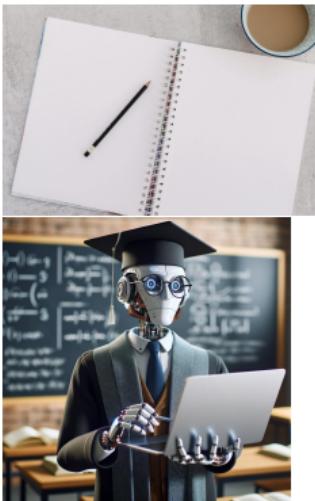
# Questions pédagogiques : le bon, la brute et le truand

## Quels usages vertueux ?

Etudiant: un professeur disponible 24h/24

- **Oser** coder/écrire – briser la peur de la page blanche
- Poser des questions, **vérifier** des solutions, ne pas hésiter à poser des "questions bêtes"
- Améliorer ses révisions en demandant des exercices et fiches de synthèse...
- **Améliorer** des lettres de motivation / équité sociale (?)

Enseignant: un assistant disponible 24h/24!



## Questions ?

- Peut-on résister à la tentation de demander la réponse ?
- Comment départager des lettres générées par ChatGPT ?



# Questions pédagogiques : le bon, la brute et le truand

## Quels usages frauduleux ?

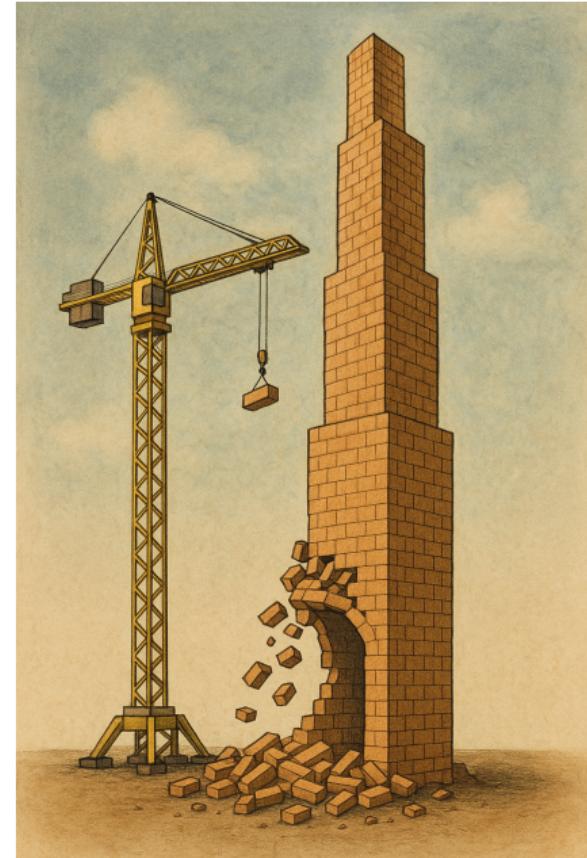
- Rédiger une dissertation / un exposé
  - Générer des réponses d'examen : code, histoire, langues étrangères, ...
  - Produire une analyse de document
- 
- Culture générale (les LLM sont compétents)
  - Travaux axés sur la forme (LLM très compétents)
  - Analyse de documents fournis ( LLM assez compétents)





# Quels défis pédagogiques ?

- Redéfinir les **priorités éducatives**, discipline par discipline, comme pour Wikipédia / calculatrice / ...
  - Accepter la **réduction/disparition de certaines compétences**
- Former les étudiants aux LLM, + en interdire régulièrement l'usage
  - Examens sur papier, soutenances avec questions individuelles, ...
- Apprendre à **reconnaître un contenu généré par LLM**, utiliser les outils de détection
  - Mécanismes de sanctions adaptés





# Détection de textes générés par *chatGPT*

Plagiat Def 2 – Saved to my Mac

Design Layout References Mailings Review View Tell me Comments

L'externalité fait référence au fait qu'une activité économique d'un agent peut avoir un impact sur d'autres personnes sans qu'il y ait de compensation financière. Cela peut être bénéfique pour les autres, comme offrir une utilité gratuitement, ou nuisible, comme causer des dommages écosystémiques, économiques ou qui ne sont pas compensés par le coût, mais

Tout cocher Trier les documents par Date de dépôt 1 - 2 sur 2

Plagiat Def 2 #4483eb 07/01/2023 19:18 par vous | 122 mots | 19,47 ko | Plus d'infos 0% Rapport

Plagiat Def 1 #f90ff3 07/01/2023 19:16 par vous | 135 mots | 16,78 ko | Plus d'infos 100% Rapport

L'externalité caractérise le fait qu'un agent économique crée, par son activité, un effet externe en procurant à autrui, sans contrepartie monétaire, une utilité ou un avantage de façon gratuite, ou au contraire une nuisance, un dommage sans compensation (coût social, coût écosystémique, pertes de ressources pas, peu, difficilement, lentement ou coûteusement renouvelables...).

De la sorte, un agent économique se trouve en position d'influer consciemment ou inconsciemment sur la situation d'autres agents, sans que ceux-ci soient parties prenantes à la décision : ces derniers ne sont pas forcément informés et/ou n'ont pas été consultés et ne participent pas à la gestion de ses conséquences par le fait qu'ils ne reçoivent (si l'influence est négative), ni ne paient (si l'influence est positive) aucune compensation.

En résumé : « Tout coûte mais tout ne se paie pas »

## Reformulation par chatGPT

## Définition de Wikipedia

Crédit: S.  
Pajak



# Détection de textes générés par *chatGPT*

GPTZero

Detect AI Plagiarism. Accurately



Chat GPT



AI Detector

- **Classifieur** de texte (comme pour tout auteur)
  - Détection de biais dans le choix des mots / tournures
- Caractérisation de la **vraisemblance** du texte ([OpenAI](#), [GPTZero](#))
  - Phrases trop fluides, connecteurs logiques trop nombreux
  - Modèle de langage = statistique ⇒ comp. de distributions (**perplexité**)
- **$\delta$ -vraisemblance** sur des textes perturbés ([DetectGPT](#))

Détecteurs ⇒ taux de détection < 100%

- + Score de confiance dans la détection
- Dépend de la longueur du texte et des modifications apportées
- ≈ Peut signaler des extraits de Wikipédia (chatGPT = *perroquet stochastique*)
- ≈ Sur-détection sur les étudiants étrangers



# Déclin / évolution cognitive

Notre cerveau va évoluer avec ces nouveaux outils...

Quelle est la portée de ces transformations? Quelles en seront les conséquences?

- Les sciences de l'éducation et la psychologie l'avaient conjecturé...  
les sciences cognitives l'ont mesuré

