

# INTRODUCTION AUX INTELLIGENCES ARTIFICIELLES GÉNÉRATIVES

11 Janvier 2024

Séminaire MERIT, Institut des Systèmes Complexes

Vincent Guigue

[vincent.guigue@agroparistech.fr](mailto:vincent.guigue@agroparistech.fr)

<https://vguigue.github.io>

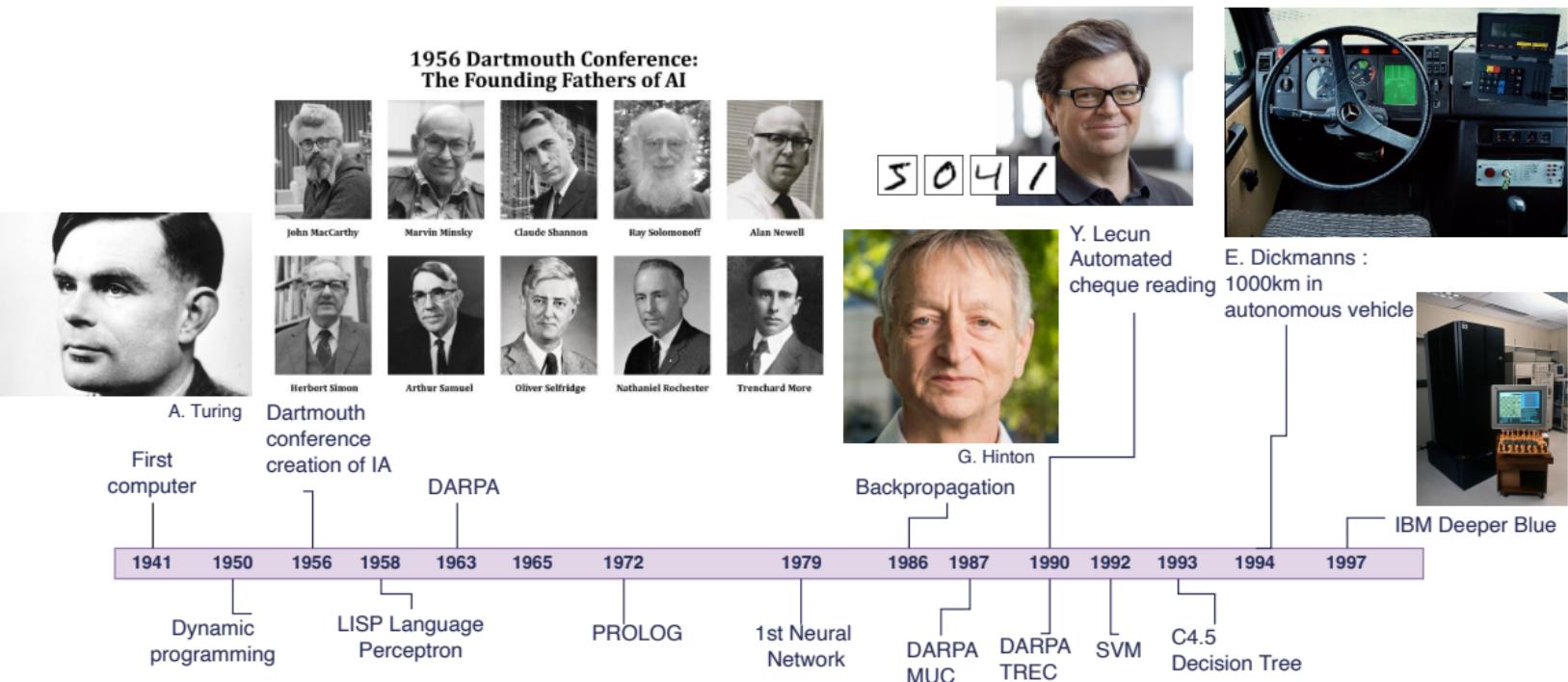


# DE L'IA AU MACHINE-LEARNING<sup>★</sup>

[<sup>\*</sup> APPRENTISSAGE STATISTIQUE/APP. AUTOMATIQUE]

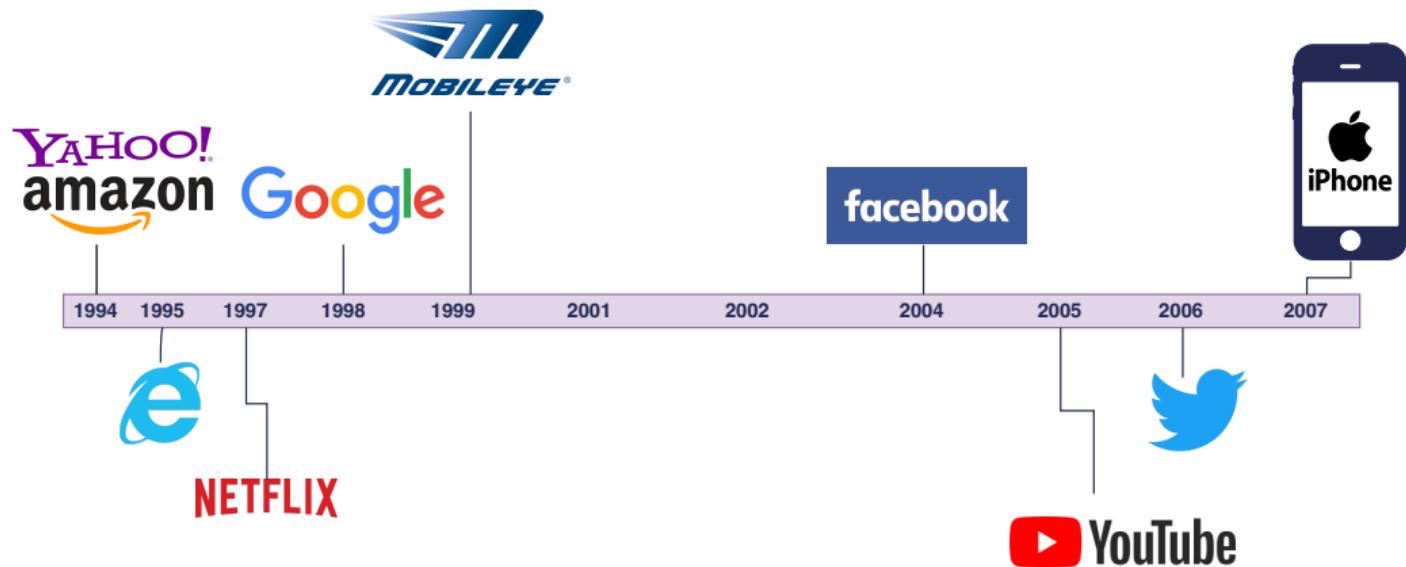
# Historique rapide de l'IA

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



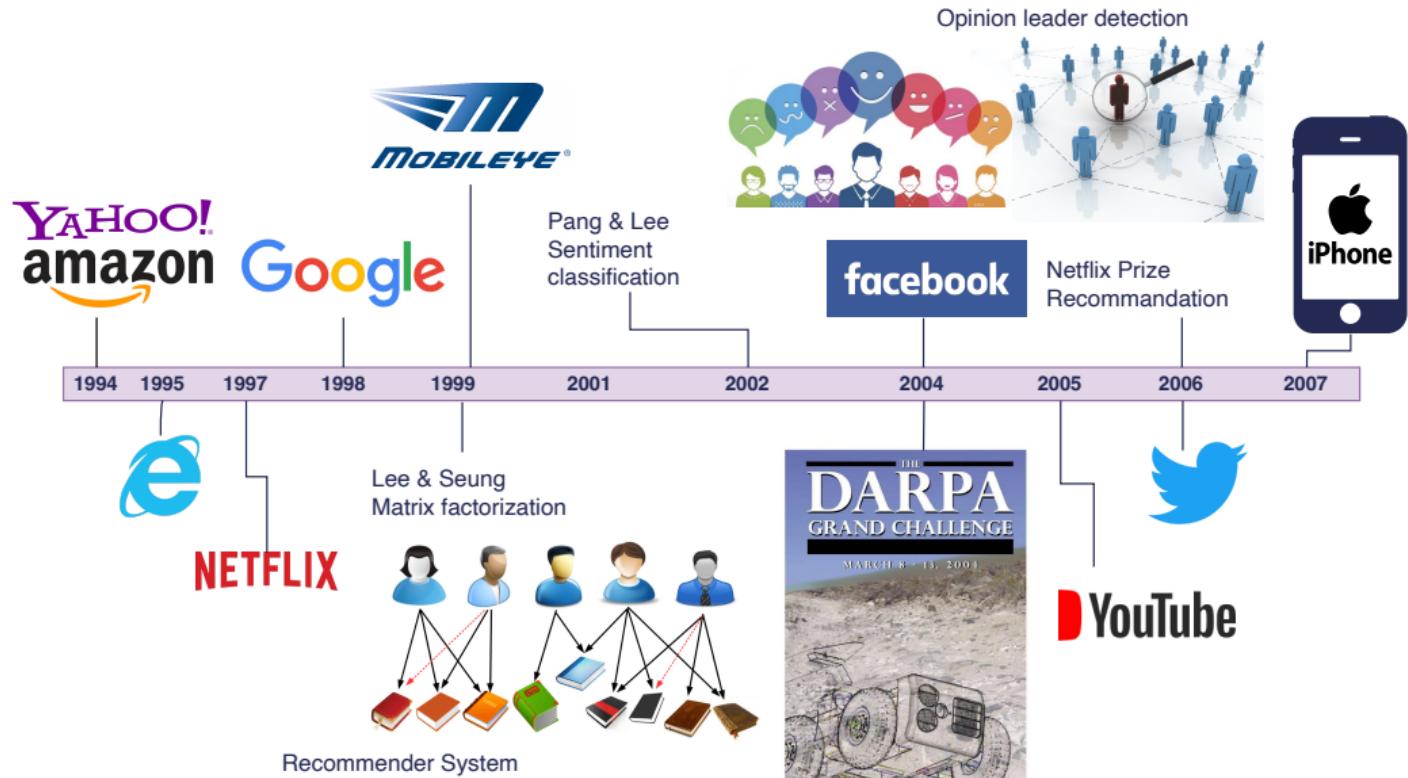
# Historique rapide de l'IA

## Emergence (ou refondation) des GAFAM/GAMMA



# Historique rapide de l'IA

## Emergence (ou refondation) des GAFAM/GAMMA

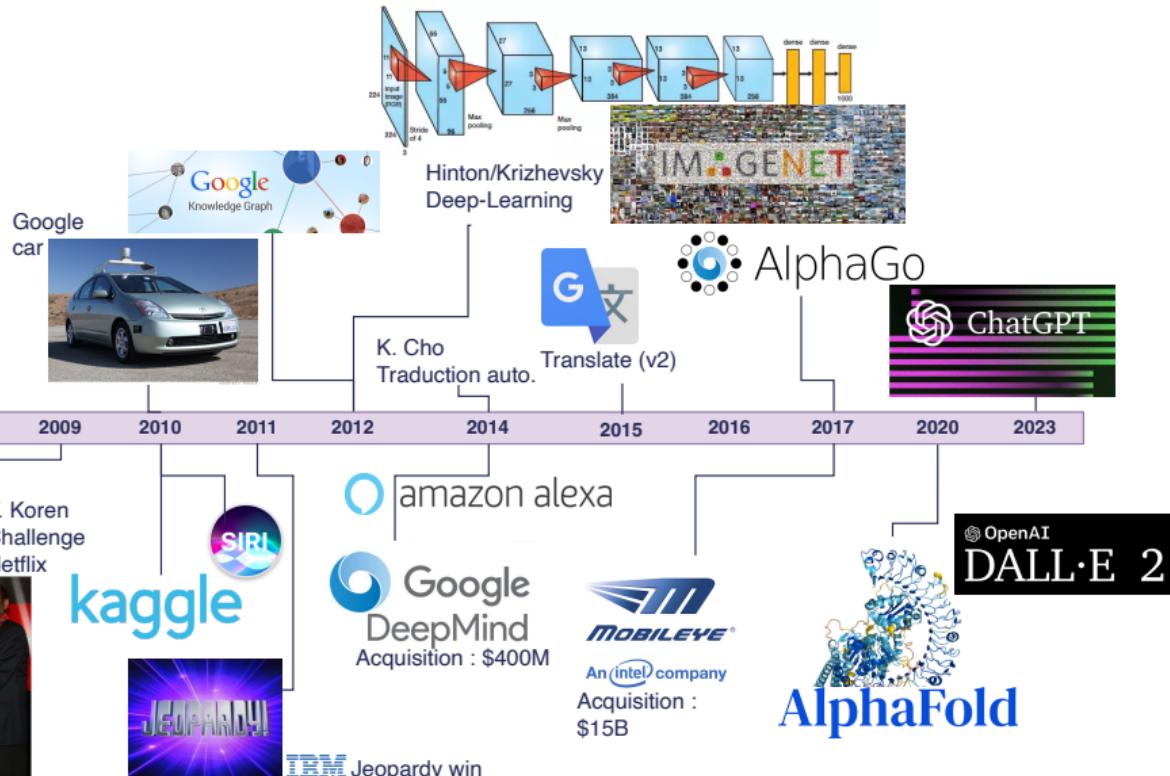


# Historique rapide de l'IA

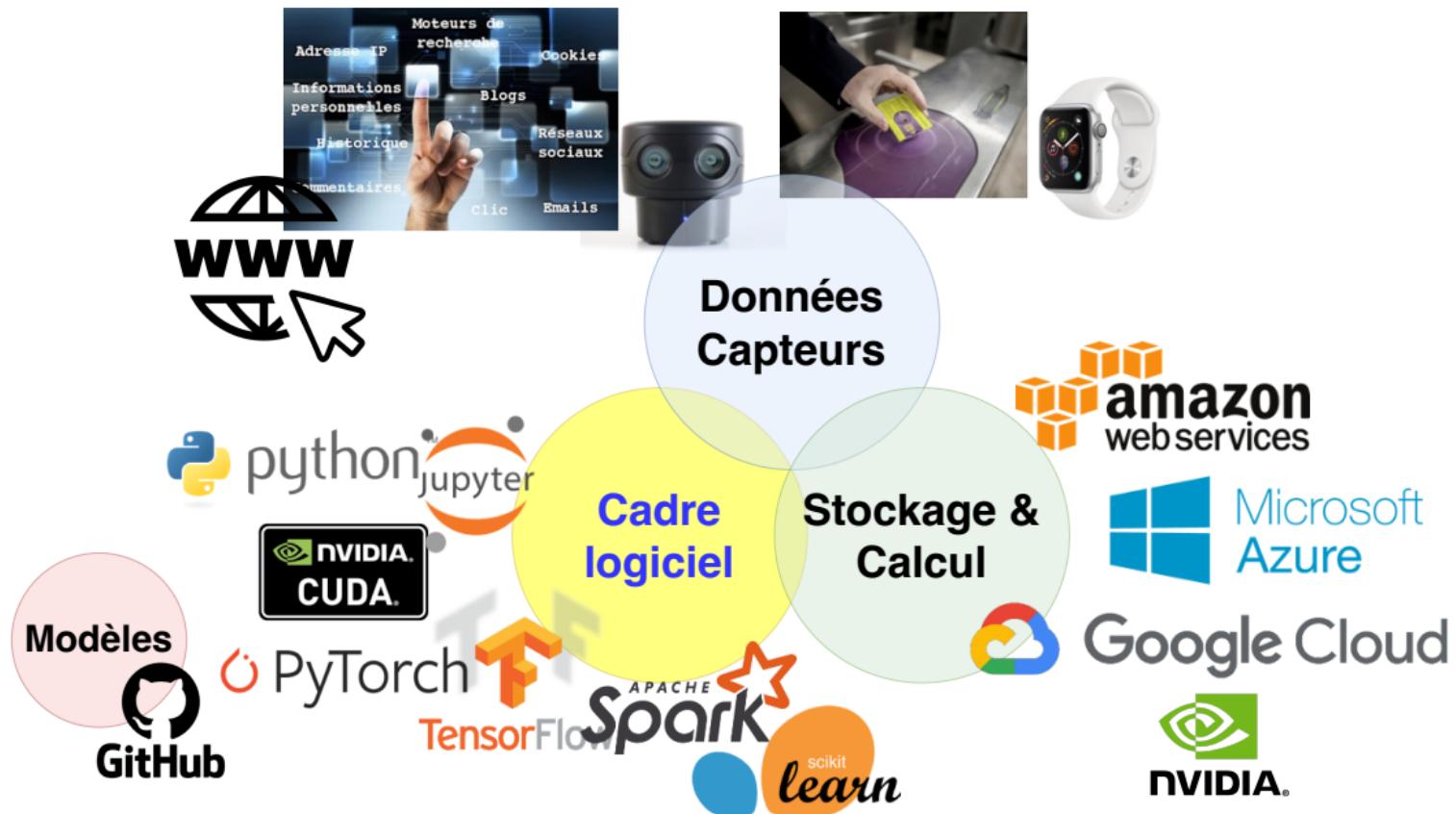
## Formation d'une vague de l'Intelligence Artificielle



**Thrun:**  
DARPA Gd Challenge  
victory

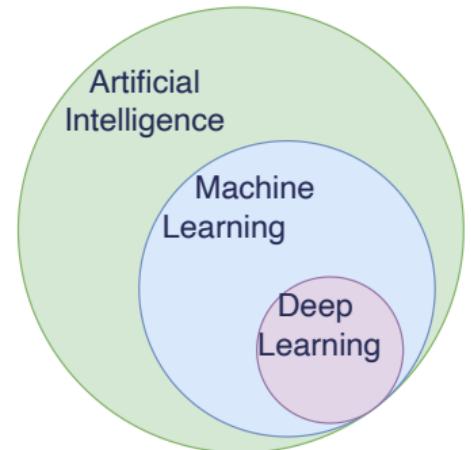


# Ingrédients de l'Intelligence Artificielle



# Intelligence Artificielle & 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.

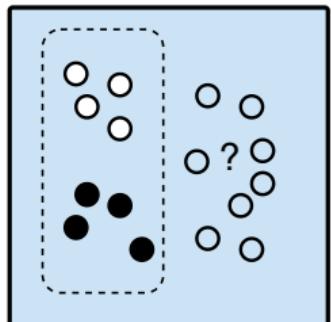
*Marvin Lee Minsky, 1956*

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

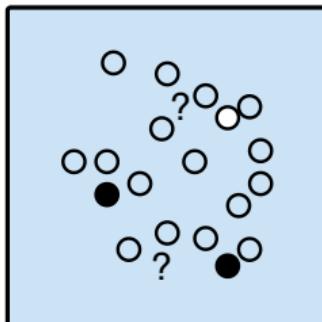
**≠ G-AI (General AI)** qui remplace l'humain dans des systèmes complexes.

*Andrew Ng, 2015*

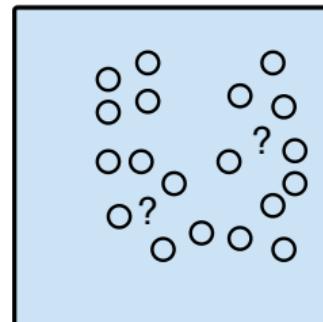
# Cadres en machine-learning



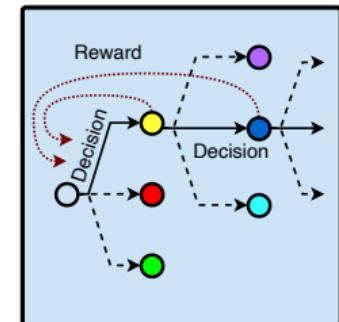
Apprentissage supervisé



Apprentissage semi-supervisé



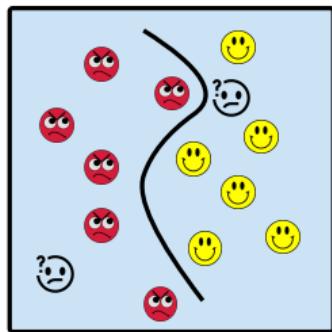
Apprentissage non-supervisé



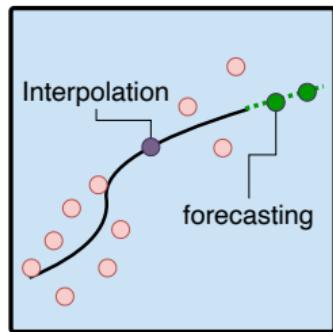
Apprentissage par renforcement

- ▶ Différentes **modalités** de données (images, textes, données numériques...)
- ▶ Différents **étiquetages**

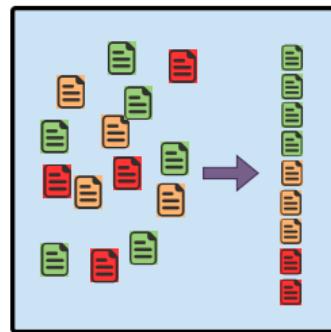
# Cadres en machine-learning



Classification



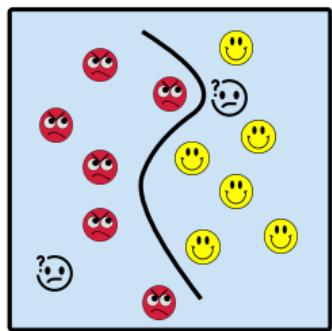
Regression



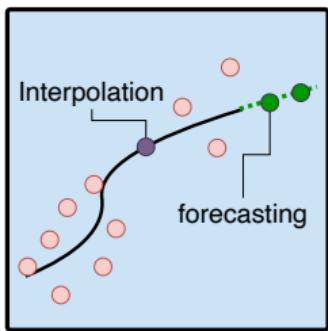
Ranking

- ▶ Différentes **modalités** de données (images, textes, données numériques...)
- ▶ Différents **étiquetages**
- ▶ Différentes types de **prédictions**

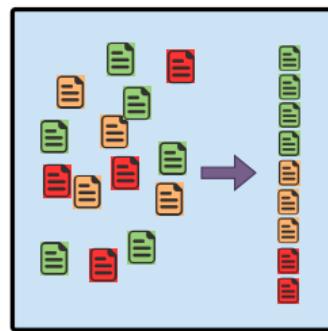
# Cadres en machine-learning



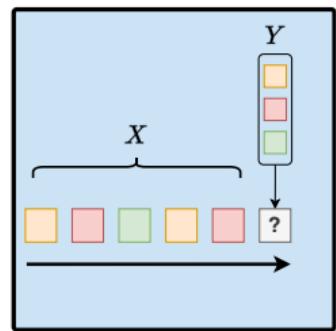
Classification



Regression



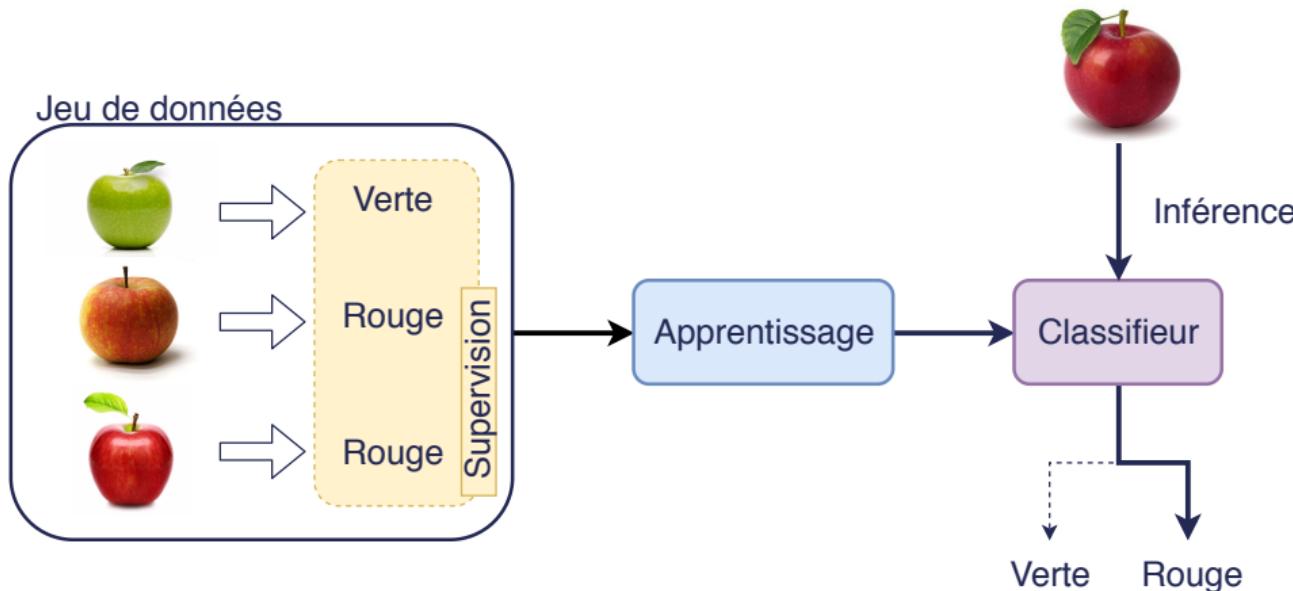
Ranking



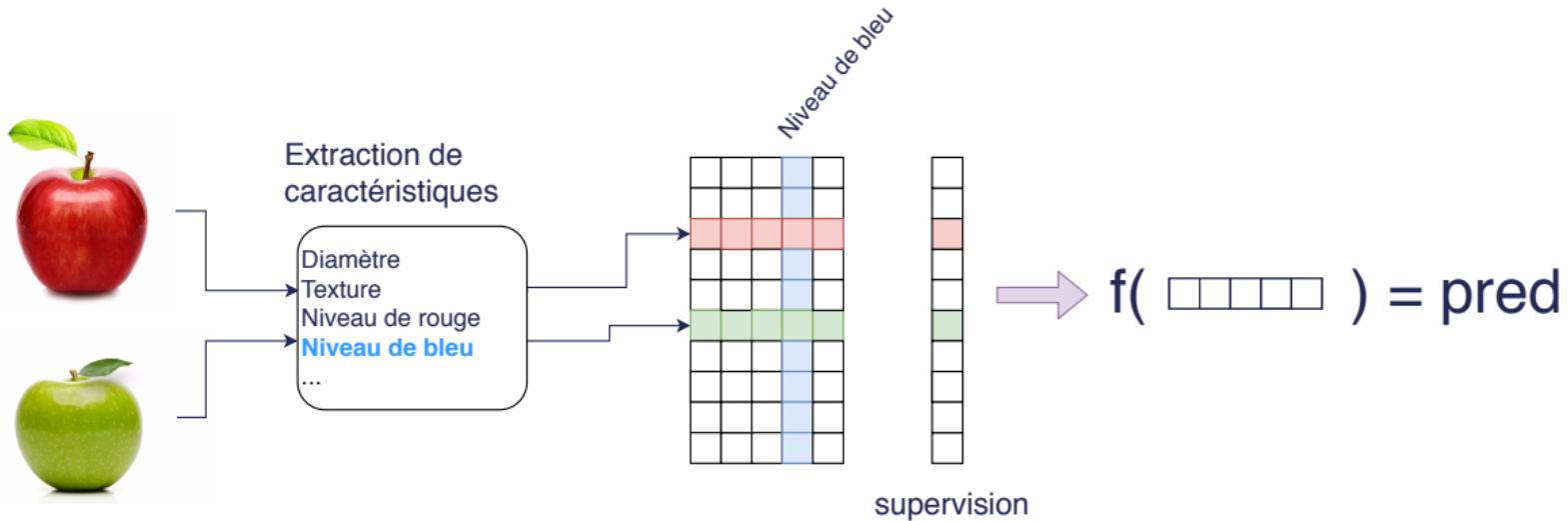
Generative AI

- ▶ Différentes **modalités** de données (images, textes, données numériques...)
- ▶ Différents **étiquetages**
- ▶ Différentes types de **prédictions**

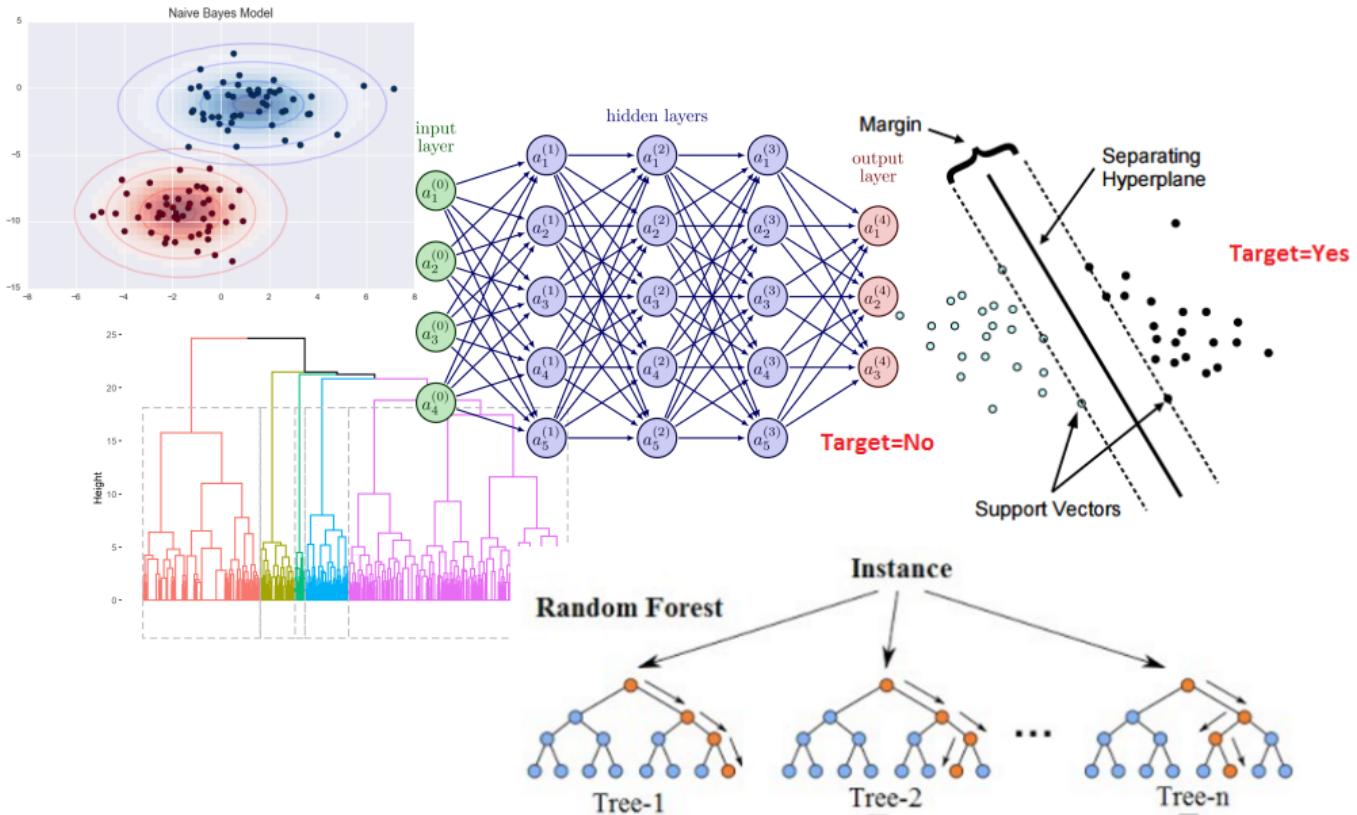
# Chaine de traitements & modèles



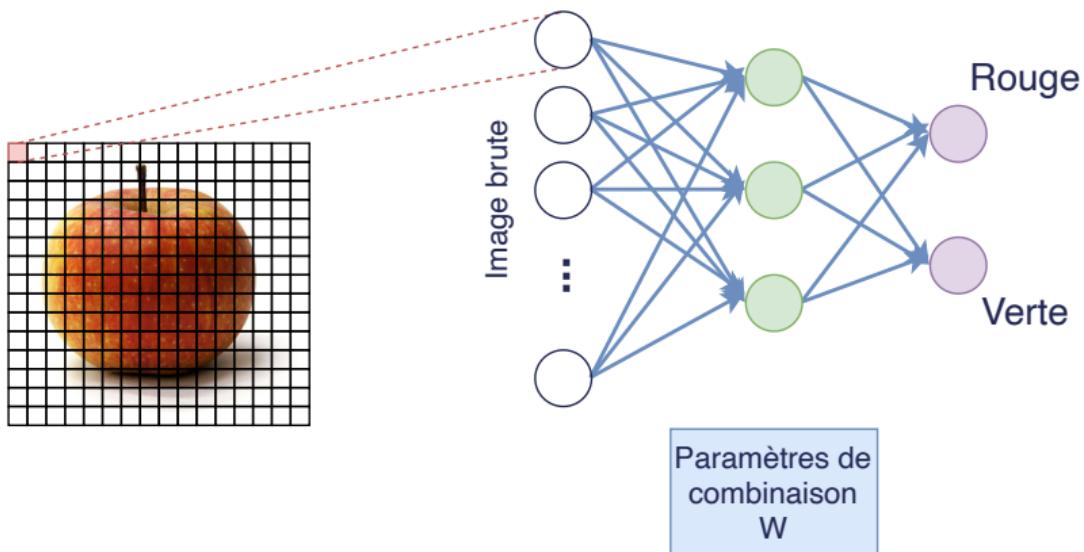
# Chaine de traitements & modèles



# Chaine de traitements & modèles



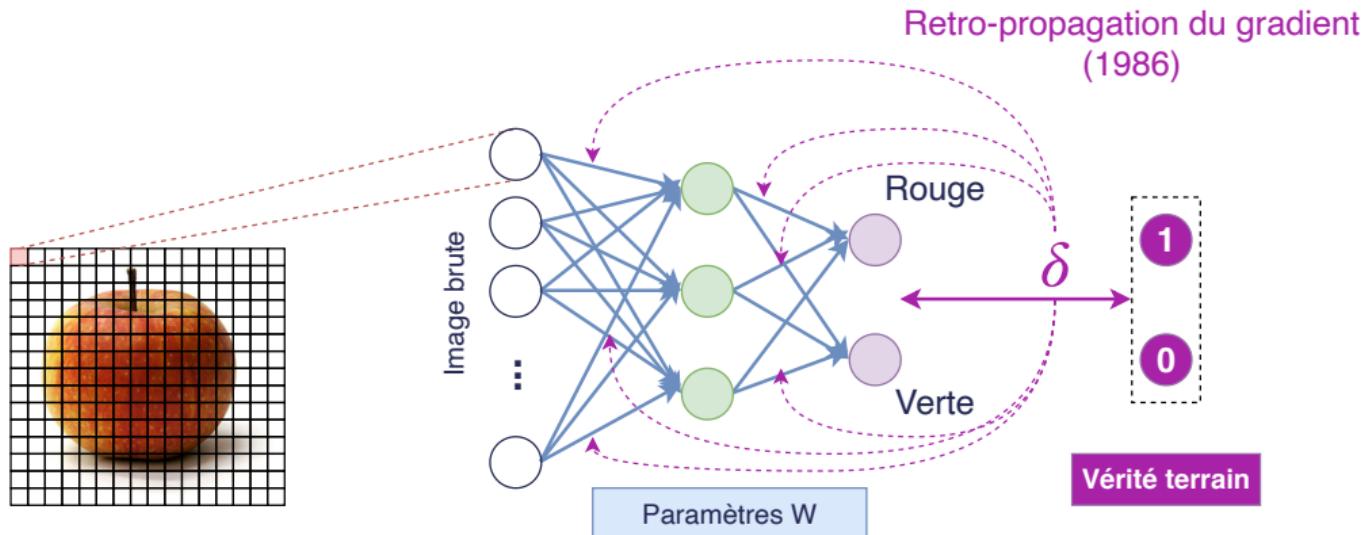
# Chaine de traitements & modèles



► Initialisation aléatoire...

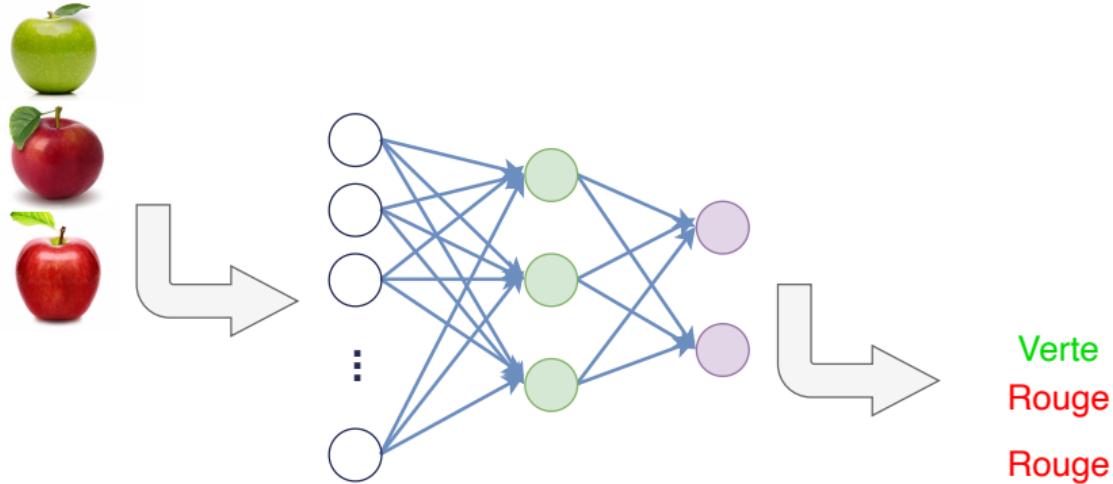
Et décision aléatoire (au début!)

# Chaine de traitements & modèles



- ▶ Mise à jour des poids
- ▶ Pas à pas epsilonesque, nombreuses itérations sur les données

# Chaine de traitements & modèles

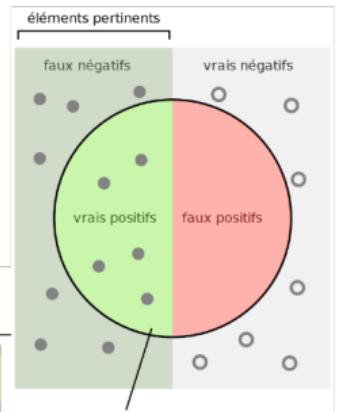
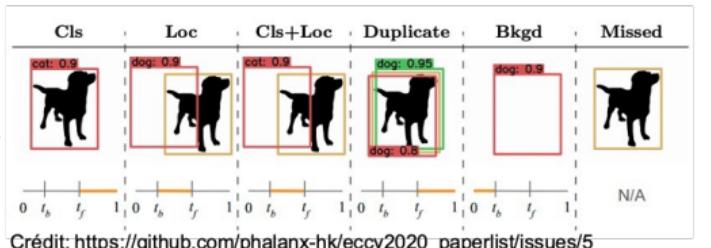
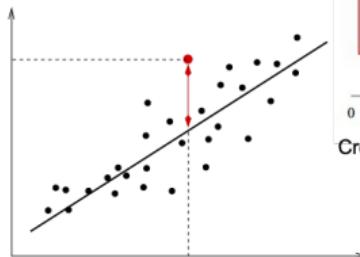


- ▶ Apprentissage lent et couteux
- ▶ Inférence (beaucoup plus) rapide

# Mesurer les performances

Estimer les performances (en généralisation)...

Est aussi important que l'apprentissage du modèle lui-même!



$$\text{Recall}@3 = 2/(2+1) = 2/3 = 0.67$$

Relevance	3	2	3	0	1
Position	1	2	3	4	5

Model

→ "the hello a cat dog fox jumps"

reference text

"the fox jumps" → ['the', 'fox', 'jumps']

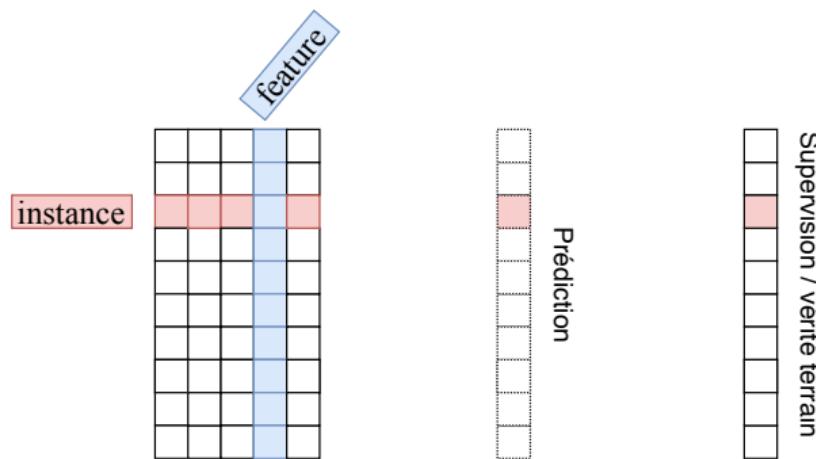
$$\frac{\text{count}_{\text{match}}(\text{gram}_n)}{\text{count}(\text{gram}_n)}$$

$$\frac{3}{7} = 43\% \text{ precision}$$

## Mesurer les performances

Estimer les performances (en généralisation)...

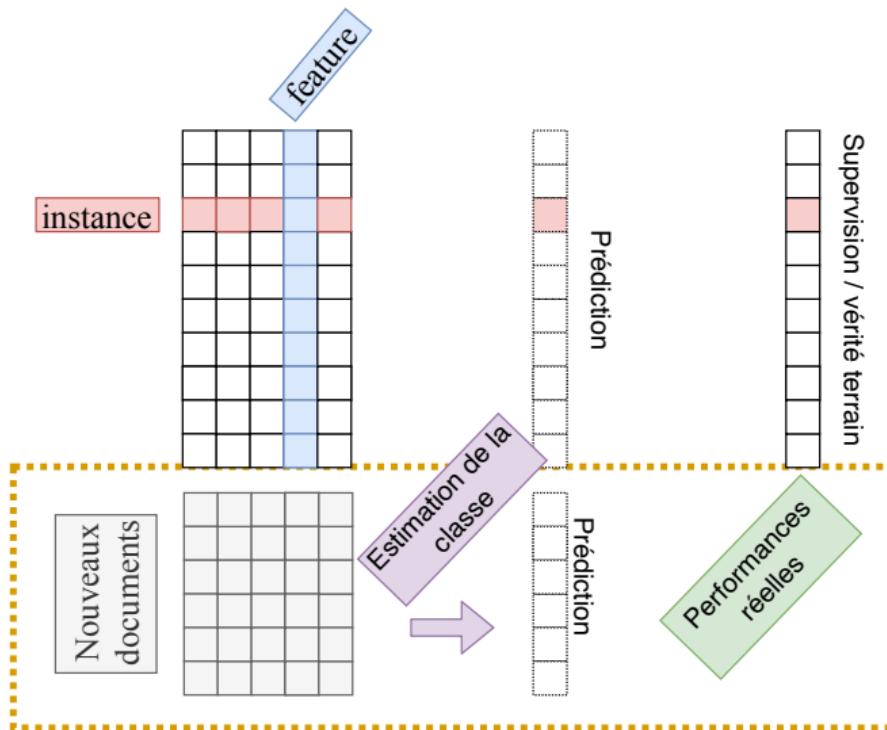
Est aussi important que l'apprentissage du modèle lui-même!



# Mesurer les performances

Estimer les performances (en généralisation)...

Est aussi important que l'apprentissage du modèle lui-même!



# DEEP LEARNING & NATURAL LANGUAGE PROCESSING<sup>★</sup>

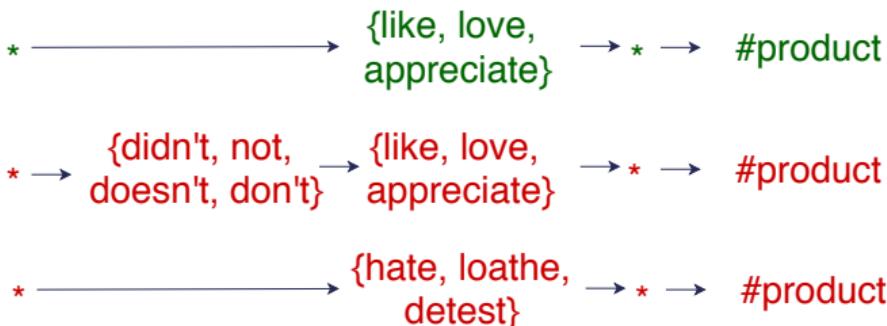
[<sup>\*</sup> TRAITEMENT AUTOMATIQUE DE LA LANGUE NATURELLE]

# L'IA et les données textuelles

TALN / NLP = plus grosse communauté scientifique de l'IA

## Linguisitique [1960-2010]

### Systèmes à base de règles:



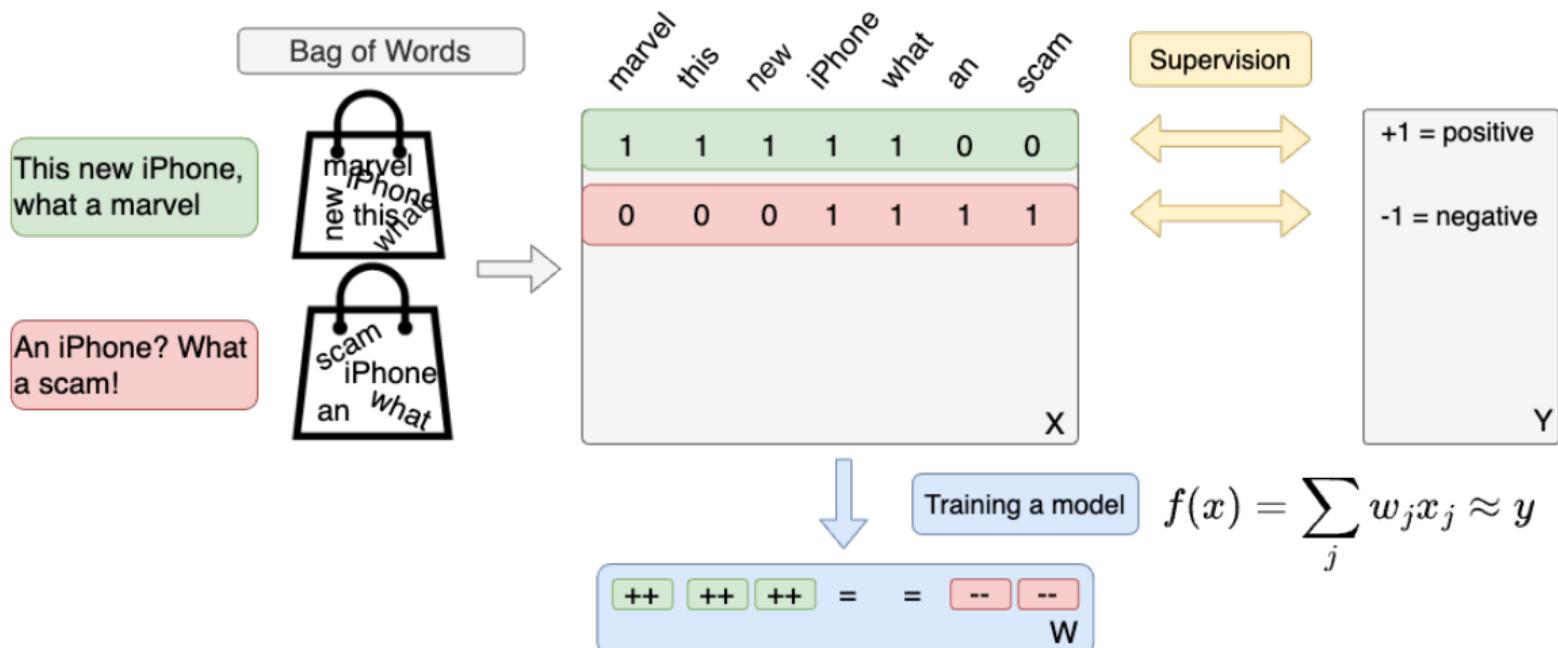
- Besoin de connaissance expert
- Extraction des règles ⇔ données très propres
- Très bonne précision
- Faible rappel
- Système interprétable



# L'IA et les données textuelles

TALN / NLP = plus grosse communauté scientifique de l'IA

## Machine Learning [1990-2015]



# L'IA et les données textuelles

TALN / NLP = plus grosse communauté scientifique de l'IA

## Linguisitique [1960-2010]

- ▶ Besoin de connaissance expert
- ▶ Extraction des règles  $\Leftrightarrow$  données très propres
- + Système interprétable
- + Très bonne précision
- Faible rappel

## Machine Learning [1990-2015]

- ▶ Peu de connaissance expert
- ▶ Extraction statistique  $\Leftrightarrow$  résiste aux données bruitées
- $\approx$  Système peu interprétable
- Précision faible
- + Meilleur rappel

Précision = critère d'acceptation par les industriels

→ Lien vers les métriques

# Deep learning et données textuelles: révolutionS 2010'

## 1. Des sacs de mots aux représentations vectorielles

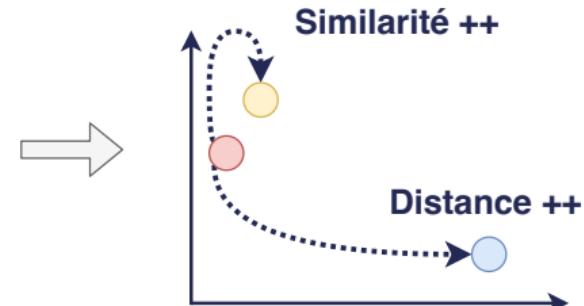
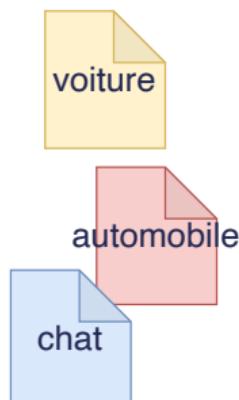
[2008, 2013, 2016]

### Corpus en sac de mots

d1	1	0	0
d2	0	0	1
d3	0	1	0

mot<sub>1</sub>    ...    voiture    ...    automobile    ...    chat    ...    mot<sub>D</sub>

Mêmes distances



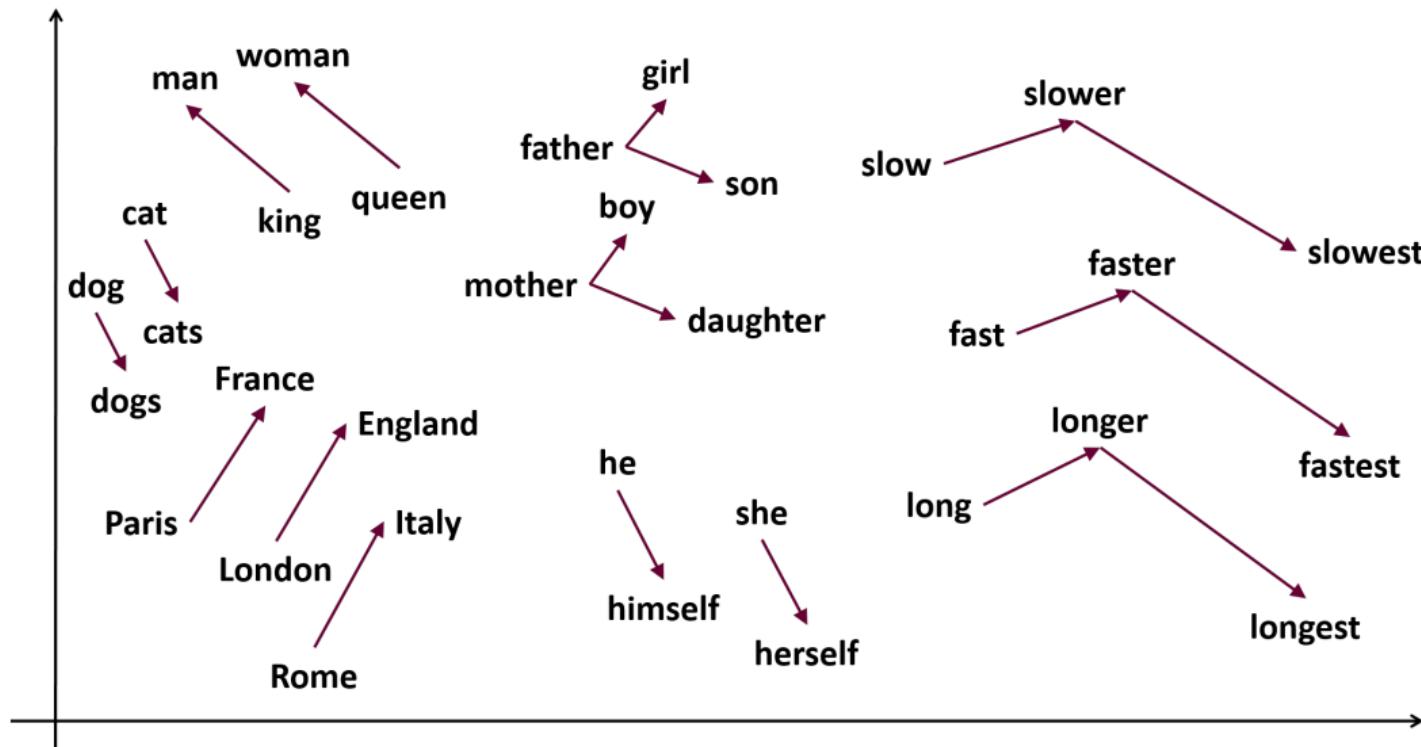
Espace vectoriel continu



# Deep learning et données textuelles: révolutionS 2010'

## 1. Des sacs de mots aux représentations vectorielles

[2008, 2013, 2016]



# Deep learning et données textuelles: révolutionS 2010'

## 2. Agrégation des représentations de mots

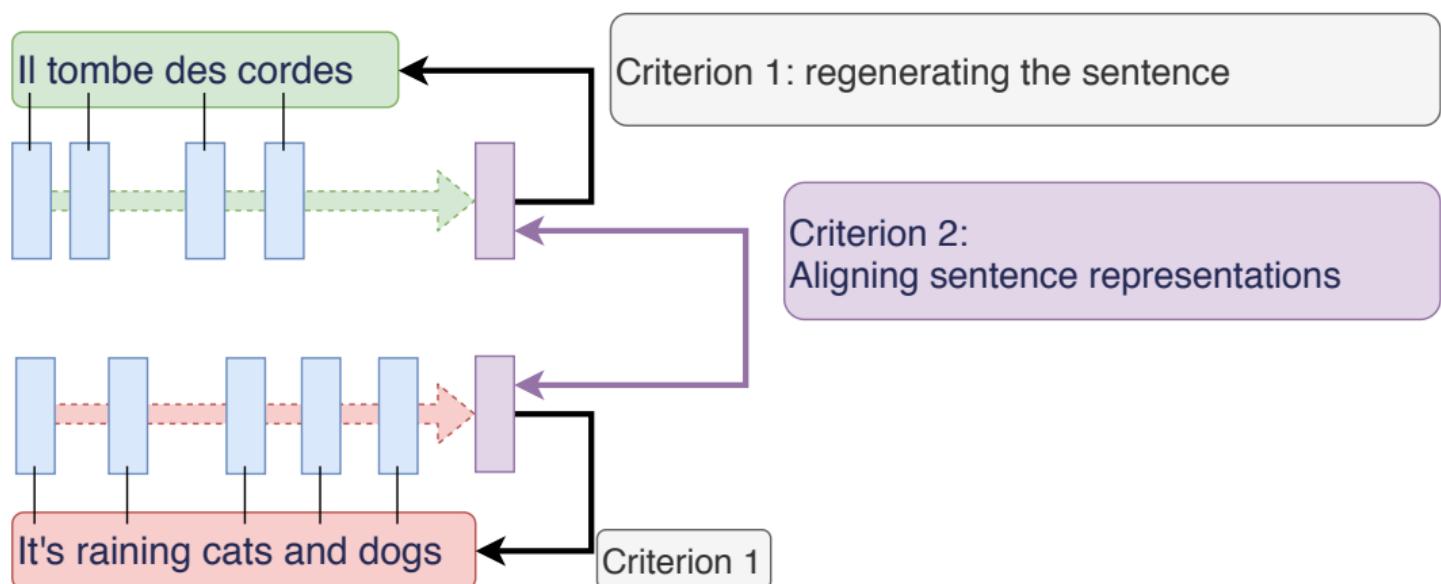
[2014, 2017]



# Deep learning et données textuelles: révolutionS 2010'

## 2. Agrégation des représentations de mots

[2014, 2017]



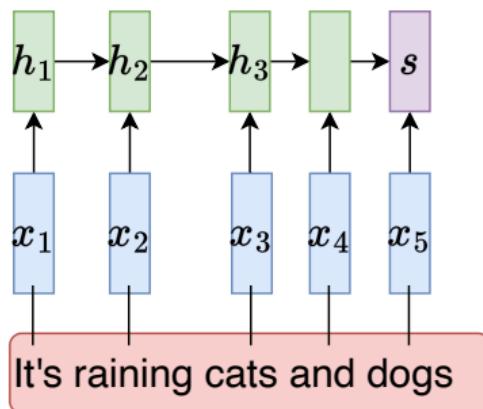
# Deep learning et données textuelles: révolutionS 2010'

## 2. Agrégation des représentations de mots... V2

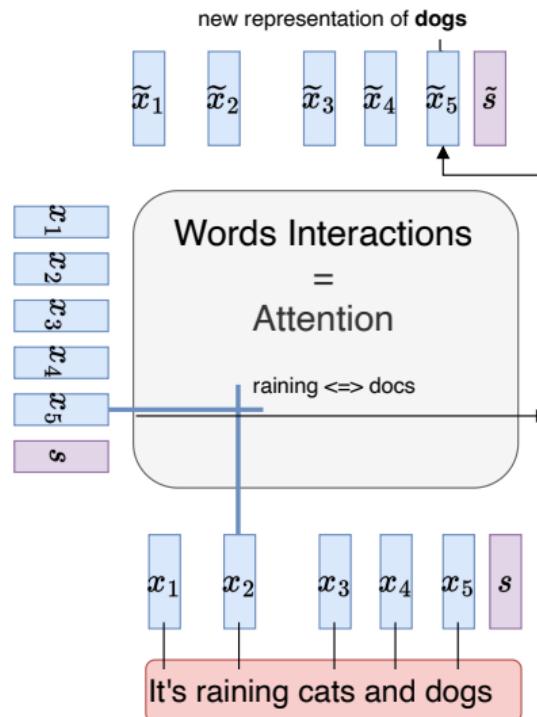
[2000, 2017-2023]

### Recurrent Neural Network:

$$h_{t+1} = h_t W_1 + x_{t+1} W_2$$



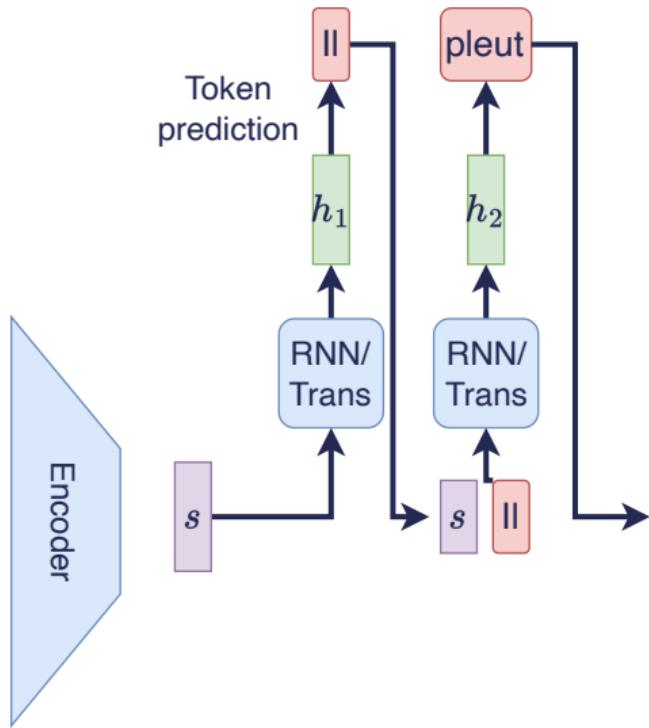
### Transformer:



# Deep learning et données textuelles: révolutionS 2010'

## 3. Des modèles génératifs

[2014, 2022]



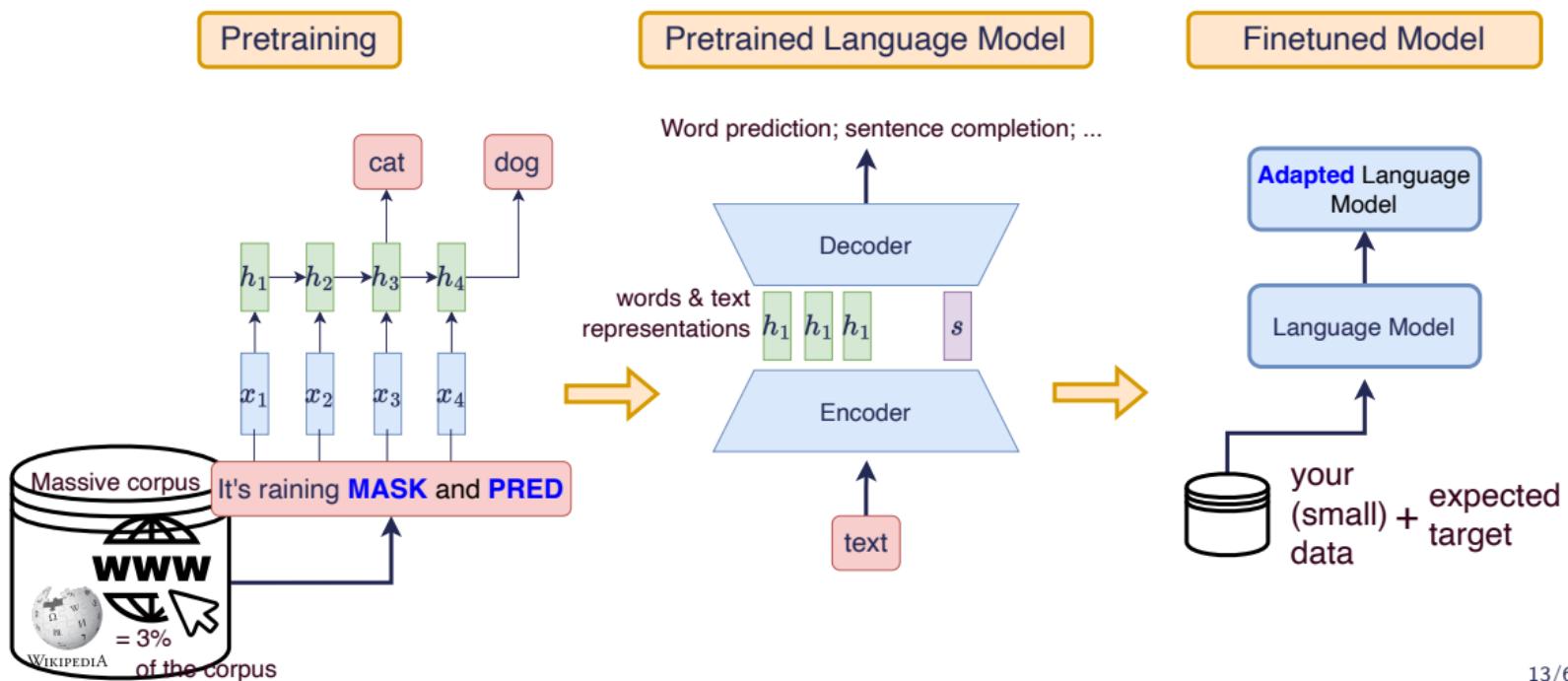
- ▶ Coût élevé (+beam search)
  - ▶ 1 appel / token
- ▶ Génération au sens du maximum de vraisemblance
- ▶ Principales tâches de NLP  $\Leftrightarrow$  reformulation en mode génératif

# Deep learning et données textuelles: révolutionS 2010'

## 4. Transfert & fine-tuning

[2008, 2012, 2018]

⇒ L'émergence des modèles de langue (larges) –LLM–



# CHATGPT

30 NOVEMBRE 2022

1 MILLION D'UTILISATEURS EN 5 JOURS

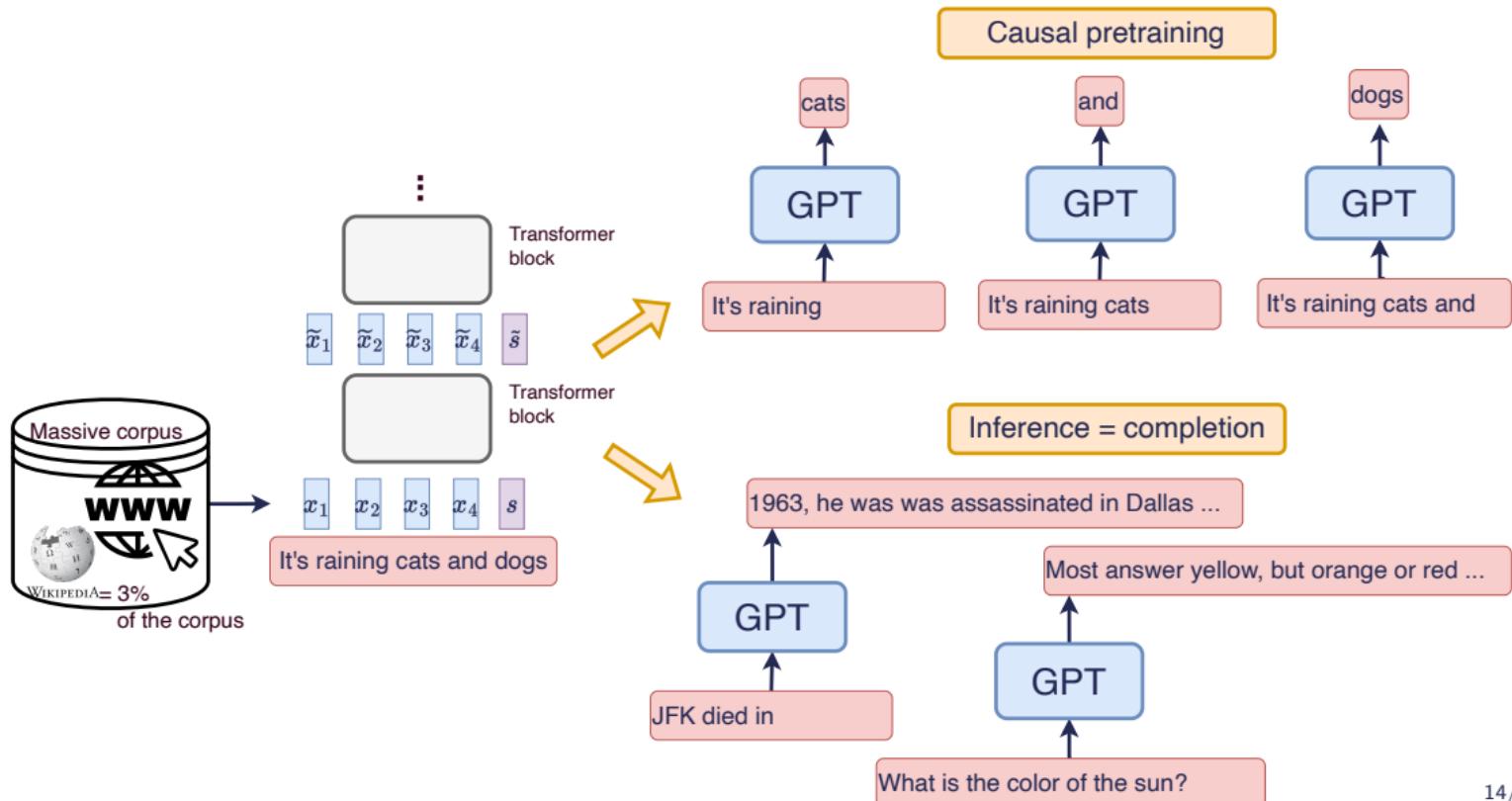
100 MILLIONS FIN JANVIER 2023

1.16 MILLIARDS EN MARS 2023



# Les ingrédients de chatGPT

## 0. Transformer + données massives (GPT)





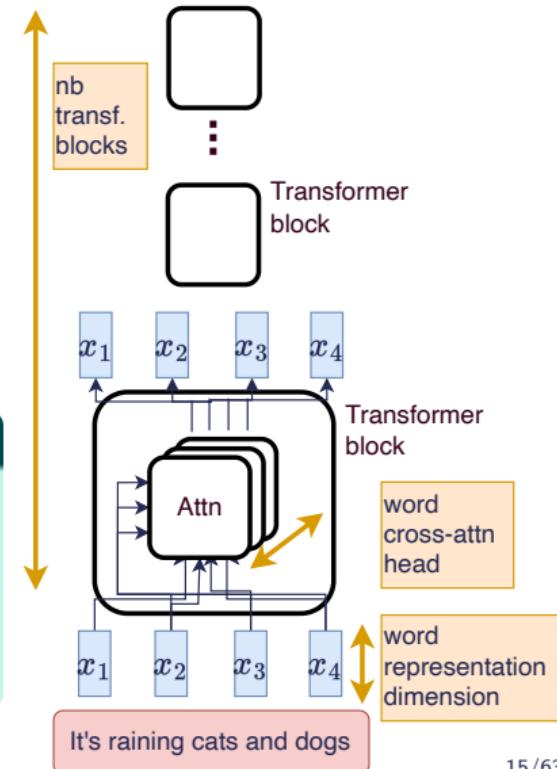
# Les ingrédients de chatGPT

## 1. Toujours plus! (GPT)

- + de mots en entrée [500  $\Rightarrow$  2k, 32k]
- + de dimensions dans l'espace des mots [500-2k  $\Rightarrow$  12k]
- + de têtes d'attention [12  $\Rightarrow$  96 (dim 128)]
- + de blocks/couches [5-12  $\Rightarrow$  96]

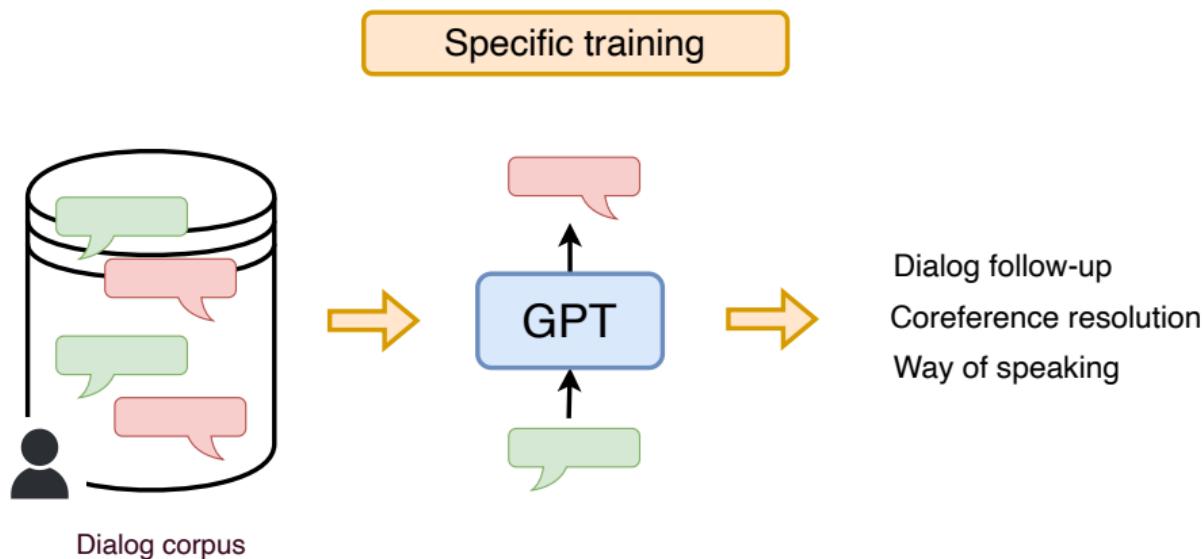
**175 Milliards** de paramètres... Ca fait quoi?

- ▶  $1.75 \cdot 10^{11} \Rightarrow 300 \text{ Go} + 100 \text{ Go}$  (stockage des données en inférence)  $\approx 400 \text{ Go}$
- ▶ GPU NVidia A100 = 80Go de mémoire ( $=20k\text{\euro}$ )
- ▶ Coût pour (1) entraînement: 4.6 Millions d' $\text{\euro}$



# Les ingrédients de chatGPT

## 2. Suivi de dialogue



► Données très propres

Données générées/validées par des humains

# Les ingrédients de chatGPT

## 3. L'affinage sur différentes tâches de raisonnement ( $\pm$ ) complexe

### 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?

### 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.

Language model

-320.4F

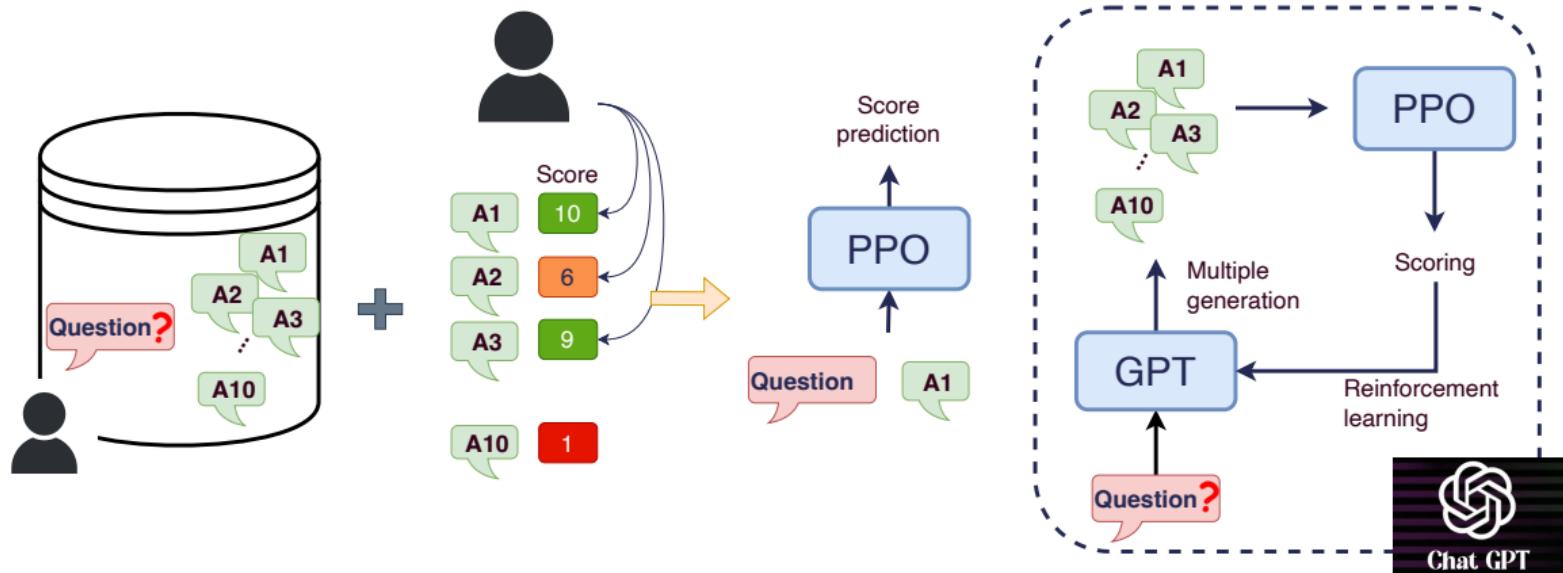
The cafeteria had 23 apples originally. They used 20 to make lunch. So they had  $23 - 20 = 3$ . They bought 6 more apples, so they have  $3 + 6 = 9$ .

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. Suivi de dialogue & amélioration des réponses

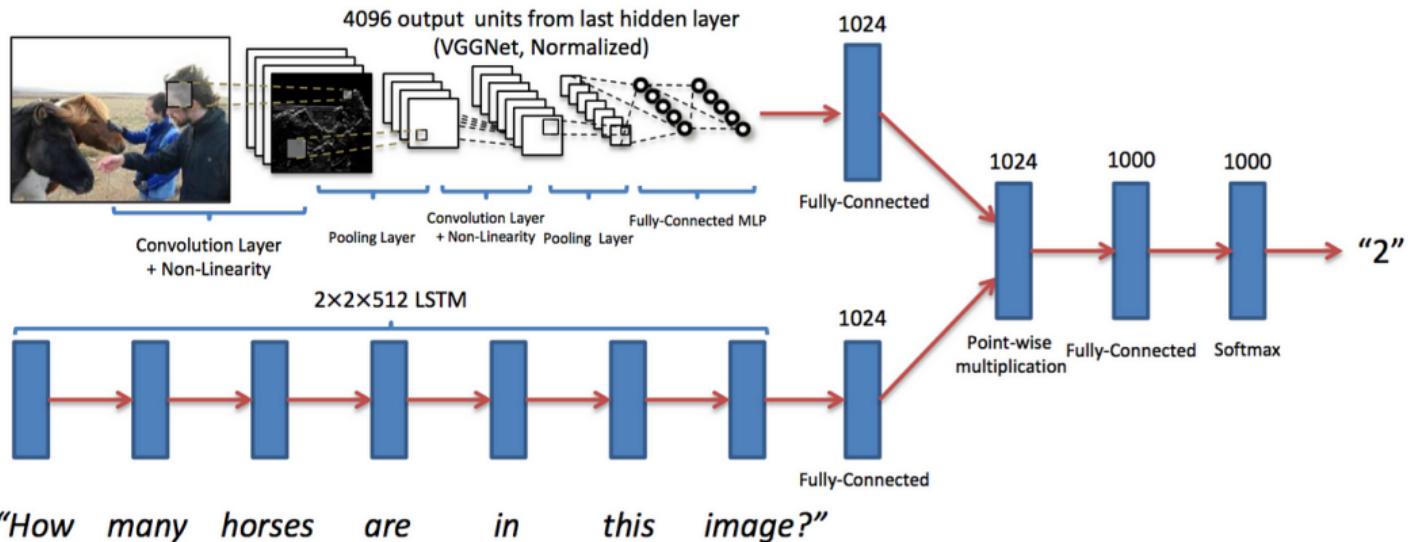


- BD faite par des humains
- Amélioration des réponses
- ... Aussi une manière d'éviter les sujets critiques

# GPT4 & la multimodalité

**Fusionner** les informations issues du texte et de l'image. **Apprendre** à exploiter les informations conjointement

*L'exemple du VQA: visual question answering*



"How many horses are in this image?"

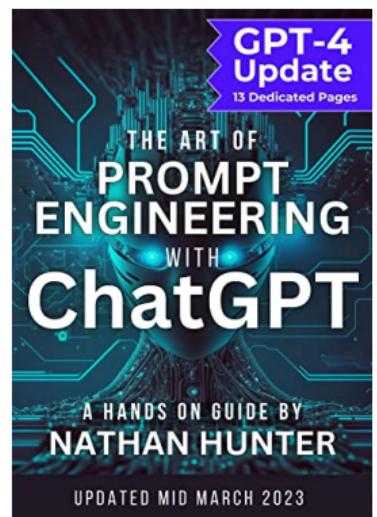
⇒ Rétro-propager l'erreur ⇒ modifier les représentations de mots + analyse de l'image

→ + de détails

# Usage de chatGPT & Prompting

- ▶ Interroger chatGPT... Ca s'apprend! = *prompting*
  - ▶ Bien poser une question: ... *en détails*, ... *step by step*
  - ▶ Spécifier nb elts e.g. : *3 qualités pour ...*, *5 éléments pour...*
  - ▶ Poser un contexte : *cellule* pour un biologiste / assistant juridique
- ▶ Ne pas s'arrêter à la première question
  - ▶ Détailler des points particuliers
  - ▶ réorienter la recherche
- ▶ Reformulation
  - ▶ Explain like I'm 5, plus formel, à la manière d'un article scientifique, bro style, ...
  - ▶ Résumer, étendre
  - ▶ Ajouter des fautes (!)

⇒ Besoin d'expériences, pas d'un livre !



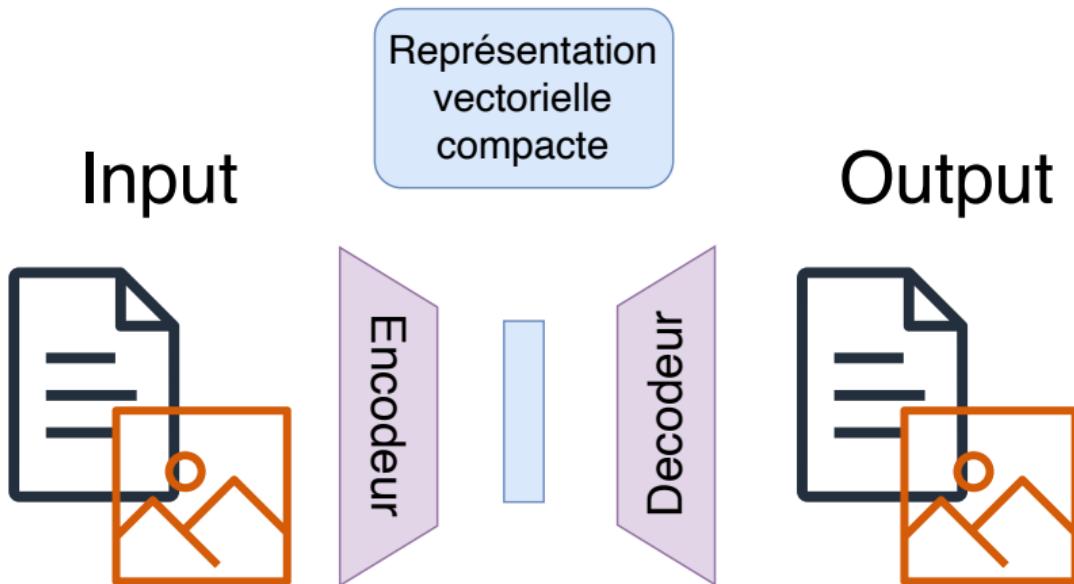
# Vers du *few-shot learning*

- ▶ Apprendre sans modifier le modèle = exemples dans le prompt

	This person is like 😊.		This person is like 😊.		This person is like	<b>Model Completion</b>	 . <EOS>
	This was invented by Zacharias Janssen.		This was invented by Thomas Edison.		This was invented by	<b>Model Completion</b>	the Wright brothers. <EOS>
	With one of these I can drive around a track, overtaking other cars and taking corners at speed		With one of these I can take off from a city and fly across the sky to somewhere on the other side of the world		With one of these I can	<b>Model Completion</b>	break into a secure building, unlock the door and walk right in <EOS>

# IAs GÉNÉRATIVES

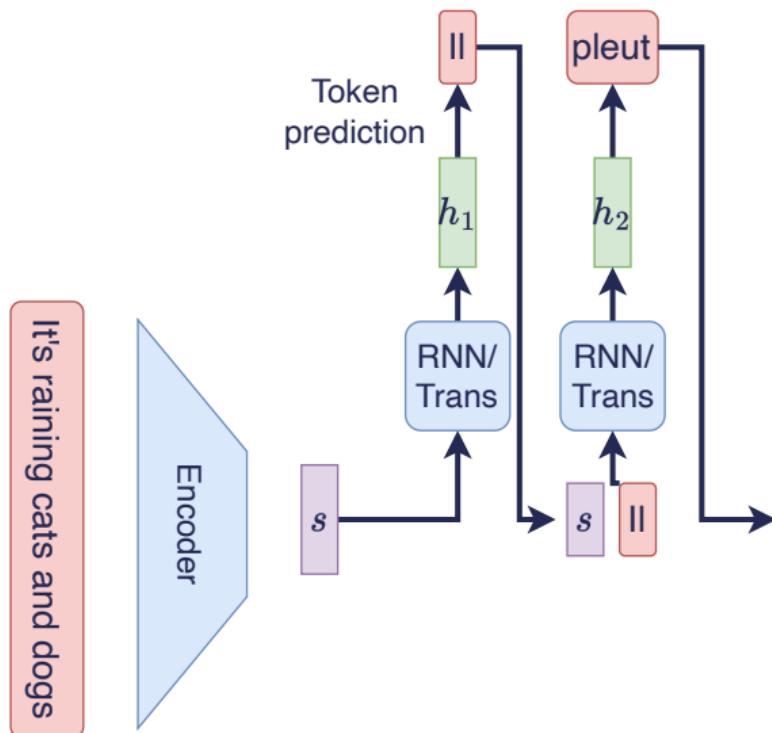
# Différentes formes d'IA génératives



- 1 Encoder une entrée = construire un vecteur
- 2 Décoder un vecteur = générer une sortie

# Différents media / différentes architectures

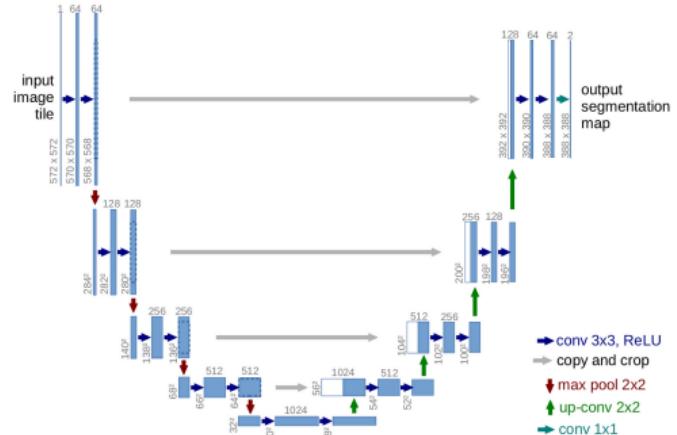
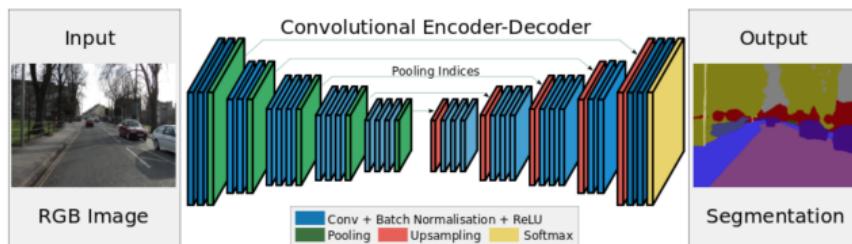
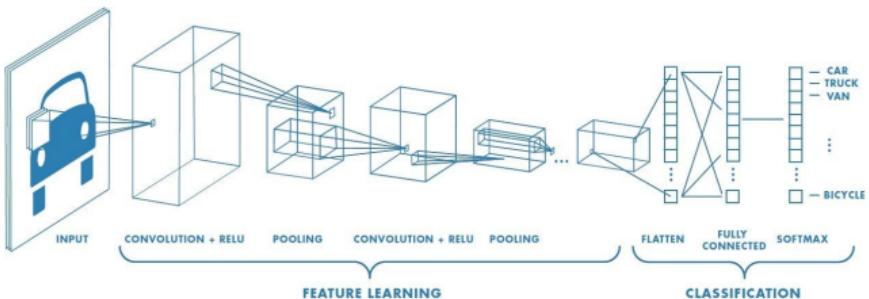
- ▶ Textes : problème classification





# Différents media / différentes architectures

- ▶ Textes : problème classification
- ▶ Images : problème de régression multi-variées



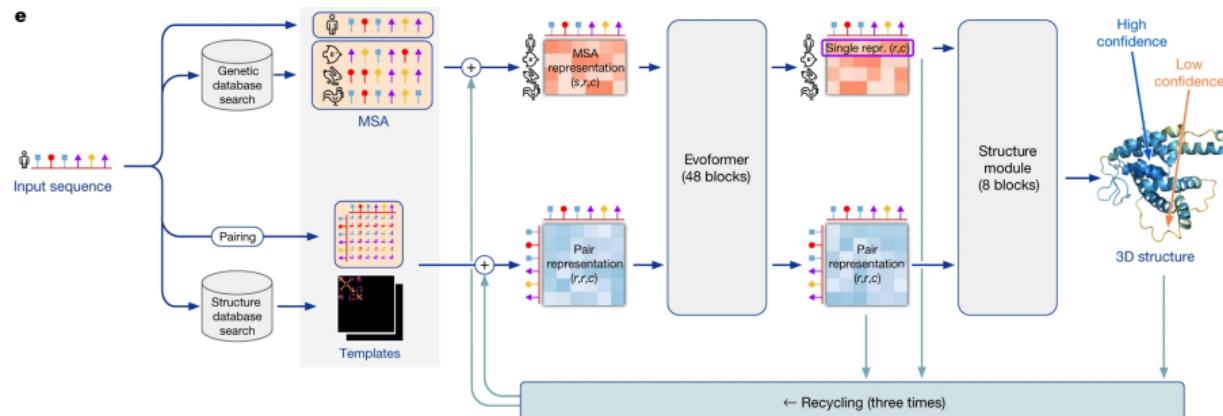
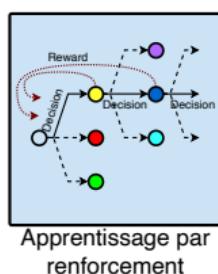
Architecture U-Net

# Différents media / différentes architectures

- ▶ Textes : problème classification
- ▶ Images : problème de régression multi-variées
- ▶ Structures complexes / 3D / graphes : problème séquentiel

- ▶ A partir d'un point donné:

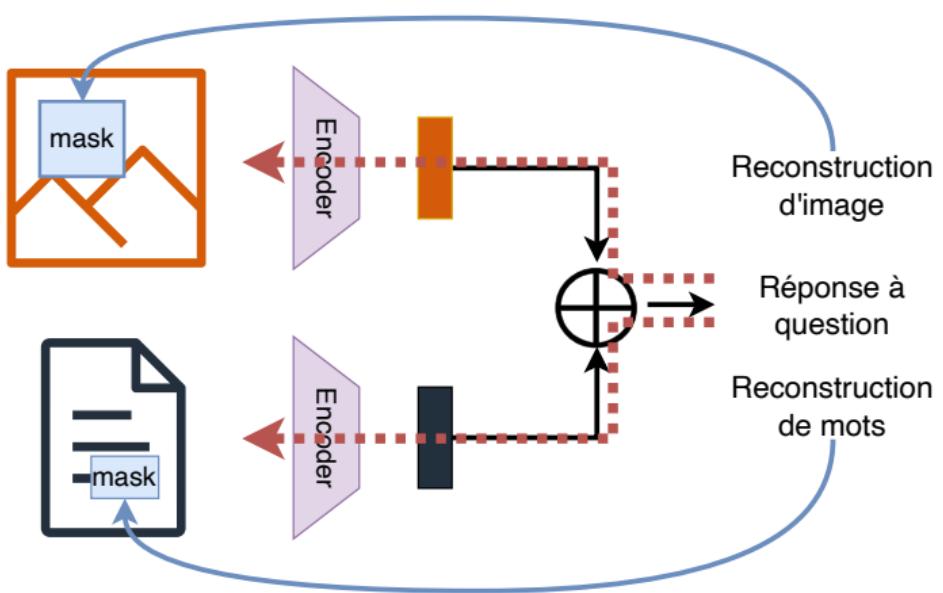
- ▶ Générer/tester une coordonnée / un noeud (=action)
- ▶ Evaluer la vraisemblance (=récompense)



AlphaFold

# Multi-modalité

- ▶ Construction d'espaces de représentation multimodaux = *grounding*
- ▶ Image ⇒ Texte : *Captioning, Visual Question Answering*
- ▶ Texte ⇒ Image : *mid-journey, dall-e, ...*

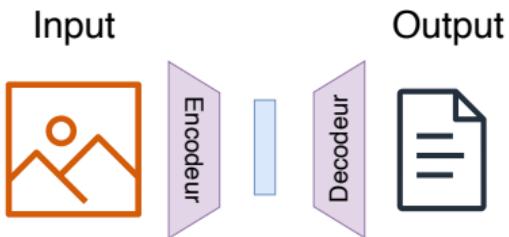


## Alignment des espaces de représentations

Word	Teraword	Knext
Spoke	11,577,917	372,042
Laughed	3,904,519	179,395
Murdered	2,843,529	16,890
Inhaled	984,613	5,617
Breathed	725,034	41,215

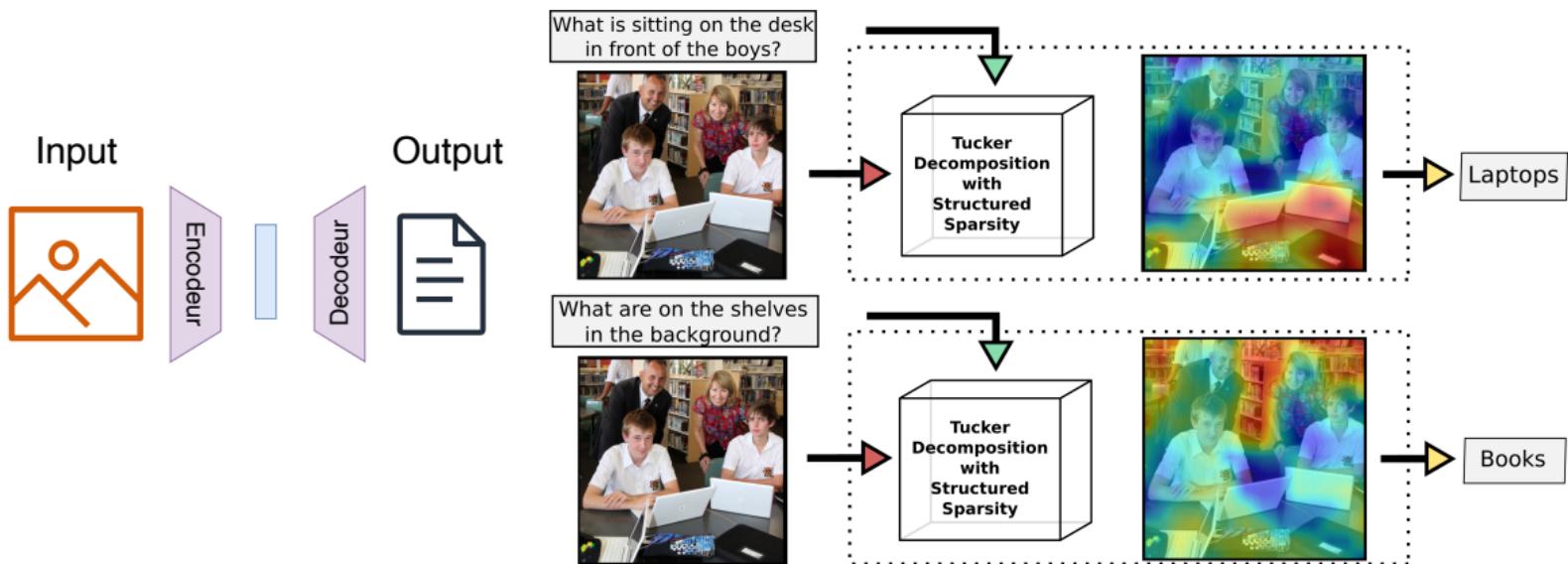
# Multi-modalité

- ▶ Construction d'espaces de représentation multimodaux = *grounding*
- ▶ Image ⇒ Texte : *Captioning, Visual Question Answering*
- ▶ Texte ⇒ Image : *mid-journey, dall-e, ...*



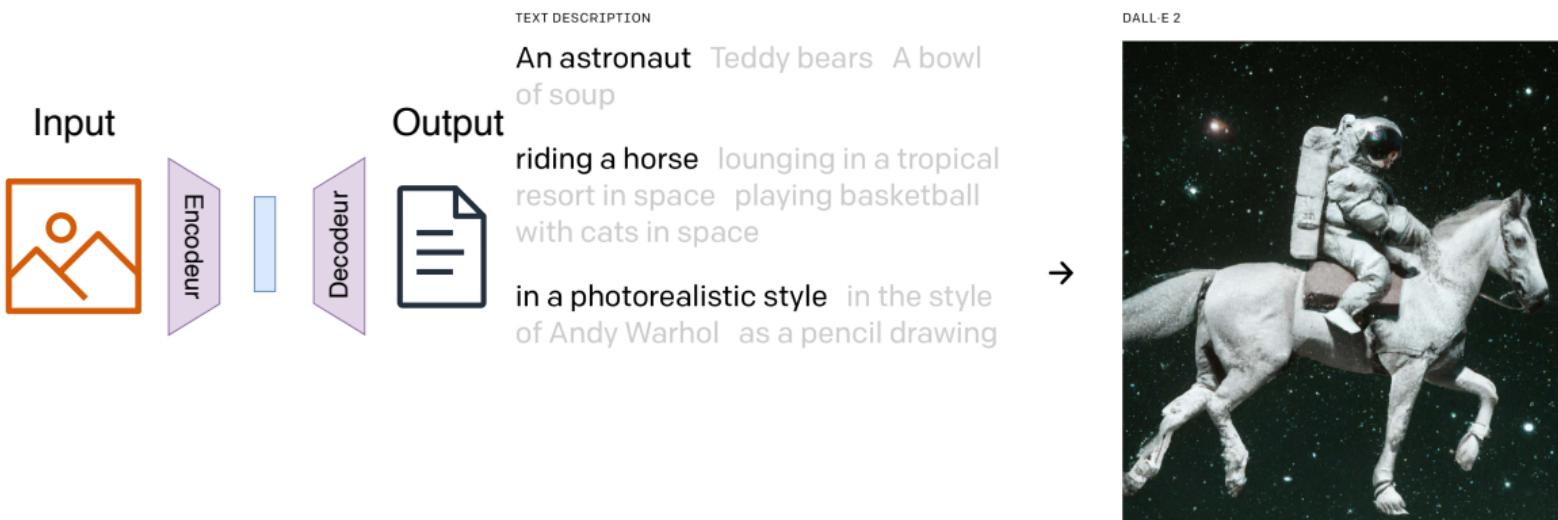
# Multi-modalité

- ▶ Construction d'espaces de représentation multimodaux = *grounding*
- ▶ Image ⇒ Texte : *Captioning, Visual Question Answering*
- ▶ Texte ⇒ Image : *mid-journey, dall-e, ...*



# Multi-modalité

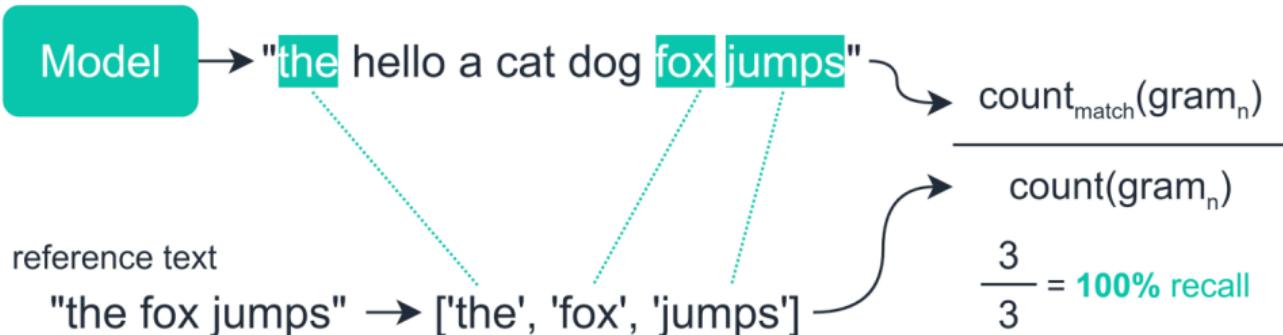
- ▶ Construction d'espaces de représentation multimodaux = *grounding*
- ▶ Image ⇒ Texte : *Captioning, Visual Question Answering*
- ▶ Texte ⇒ Image : *mid-journey, dall-e, ...*



# IA Génératives: comment évaluer les performances?

## Le point critique aujourd'hui

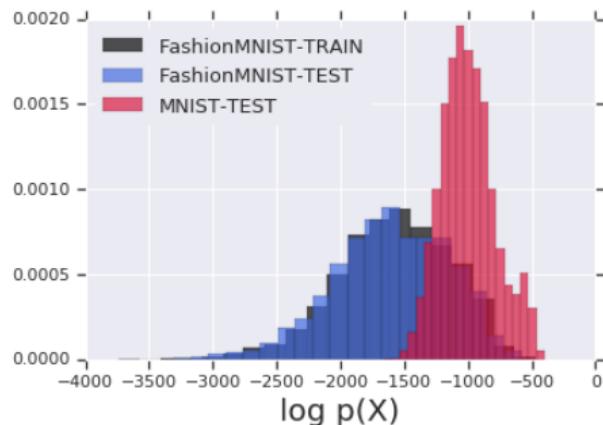
- ▶ Comment s'évaluer par rapport à une vérité terrain?
- ▶ Comment évaluer la confiance du système / vraisemblance de la génération ?



# IA Génératives: comment évaluer les performances?

## Le point critique aujourd'hui

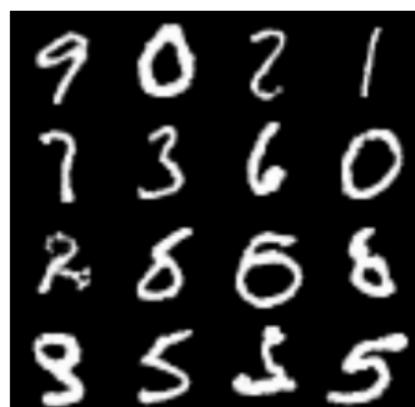
- ▶ Comment s'évaluer par rapport à une vérité terrain?
- ▶ Comment évaluer la confiance du système / vraisemblance de la génération ?



Vraisemblance



Train



Test

# LIMITES

# Limites générales de chatGPT

## ► Manque de **vérité / fiabilité**

Excellent es performances en moyenne... Mais des erreurs bêtes.

Quelle chance de faire une erreur sur la prochaine question?

**Vérité ≠ Vraisemblance** – Génération d'**hallucinations**

## ► Manque de **stabilité/prédicibilité**

IA des jeux vidéo ≠ machine learning

*how old is Obama VS how old is Obama?*

## ► Manque d'**explicabilité/interprétabilité**

Pourquoi cette réponse?

Pourquoi n'a-t-il pas mis de négation dans la phrase?

## ► Manque de **transparence**

Code privé, pas encore d'article scientifique

## ► Limite dans la **dimension des textes**

Limite à **2000 mots** en entrée (32000 avec GPT-4)

# chatGPT et le rapport à la vérité

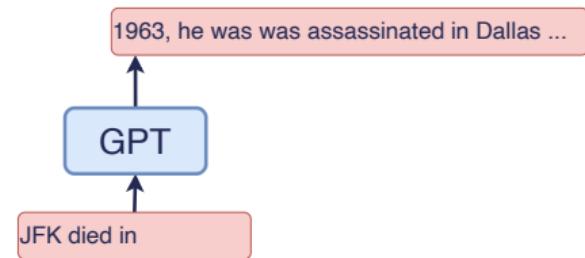
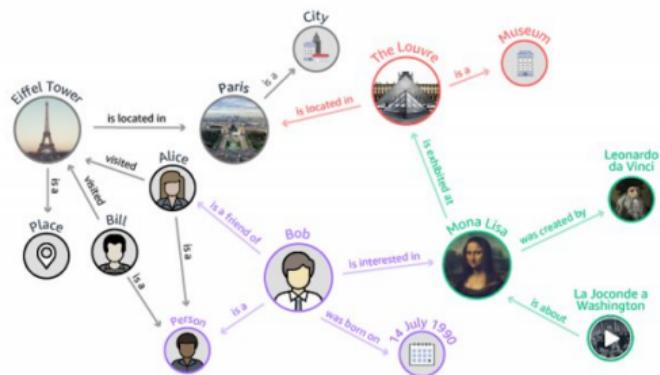
- 1 Vraisemblance = grammaire, accords, concordance des temps, enchaînements logiques...  
⇒ Connaissances répétées  
≈ grammaire

- 2 Prédire le mot le plus **vraisemblable**...  
⇒ produit des **hallucinations**

- 3 Fonctionnement **hors-ligne**

- 4 chatGPT =  
loin des **graphes de connaissances**

- 5 Des réponses brillantes...  
Et des erreurs bêtes!  
+ on ne sait pas prédire les erreurs



# chatGPT et le rapport à la vérité

1 Vraisemblance = grammaire, accords,  
concordance des temps,  
enchaînements logiques...

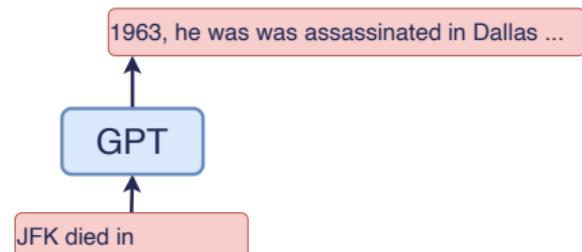
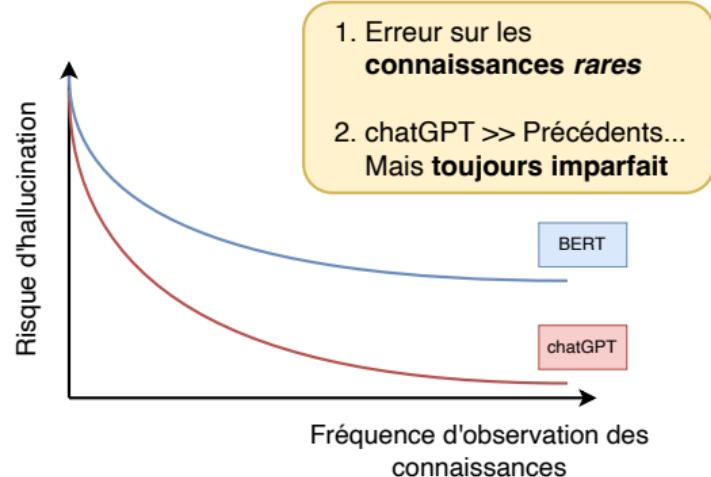
⇒ Connaissances répétées  
≈ grammaire

2 Prédire le mot le plus **vraisemblable**...  
⇒ produit des **hallucinations**

3 Fonctionnement **hors-ligne**

4 chatGPT =  
loin des **graphes de connaissances**

5 Des réponses brillantes...  
Et des erreurs bêtes!  
+ on ne sait pas prédire les erreurs



# Stabilité/prédictibilité

- ▶ Difficile de borner un comportement
- ▶ Impossible de prédire les bonnes/mauvaises réponses
- ⇒ Peu/pas utilisé en jeux vidéo



V

how old is Obama



Barack Obama was born on August 4, 1961, making him 61 years old as of February 2, 2023.



# Stabilité/prédicibilité

- ▶ Difficile de borner un comportement
  - ▶ Impossible de prédire les bonnes/mauvaises réponses
- ⇒ Peu/pas utilisé en jeux vidéo



V how old is obama?  
==



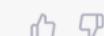
As of 2021, Barack Obama was born on August 4, 1961, so he is 60 years old.



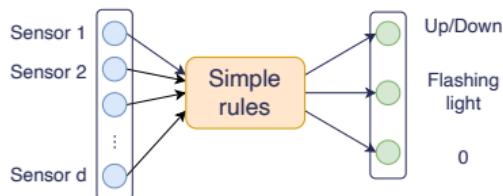
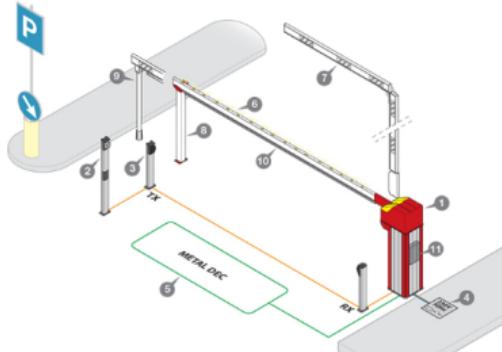
V and today?



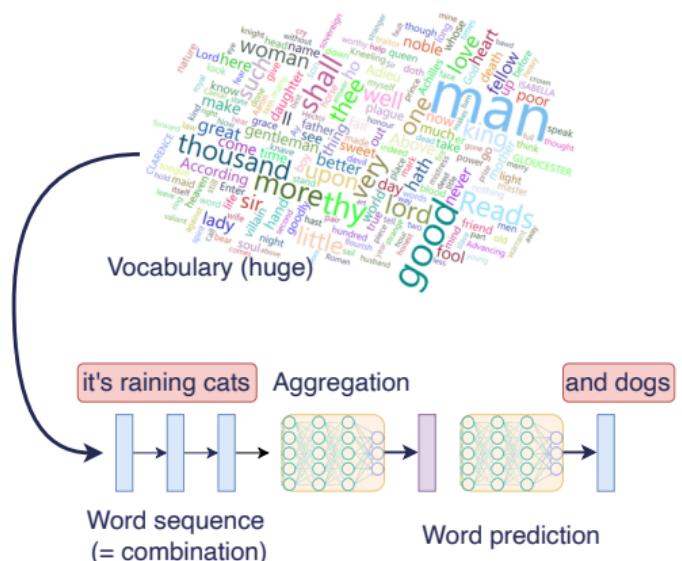
As a language model AI, I don't have real-time access to current dates. However, Barack



# Stabilité, explicabilité... Et complexité



- ▶ Système *simple*
- ▶ Test exhaustif des entrées/sorties
- ▶ Comportement **prédictible** & **explicable**



- ▶ Grande dimension
- ▶ Combinaisons non-linéaires complexes
- ▶ Comportement **non-prédictible** & **non-explicable**

# Stabilité, explicabilité... Et complexité

## Interprétabilité vs Explication post'hoc

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

*trop de combinaisons pour anticiper*

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

*roles des entrées dans une décision sur un exemple*



[Accident Uber, 2018]

- ▶ Système *simple*
- ▶ Test exhaustif des entrées/sorties
- ▶ Comportement **prédictible** &  
**explicable**

- ▶ Grande dimension
- ▶ Combinaisons non-linéaires complexes
- ▶ Comportement **non-prédictible** &  
**non-explicable**

# USAGES & RISQUES

# Usages

Une fois admis le besoin de relecture, comment exploiter l'outil?

## 1 Assistant pédagogique / accès information

- ▶ Wikipedia ++ / (coté enseignant) Génération de quizz...



## 2 Développement argumentaire (et recherche de contradiction)

- ▶ Proposition de motivation sur un sujet
- ▶ Reformulation de paragraphe



## 3 Assistant pour le développement informatique

- ▶ Génération de code, recherche d'erreur, ...

## 4 Assistant personnel

- ▶ Courrier standard, lettres de recommandation, de motivations...
- ▶ Aide à la rédaction: article, projet, compte-rendu

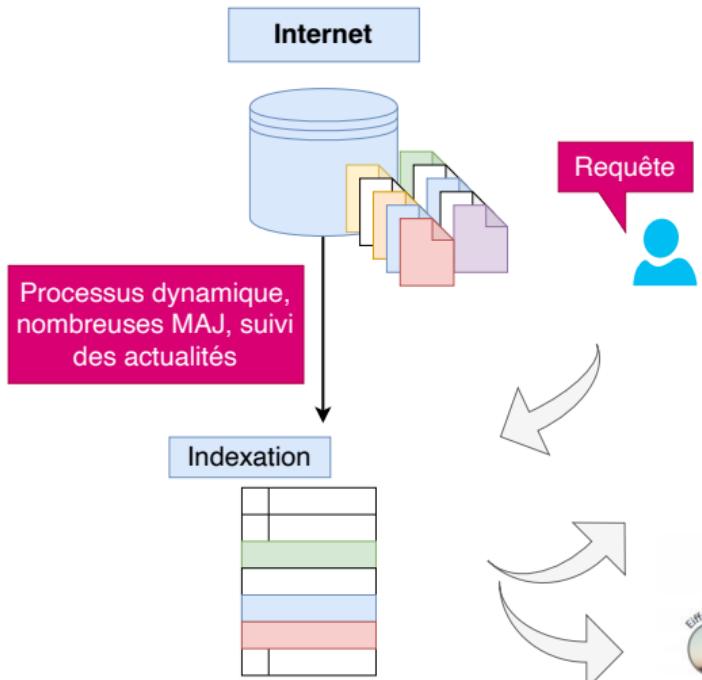


## 5 Analyse de documents

- ▶ Extraction d'information, question answering, ...



# Usage en accès à l'information



Google search results for "information":

- Environ 25 270 000 000 résultats (0,30 secondes)
- Information
- 1. Action d'informer quelqu'un, un groupe, de le tenir au courant des événements : La presse est un moyen d'information. 2. Indication, renseignement, précision que l'on donne ou que l'on obtient sur quelqu'un ou quelque chose : Manquer d'informations sur les causes d'un accident.
- Larousse
- https://www.larousse.fr/dictionnaires/francais/infor... ↗
- information, informations - Dictionnaire de français Larousse
- À propos des extraits optimisés • Commentaires
- Franceinfo
- https://www.franceinfo.fr ↗
- Franceinfo - Actualités en temps réel et info en direct
- Pour savoir ce qui se passe maintenant - Toutes les infos live toutes minute par minute par la rédaction de Franceinfo. Photos, vidéos, tweets et vos ...
- Direct Radio - Direct TV - En direct - Fait-divers

Résultats sourcés

20 Minutes

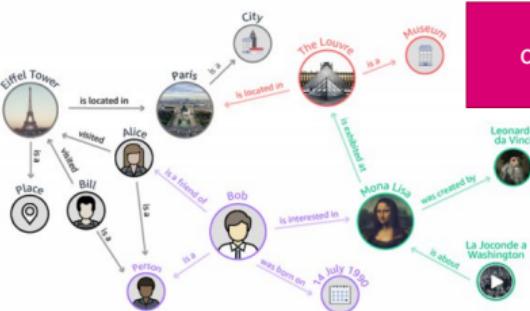
https://www.20minutes.fr ↗

20 Minutes - Toute l'actualité en direct et les dernières infos en ...

Suivez l'actualité du jour sur 20 Minutes, média gratuit et indépendant. Politique, Sport, Culture, High Tech, Ecologie... toute l'info en continu.

Le direct - Actualité générale - Jeux - Guerre en Ukraine

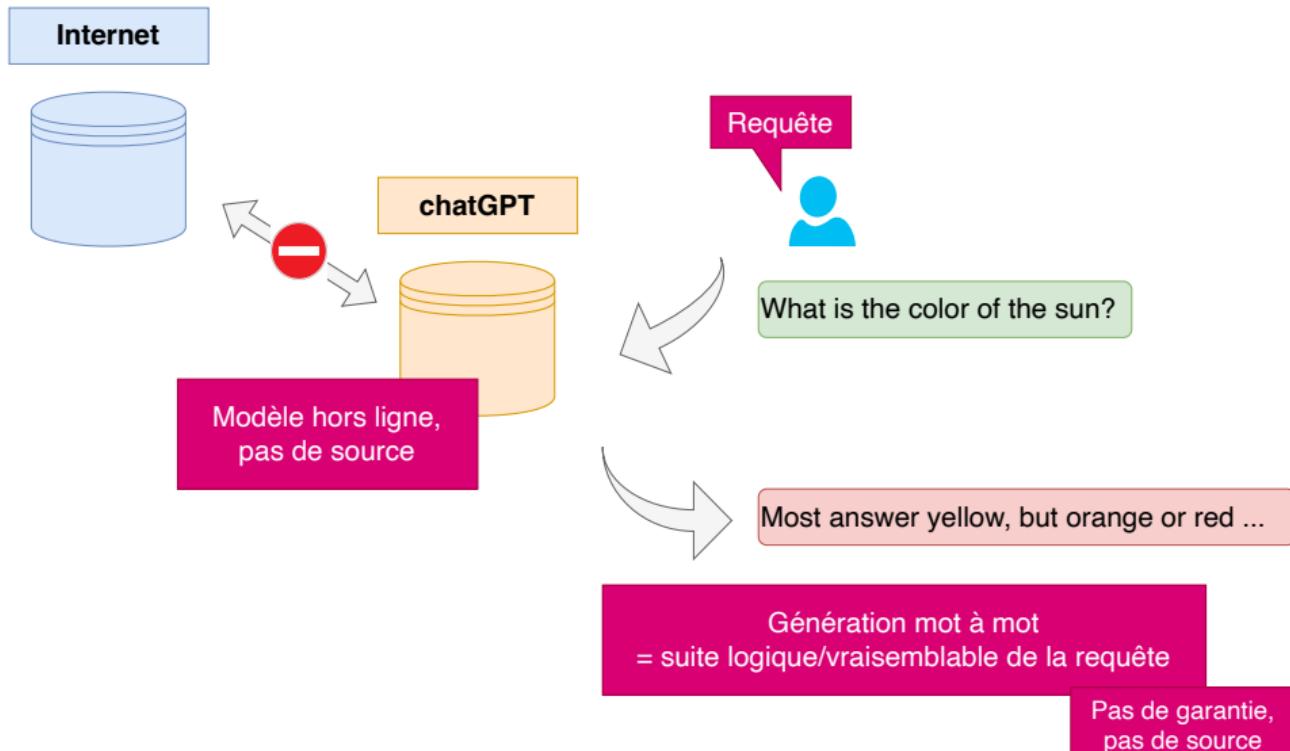
## Résultats sourcés



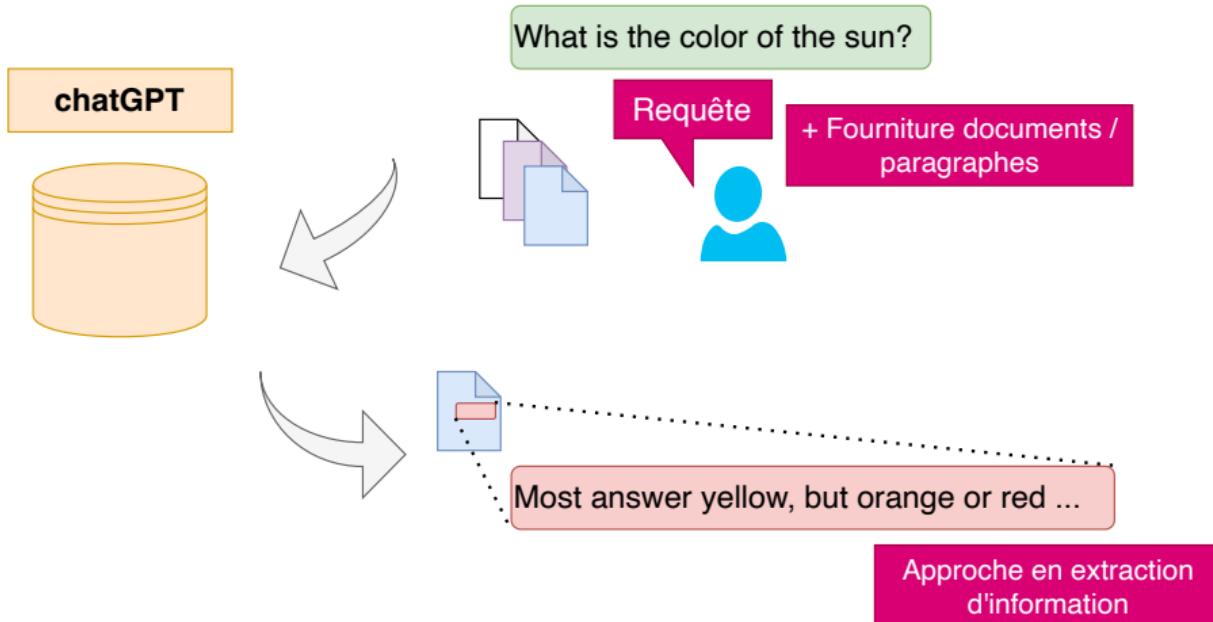
## Graphe de connaissances vérifiées

# Usage en accès à l'information

- Demander des informations à chatGPT... Un usage étonnant !



# Usage en accès à l'information



- Requête web + analyse, résumé automatique, reformulation, compte-rendus de réunion...
- Limite (actuelle) sur la taille des entrées (2k puis 32k tokens)

# Typologie des risques de l'IA en TAL (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.

# Quelques risques d'usages de chatGPT

## Accès à l'information

- ▶ Construire une bombe, technique de hacking, ... ⇒ comme wikipedia ou internet, plus ou moins
- ▶ Réponses d'examen, ...
- ▶ Accéder à des informations personnelles relativement inaccessibles
- ▶ Remise en cause du droit à l'oubli

## Génération en masse

- ▶ Fabrication de fausses informations ⇒ deepfake en version industrielle
  - ▶ Harcèlement en ligne
- ▶ Dead-bot / Usurpation d'identité ⇒ Apprendre à parler comme quelqu'un;  
Assimiler les données des réseaux sociaux d'une personne

## Automatisation & perte de sens

- ▶ Automatisation des réponses aux devoirs / lettres de motivation / ...
  - ▶ Reformulation chatGPT = peu détectable
- ▶ ... Mais automatisation de la correction / évaluation des candidats / ...

# Machine-learning & biais

## Biais dans les données

- ▶ Biais dans les réponses



Oreilles pointues,  
moustaches, texture de poils  
=  
Chat



Homme blanc, +40ans,  
costume  
=  
Cadre supérieur

Le machine-learning est basé sur l'extraction de biais statistiques...

⇒ Lutter contre les biais = forcer l'algorithme à la main

# Machine-learning & biais

## Biais dans les données

### ► Biais dans les réponses



The nurse and the doctor

⇒

L'infirmière et le docteur

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

Le machine-learning est basé sur l'extraction de biais statistiques...

⇒ Lutter contre les biais = forcer l'algorithme à la main

# Correction des biais & ligne éditoriale

## Correction des biais:

- ▶ Sélection de données spécifiques, ré-équilibrage
- ▶ Censure de certaines informations
- ▶ Censure des résultats de l'algorithme

⇒ Travail éditorial...



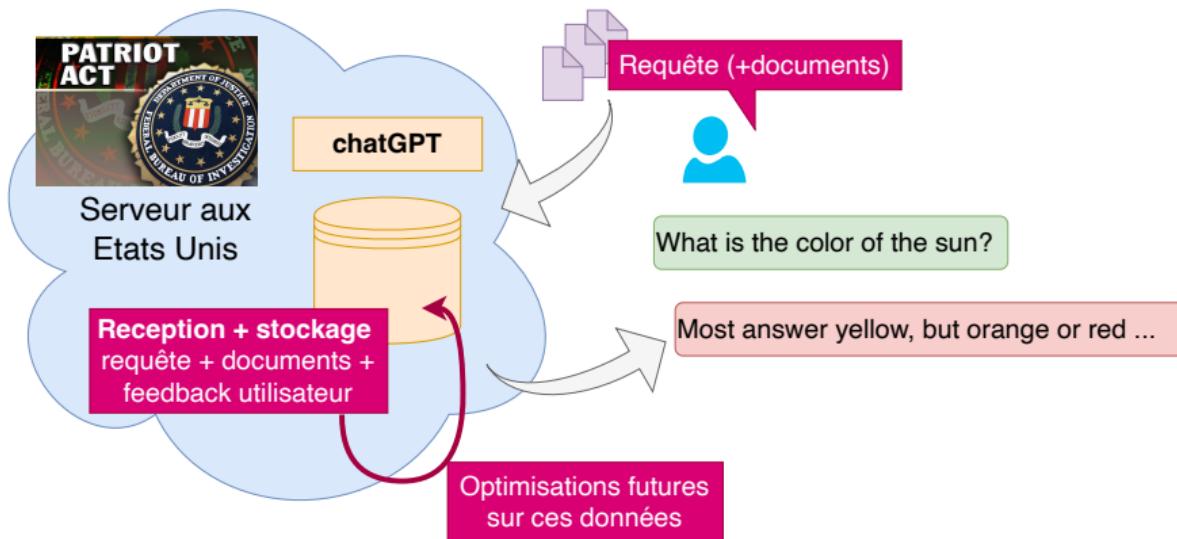
Effectué par qui?

- ▶ Experts métiers / cahier des charges
- ▶ Ingénieurs, lors de la conception des algorithmes
- ▶ Groupe éthique, lors de la validation des résultats
- ▶ Groupe communication / réaction aux utilisateurs

⇒ Quelle légitimité? Quelle transparence? Quelle efficacité?



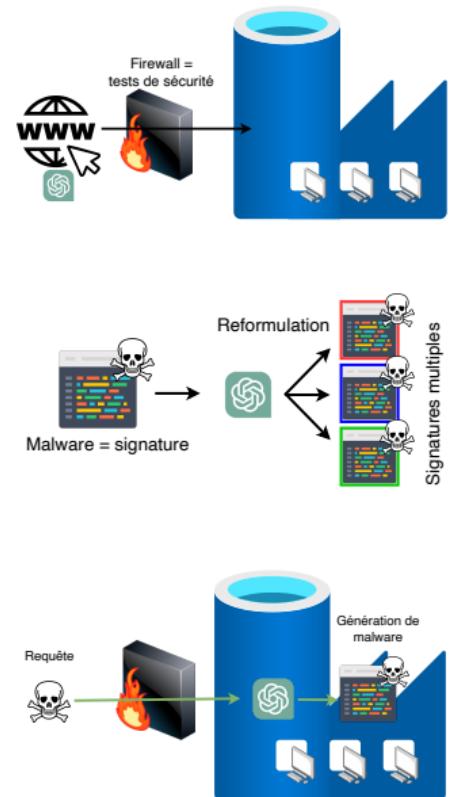
# Fuite(s) de données



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

# Problèmes de sécurité

- ▶ Plug-ins ⇒ Souvent des failles de sécurité béantes pour l'utilisateur
  - ▶ Accès mail / transferts d'information sensibles etc...
- ▶ Problème de gestion pour les entreprises
  - ▶ Sécurisation des (très) gros fichiers
- ▶ Opportunité de multiplication des signatures de malware
  - ▶ ≈ reformulation logicielle
- ▶ Nouveaux problèmes !
  - ▶ Génération directe de malware



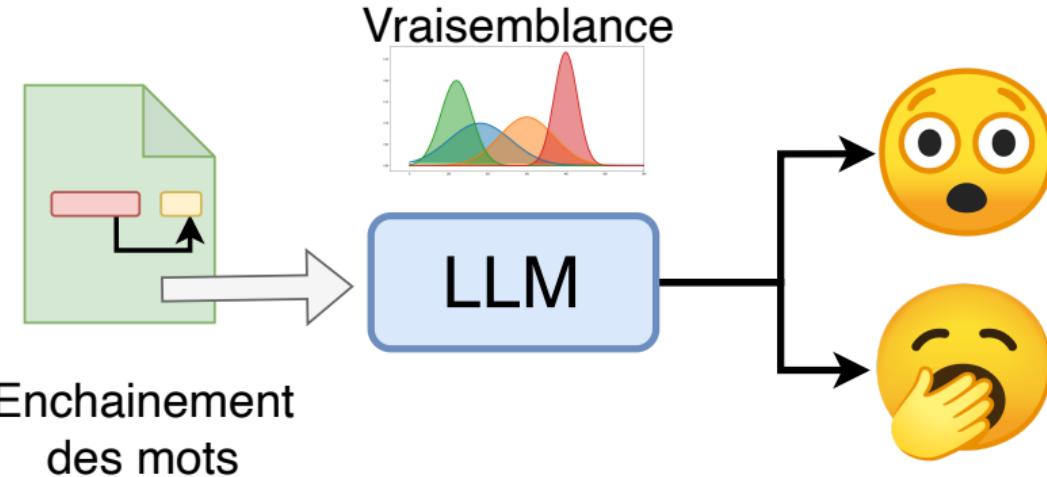
# Détection des textes générés par chatGPT

- ▶ **Classifieur** de texte (comme pour n'importe quel auteur)
  - ▶ Détection des biais dans le choix des mots / tournure de phrases
- ▶ Caractérisation de la **vraisemblance** des textes ([OpenAI](#), [GPTZero](#))
  - ▶ Hyper-fluidité des phrases, sur-abondance de connecteurs logiques
  - ▶ Modèle de langue = statistique ⇒ mesure entre distribution (**perplexité**)
- ▶  $\delta$ -**vraisemblance** sur des textes perturbés ([DetectGPT](#))
- ▶ Evaluation vs estimation de la vraisemblance
- ▶ [chatGPT](#) devrait rapidement intégrer des **empreintes** dans les textes générés

Détecteurs ⇒ < 100% de détection

- + niveau de confiance dans la détection
- depend de la longueur du texte et des modifications apportées
- ≈ détecte des bouts de wikipedia (*chatGPT = stochastic parrot*)

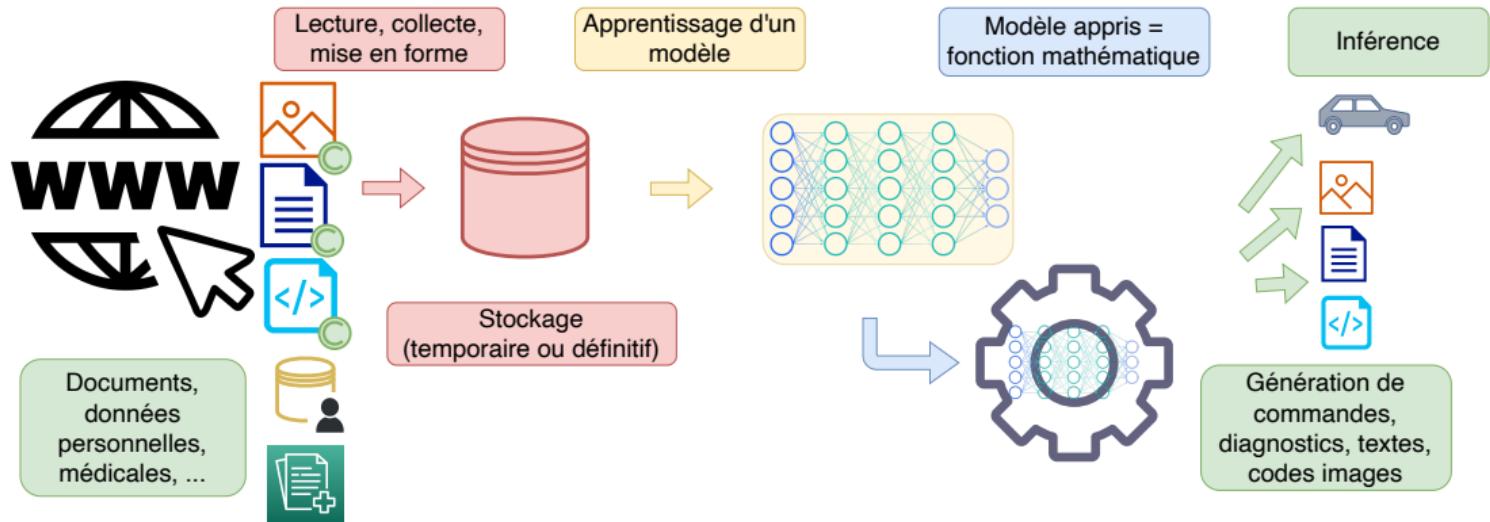
# Détection des textes *utilisés par chatGPT*



Enchainement  
des mots

- Corpus fermés ⇒ enjeu de **détection des textes utilisés en apprentissage**
- Détection de la **vraisemblance/surprise des enchainements de mots observés**

# Risques/Questions juridiques



Droit d'auteur, droit des bases de données

Droit de collecte, droit de copie, consentement

Droit d'utiliser les données dans un algorithme

Modèle = émanation des données ?



- Reproductions d'extraits non traçables
- Régulation des usages
- Responsabilité en cas d'erreur

# Questions économiques

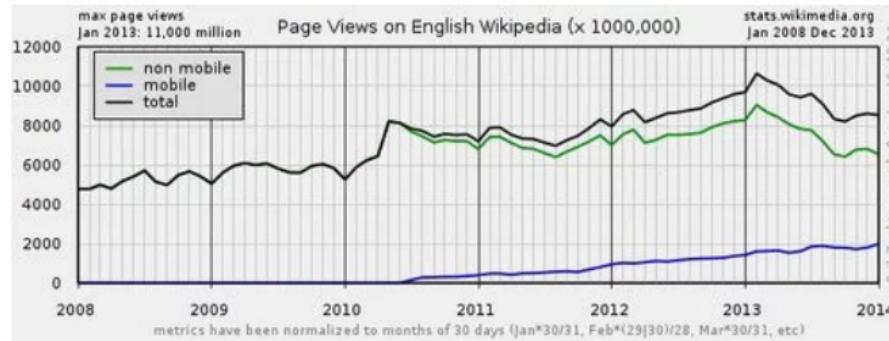
## Financement des sources d'information = publicité

- ▶ Publicité ⇔ **visites** des internautes
- ▶ Google knowledge graph (2012) ⇒ – de visites, – de revenu
- ▶ chatGPT = encodage des informations du web... ⇒ beaucoup moins de visites?

⇒ Quel modèle économique pour les sources d'information avec chatGPT?

### Google Knowledge Graph aurait causé une baisse du trafic de Wikipedia en 2013

Guillaume Belfiore  
13 janvier 2016 à 10h16



# IA, domaines impactés & ressources humaines

## (1) Ne pas se tromper de sujet / domaine

**Numérisation / digitalisation** ≠ **Intelligence Artificielle**

- ▶ **Supermarché:** Automatisation des caisses, de l'étiquetage des prix...
- ▶ **Secrétariats délocalisés:** Doctolib, ...
- ▶ **Dématérialisation des courriers**
- ▶ **Service public:** Sites webs, FAQ, échanges numériques
- ▶ **Jeux vidéos**
- ▶ **Manutention:** Robotisation 80'⇒2000'
- ▶ **Tous domaine:** Chatbot <2020'
- ▶ **Logistique:** Optimisation
- ▶ **Services en ligne:** Traduction, filtre spam, systèmes de recommandation
- ▶ **Publicité:** ciblage, personnalisation
- ▶ **Sécurité:** reconnaissance faciale
- ▶ **Manutention:** Robotisation >2010'
- ▶ **Tous domaine:** Chatbot >2020'

# IA, domaines impactés & ressources humaines

## (2) Ne pas se tromper d'époque

**IA aujourd'hui ≠ IA demain**

- ▶ Publicité en ligne

La publicité représente entre 80 et 90% du CA de Google

- ▶ Reconnaissance faciale
- ▶ Service en ligne : traduction, filtre spam, systèmes de recommandation
- ▶ Logistique
- ▶ Robotique (émergent)
- ▶ chatbot (émergent)

- ▶ Véhicule autonome
- ▶ Aide au diagnostic médical
- ▶ Traitement automatique des mails, CR réunion, rédaction des courriers
- ▶ Bio-informatique: molécule
- ▶ Nouvelle agriculture
- ▶ Conception de structures (optimisation matériaux/résistance...)
- ▶ Jeux vidéos? Robotique avancée?
- ▶ **Ressources Humaines**

# Les défis en Ressources Humaines

## 1 Nouveaux métiers

- ▶ Grande tension sur le marché de l'emploi
- ▶ Gérer le turn-over & les grilles de salaire
- ▶ Secteur en cours de structuration
  - ▶ Far-west en 2010... Loin d'être stable en 2020

## 2 Transformation des métiers existants

- ▶ Impact très large: secteurs primaire, secondaire & tertiaire
  - + Tous les niveaux hiérarchiques
- ▶ Besoins de formations à la fois très large et très urgent
  - ▶ Révolution industrielle la plus rapide de l'histoire
- ▶ Impact sur la productivité
  - ▶ Pression sur les formations
  - ▶ Question sur les effectifs

## 3 Métiers en voie de disparition

- ▶ Très peu à court (et moyen) termes (!)
- ▶ IA manque de robustesse / ≠ adaptée tâches critiques

**L'IA est un assistant puissant mais pas un bon remplaçant**

# CONCLUSION

# Souveraineté numérique: l'ensemble de la chaîne

## Construction du modèle de base

## Raffinement du modèle

## Exploitation du modèle

**Maitrise des données**

- Collecte/équilibrage
- Nettoyage

**Entrainement**

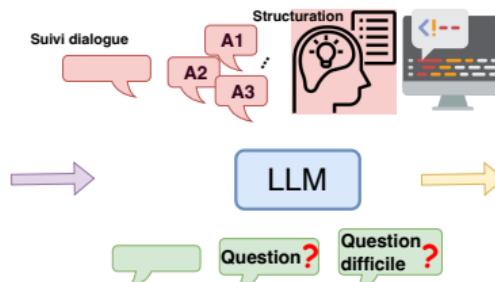
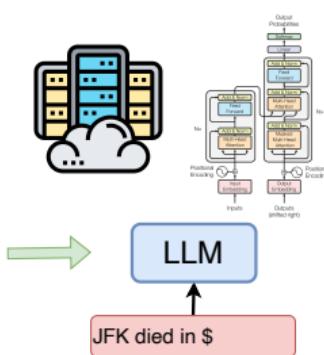
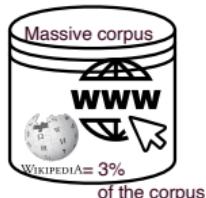
- Puissance machine (milliers de GPU)
- Architecture/recherche ML

**Maitrise & construction des données**

- interactions humaines +++
- prix des données
- spécialisation à la demande

**Optimisation / Limitation du coût**

- compétence MLOps
- déploiement local



# Une multitude de métiers

## Architecte/Manageur de la donnée

- Gère la donnée et son infrastructure (stockage)

## Ingénieur de la donnée / Ingénieur Data

- Met à jour et interroge la donnée

## Analyste de la donnée

- Visualise la données (graphiques)
- Analyse statistiquement les tendances

## Data Scientist

- Interroge la donnée
- Implémente les algorithmes
- Visualise la donnée

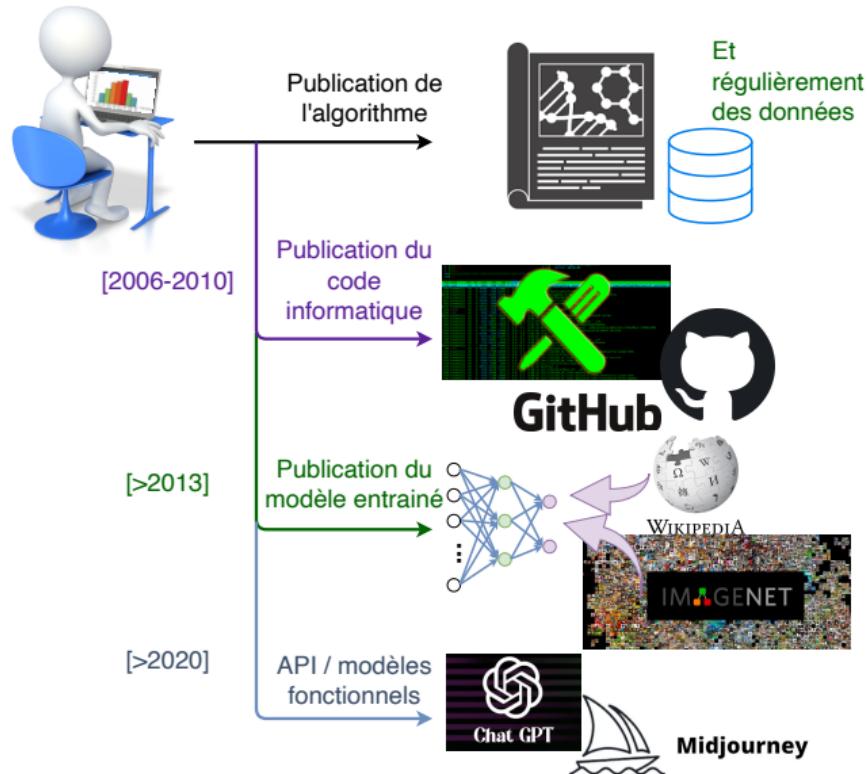
## Ingénieur Machine Learning

- Crée ou sélectionne les algorithmes de machine learning
- Évalue les algorithmes
- Industrialise les solutions

## Utilisateur / Ingénieur Prompt

- Crée des prompt = nouvelles applications

# Une multitude de métiers



# Facteurs d'acceptabilité des IA générative

## 1 Utilitarisme :

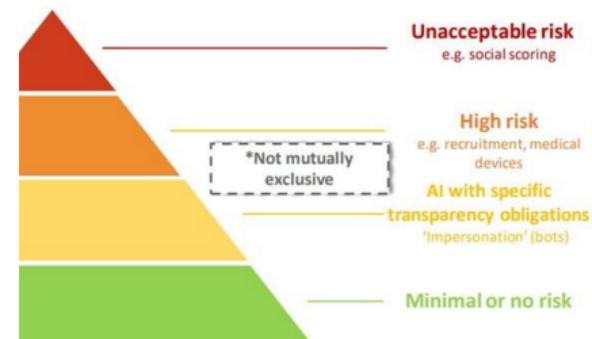
- ▶ Performances (facteur d'acceptation de chatGPT)
- ▶ Fiabilité / auto-évaluation

## 2 Non-dangerosité :

- ▶ Biais / correction
- ▶ Transparence (ligne éditoriale, confusion humain/machine)
- ▶ Implémentation fiable
- ▶ Souveraineté (?)
- ▶ Régulation (AI act)
  - ▶ Eviter les applications dangereuses

## 3 Savoir-faire :

- ▶ FormationS (usage/développement)



# chatGPT: une simple étape

## ► Cout d'entraînement & réglages

4-5 Million d'euros / apprentissage ⇒ chatGPT est **mal appris !**

## ► Efficacité sur les données

chatGPT > 1000x les lectures d'un homme dans une vie

## ► Repérer les entités, citer les sources

Ancrage des réponses dans les bases de connaissances

Ancrage des réponses dans des sources



Sam Altman ✅  
@sama

ChatGPT launched on wednesday. today it crossed 1 million users!

8:35 AM · Dec 5, 2022

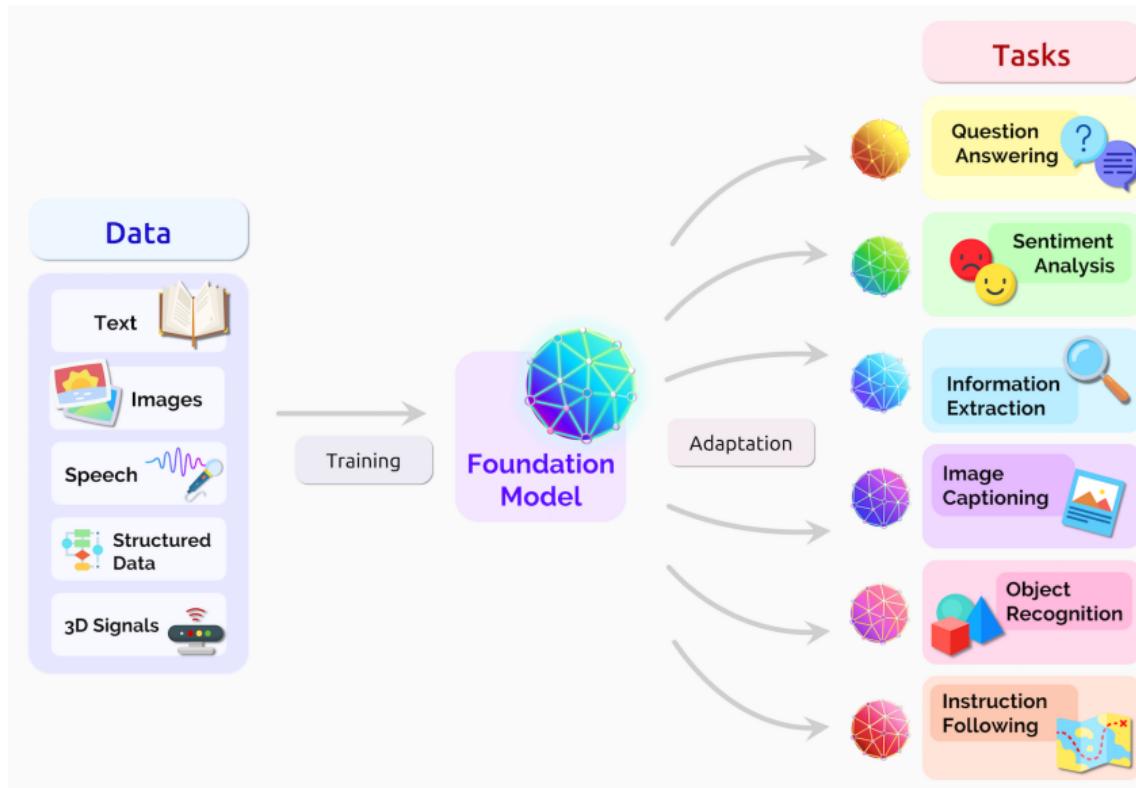
3,457 Retweets 573 Quote Tweets 52.8K Likes

...

► Multiplication des initiatives: GPT, LaMBDA, PaLM, BARD, BLOOM, Gopher, Megatron, OPT, Ernie, Galactica...

► Implication du public,  
impact sur l'accès à l'information

# Vers des modèles de fondation plus larges?



# EXEMPLES

# Encyclopédie statique (+recherche avancée !)

[S. Pajak]

SE

C'est quoi le taux marginal de substitution en microéconomie ?



Le taux marginal de substitution (TMS) est une notion importante en microéconomie qui mesure la valeur que quelqu'un attribue à un bien ou un service par rapport à un autre.



Plus précisément, le TMS mesure la quantité de bien A qu'une personne est prête à abandonner pour obtenir une unité supplémentaire de bien B.

Le TMS peut être utilisé pour déterminer la courbe d'indifférence. Une représentation graphique de tous les paniers de consommation possibles sur une courbe d'indifférence est une droite qui passe par toutes les mêmes valeurs pour une personne. Le TMS est la pente de cette droite.

Le TMS est un concept important en microéconomie car il explique comment les individus prennent des décisions d'achat et de consommation. Il est également utilisé pour déterminer les prix et les quantités dans un marché.

SE

Qu'est-ce que le dol en droit français ?



Le dol est un concept juridique qui désigne l'intention de tromper ou de frauder. Il peut être commis de manière active (c'est-à-dire en trompant ou en fraudant volontairement) ou de manière passive (c'est-à-dire en laissant volontairement une personne se tromper ou se faire frauder).



En droit français, le dol peut être invoqué dans le cadre de différentes situations, telles que l'annulation d'un contrat, la nullité d'une donation ou encore la responsabilité civile délictuelle. Dans ces cas, la preuve du dol est essentielle pour obtenir gain de cause.

# Encyclopédie statique (+recherche avancée !)

[S. Pajak]

Question de comparaison

Réponse à la question  
spécifique



Définitions « simples »

SE Dans le système juridique français, lequel est supérieur entre un arrêté et un décret ? 

 Dans le système juridique français, un décret est supérieur à un arrêté.  

Un décret est un acte administratif pris par le gouvernement ou par le Premier ministre, qui a force de loi. Il est destiné à réglementer des matières de droit public et doit être publié au Journal officiel de la République française pour entrer en vigueur.

Un arrêté est un acte administratif pris par le gouvernement, un membre du gouvernement ou un haut fonctionnaire de l'État, qui a force de loi dans un domaine précis. Il est destiné à mettre en œuvre une politique ou une réglementation définie par un décret et doit être publié au Journal officiel de la République française pour entrer en vigueur.

En résumé, un décret est supérieur à un arrêté dans la hiérarchie des normes en France, car il a une portée générale et réglemente des matières de droit public. Un arrêté, quant à lui, a une portée limitée et met en œuvre une politique ou une réglementation définie par un décret.

# Encyclopédie statique (+recherche avancée !)

[S. Pajak]

SE Explique-moi pourquoi les capitaux propres sont au passif et pas à l'actif ?

Les capitaux propres (aussi appelés fonds propres ou equity en anglais) sont une partie du bilan d'une entreprise qui représente les ressources financières dont elle dispose. Ils sont composés des fonds apportés par les propriétaires de l'entreprise (les actionnaires) ainsi que de tous les bénéfices non distribués de l'entreprise.

Les capitaux propres sont considérés comme étant au passif du bilan, car ils représentent une dette de l'entreprise envers ses propriétaires. En effet, les actionnaires ont investi de l'argent dans l'entreprise en achetant des actions et ils ont droit à une partie des bénéfices de l'entreprise. C'est pourquoi les capitaux propres sont considérés comme une dette de l'entreprise envers ses propriétaires.

En revanche, l'actif du bilan représente les ressources financières de l'entreprise qui sont utilisées pour générer des bénéfices, comme les biens et les équipements de l'entreprise, les comptes clients, les stocks, etc. Les capitaux propres ne sont pas considérés comme faisant partie de l'actif, car ils ne sont pas directement utilisés pour générer des bénéfices.

Question problématisée

Définition « simple »

Réponse personnalisée, qui traite la problématique dans la question



# Pédagogie

[S. Pajak]

SE

Comment on fait une regression OLS simple en R de y sur x



Pour effectuer une régression OLS simple en utilisant R, vous pouvez utiliser la fonction `lm()` (qui signifie "linear model" en anglais) de R. Cette fonction permet de créer un modèle de régression linéaire.

Voici un exemple de comment utiliser la fonction `lm()` pour créer un modèle régression OLS simple pour y en fonction de x :



```
R                                         Copy code

# Tout d'abord, il faut charger les données
x <- c(1, 2, 3, 4, 5)
y <- c(2, 4, 6, 8, 10)

# Ensuite, on utilise la fonction lm() pour créer un modèle de régression OLS
model <- lm(y ~ x)

# On peut maintenant afficher les résultats du modèle
summary(model)
```

La sortie affichera les informations sur le modèle, comme les coefficients, les erreurs standard, les t-values et les p-values.

Il est important de noter que cet exemple est très simpliste, il est souvent nécessaire de prétraiter les données (par exemple en enlevant les outliers ou les valeurs manquantes) et de vérifier les hypothèses du modèle (linéarité, normalité, homoscédasticité, indépendance des erreurs) avant de procéder à la régression.

## ***U.S. Pours Money Into Chips, but Even Soaring Spending Has Limits***

SE

Please explain "Soaring Spending" in the sentence "U.S. Pours Money Into Chips, but Even Soaring Spending Has Limits"



In the context of this sentence, "soaring spending" refers to the act of spending a large amount of money. The word "soaring" suggests that the spending is increasing rapidly or at a high rate.

The sentence mentions that the U.S. is pouring money into chips, which means that it is investing a large amount of money in the development and production of chips. However, the sentence also suggests that this spending has limits, which means that there is a maximum amount of money that the U.S. is willing or able to invest in chips.

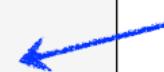
Overall, the phrase "soaring spending" refers to the high level of investment that the U.S. is making in chips, but it also implies that this level of investment is not unlimited.



Définition



Reformulation  
complète du titre



(d'ailleurs fausse en partie)

# Détection du plagiat

[S. Pajak]

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 sociaux ou environnementaux.

Tout cocher :

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

Plagiat	Date	Mots	Ko	Action
Plagiat Def 2 #4483eb	07/01/2023 19:18 par vous	122 mots	19,47 ko	<a href="#">Plus d'infos</a>
Plagiat Def 1 #f90ff3	07/01/2023 19:16 par vous	135 mots	16,78 ko	<a href="#">Plus d'infos</a>

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

Plagiat Def 2 #4483eb 07/01/2023 19:18 par vous | 122 mots | 19,47 ko | [Plus d'infos](#)

Plagiat Def 1 #f90ff3 07/01/2023 19:16 par vous | 135 mots | 16,78 ko | [Plus d'infos](#)

## Définition de Wikipedia

# Hallucination

V

Peux tu me donner une bibliographie des VAE sur les séries temporelles?



Je m'excuse pour la confusion. Voici donc une sélection de références sur les Variational Autoencoders (VAE) pour les séries temporelles :

1. "Variational Autoencoder for Deep Learning of Images, Labels and Sequences" par Zhao et al. (2017) - Cette étude présente une approche VAE pour la modélisation de séquences temporelles en utilisant une architecture de réseaux de neurones récurrents.
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 pour la phase d'entraînement.
3. "Generative Modeling for Time Series:

## Variational recurrent auto-encoders

O Fabius, JR Van Amersfoort - 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 ...

☆ Enregistrer ☰ Citer Cité 302 fois Autres articles Les 2 versions ☰

- Souvent sur des connaissances rares ⇒ difficile à mémoriser

# Assistant personnel intelligent

[S. Pajak]

SE Ecris un mail à l'administration pour décaler mon cours d'économie du lundi 13h au jeudi 9h. 1

Objet : Demande de décalage de cours d'économie thumb up thumb down

Monsieur ou Madame,

Je vous écris aujourd'hui pour vous demander de bien vouloir décaler mon cours d'économie du lundi 13h au jeudi 9h.

En effet, j'ai un empêchement le lundi à cette heure-là et je suis dans l'impossibilité de suivre mon cours à cette heure. Le jeudi 9h serait une plage horaire beaucoup plus adaptée à mon emploi du temps.

Je vous remercie par avance pour votre compréhension et reste à votre disposition pour toutes informations complémentaires.

Dans l'attente de votre réponse, veuillez agréer, Monsieur ou Madame, l'expression de mes salutations distinguées.

Cordialement,

[Votre nom]

# Assistant personnel intelligent

[S. Pajak]

SE

Ecris un mail à l'administration pour décaler mon cours d'économie du lundi 13h au jeudi 9h.

1



Objet : Demande de décalage de cours d'économie



Monsieur ou Madame,

Je vous écris aujourd'hui pour vous demandez de bien vouloir décaler mon cours d'économie du lundi 13h au jeudi 9h.

En effet, j'ai un empêchement le lundi à cette heure-là et je suis dans l'impossibilité de suivre mon cours à cette heure. Le jeudi 9h serait une plage horaire beaucoup plus adaptée à mon emploi du temps.

Je vous remercie par avance pour votre compréhension et reste à votre disposition pour toutes informations complémentaires.

Dans l'attente de votre réponse, veuillez agréer, Monsieur ou Madame, l'expression de mes salutations distinguées.

Cordialement,

[Votre nom]

SE

Ecris un mail aux étudiants pour décaler mon cours d'économie du lundi 13h au jeudi 9h.

2



Objet : Modification de l'horaire du cours d'économie



Chers étudiants,

Je vous informe que le cours d'économie prévu le lundi 13h a été décalé au jeudi 9h.

Je vous prie de bien vouloir noter cette modification dans vos agendas et de vous présenter au cours à cette nouvelle heure.

Je vous remercie pour votre compréhension et vous dis à très bientôt.

Cordialement,

[Votre nom]

# Assistant personnel intelligent

[S. Pajak]

SE

Ecris un mail de candidature spontanée à un stage d'assistant comptable en cabinet.



Objet : Candidature pour un stage d'assistant comptable



Monsieur ou Madame,

Je vous écris aujourd'hui pour vous présenter ma candidature pour un stage d'assistant comptable au sein de votre cabinet.

Actuellement étudiant(e) en comptabilité et finance, je suis à la recherche d'un stage de fin d'année afin de mettre en pratique les connaissances acquises lors de mes études et de découvrir le monde de la comptabilité en cabinet.

Je suis particulièrement intéressé(e) par votre cabinet pour sa réputation et la qualité de ses services. Je suis persuadé(e) que cette expérience serait pour moi une occasion unique de développer mes compétences et de m'épanouir professionnellement.

Je suis rigoureux, organisé et j'aime travailler en équipe. Je suis également doté(e) d'un excellent relationnel et je suis persuadé(e) que je pourrais apporter une valeur ajoutée à votre équipe.

Je vous remercie par avance pour l'attention que vous porterez à ma candidature et me



# Assistant personnel intelligent

[S. Pajak]

SE

Ecris un mail de candidature spontanée



Objet : Candidature pour un stage d'ass

Monsieur ou Madame,

Je vous écris aujourd'hui pour vous présenter mon dossier pour un stage en comptabilité au sein de votre cabinet.

Actuellement étudiant(e) en comptabilité

fin d'année afin de pouvoir mieux me préparer et de découvrir le monde professionnel.

Je suis particulièrement intéressé par les services que vous proposez. Je suis convaincu que votre cabinet est unique de développer une telle expertise.

Je suis rigoureux, attentionné et j'ai des excellentes relations avec mes collègues et ma

votre équipe.

Je vous remercie pour l'accord que vous m'accordez et je tiens à votre disponibilité.

SE

Propose moi 3 phrases différentes pour dire ce que l'entreprise gagnerait à m'accepter en stage.



1. "Je suis convaincu que ma motivation et mon dynamisme pourraient être un atout pour votre entreprise et que je pourrais apporter une valeur ajoutée à votre équipe." 👍 👎
2. "Ma rigueur et mon organisation me permettraient de m'intégrer rapidement et de prendre en main efficacement les tâches qui me seraient confiées."
3. "Je suis passionné par mon domaine d'études et j'aimerais mettre mes connaissances au service de votre entreprise."

SE

Donne une phrase professionnelle pour dire que j'ai déjà travaillé comme assistant comptable l'année dernière et que j'ai aimé cette expérience.



Voici une phrase professionnelle qui exprime que vous avez déjà travaillé comme assistant comptable l'année dernière et que vous avez apprécié cette expérience :



"J'ai eu l'opportunité de travailler comme assistant comptable l'année dernière et j'ai vraiment apprécié cette expérience. Cela m'a permis de mettre en pratique mes connaissances et de développer de nouvelles compétences, tout en travaillant dans un environnement stimulant et professionnel."

# Assistant personnel intelligent

[S. Pajak]

SE

propose moi un titre pour une présentation qui traite des conséquences de l'IA dans la pédagogie et les enseignements à l'université.



Voici quelques suggestions de titres pour une présentation sur les conséquences de l'IA dans la pédagogie et l'enseignement à l'université :



1. "L'IA et la pédagogie universitaire : nouvelles opportunités pour l'apprentissage"
2. "Comment l'IA peut transformer l'enseignement universitaire"
3. "L'IA et l'université : les défis de la pédagogie de demain"
4. "L'IA et l'enseignement supérieur : une nouvelle ère pour l'apprentissage"
5. "L'apport de l'IA à l'enseignement universitaire : quelles conséquences pour les étudiants et les enseignants?"

# Assistant personnel intelligent

[S. Pajak]

SE

Les titres doivent tenir sur une ligne



Enchaînement à partir du résultat précédent



Voici quelques suggestions de titres pour une présentation sur les conséquences de l'IA dans la pédagogie et l'enseignement à l'université, qui tiennent sur une seule ligne :



1. "L'IA et l'enseignement universitaire: opportunités et défis"
2. "Transformer l'enseignement universitaire avec l'IA"
3. "L'IA et la pédagogie de demain: enjeux et perspectives"
4. "L'IA et l'apprentissage universitaire: nouvelles possibilités"
5. "L'IA et l'enseignement supérieur: révolution ou évolution?"

# Assistant personnel intelligent

[S. Pajak]

SE

des titres plus créatifs stp



Prompt en language naturel



Voici quelques propositions de titres créatifs pour une présentation sur les conséquences de l'IA dans la pédagogie et les enseignements à l'université : like dislike

- "Le cerveau augmenté : l'IA au service de l'apprentissage universitaire"
- "Apprendre autrement grâce à l'IA : nouvelles perspectives pour l'enseignement universitaire"
- "L'IA, une supercherie pour l'enseignement universitaire ?"
- "L'IA, un atout ou un piège pour l'enseignement universitaire ?"
- "AI University : quand l'IA révolutionne l'enseignement universitaire"
- "L'IA, un allié ou un rival pour l'enseignement universitaire ?"

# Assistant personnel intelligent

[S. Pajak]

SE

turn the following citation into BibTex format: Anne Boring. (2017) Gender Biases in Student Evaluations of Teaching, Journal of Public Economics, 145, 27-41.

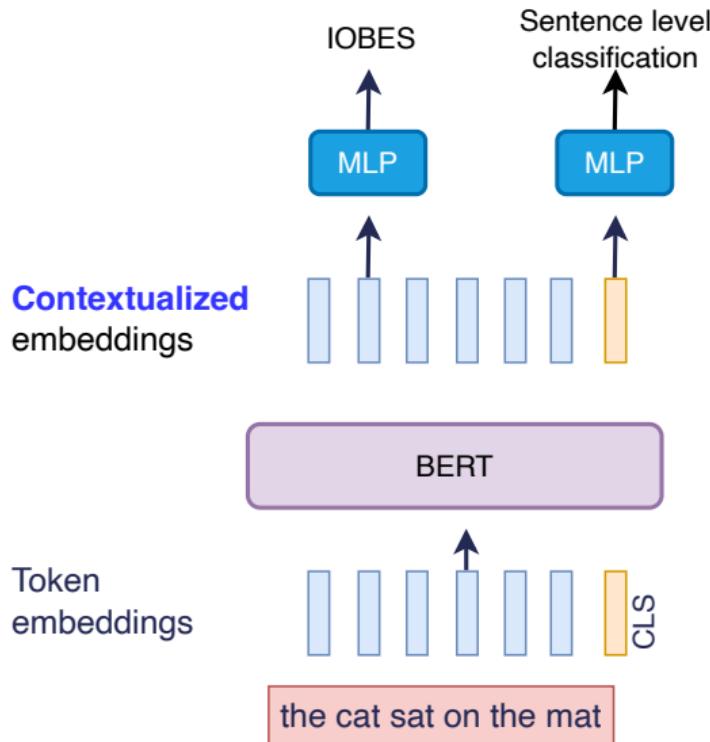


```
@article{boring2017gender,  
title={Gender Biases in Student Evaluations of Teaching},  
author={Boring, Anne},  
journal={Journal of Public Economics},  
volume={145},  
pages={27-41},  
year={2017}  
}
```



# NLP *vs* GÉNÉRATIF

# Extraction de connaissances



Words	BIOES Label
Jane	B-PER
Villanueva	E-PER
of	O
United	B-ORG
Airlines	I-ORG
Holding	E-ORG
discussed	O
the	O
Chicago	S-LOC
route	O
.	O

- ▶ Etat de l'art en NER
- ▶ Efficace sur l'extraction de relation

# Extraction de connaissances

V

Extract entities from the following wikipedia text:

Mozart naît le 27 janvier 1756 à 8 heures du soir au numéro 9 de la Getreidegasse à Salzbourg. Il est le fils de Léopold Mozart, musicien, compositeur et pédagogue originaire d'Augsbourg, ville libre d'Empire, qui occupe alors la fonction de vice-maître de chapelle à la cour du prince-archevêque de Salzbourg, et d'Anna Maria Pertl, sa femme<sup>4</sup>, fille d'un fonctionnaire de la cour de Salzbourg.



From the given text, the entities that can be extracted are:



- Person: Mozart, Léopold Mozart, Anna Maria Pertl
- Location: Salzbourg, Augsbourg, Empire
- Date: 27 janvier 1756
- Time: 8 heures du soir
- Organization: la cour du prince-archevêque de Salzbourg, la cour de Salzbourg

# Extraction de connaissances

V

can you give me the relations between those entities?



Sure, here are the relationships between the entities extracted from the text:



- Mozart is the son of Léopold Mozart and Anna Maria Pertl
- Léopold Mozart is a musician, composer, and pedagogue from Augsburg and was working as the Vice-Kapellmeister at the court of the Prince-Archbishop of Salzburg
- Anna Maria Pertl was the wife of Léopold Mozart and daughter of a court official in Salzburg
- Mozart was born on 27th January 1756 at 8 PM at number 9 of Getreidegasse in Salzburg
- Léopold Mozart and Anna Maria Pertl lived in Salzburg, which was the capital of the Prince-Archbishopric of Salzburg
- Augsburg was Léopold Mozart's birthplace and was a free city of the Empire
- The Prince-Archbishopric of Salzburg was the political jurisdiction where Salzburg was located.

# Question Answering

Les modèles de langue savent répondre aux questions... Dans un certain contexte :)

## Passage Sentence

In meteorology, precipitation is any product of the condensation of atmospheric water vapor that falls under gravity.

## Question

What causes precipitation to fall?

## Answer Candidate

gravity

- ▶ Est-il possible/facile d'extraire des passages avec un moteur de recherche?
- ▶ Est-ce une alternative aux bases de connaissances?
- ▶ Quelles différences avec chatGPT?



Rajpurkar et al., ACL 2018

Know What You Don't Know: Unanswerable Questions for SQuAD



# Question Answering & chatGPT

V

how old is Obama



Barack Obama was born on August 4, 1961, making him 61 years old as of February 2, 2023.



V

what is your source?



I am a machine learning model created by OpenAI, I was trained on a diverse range of texts from the internet. The information I provide is based on the patterns and relationships in that text, and is not based on any one specific source.





# Question Answering & chatGPT

V how old is Obama   



Barack Obama was born on August 4, 1961, making him 61 years old as of February 2,  
2023.



V how old is obama?   



As of 2021, Barack Obama was born on August 4, 1961, so he is 60 years old.



V and today?



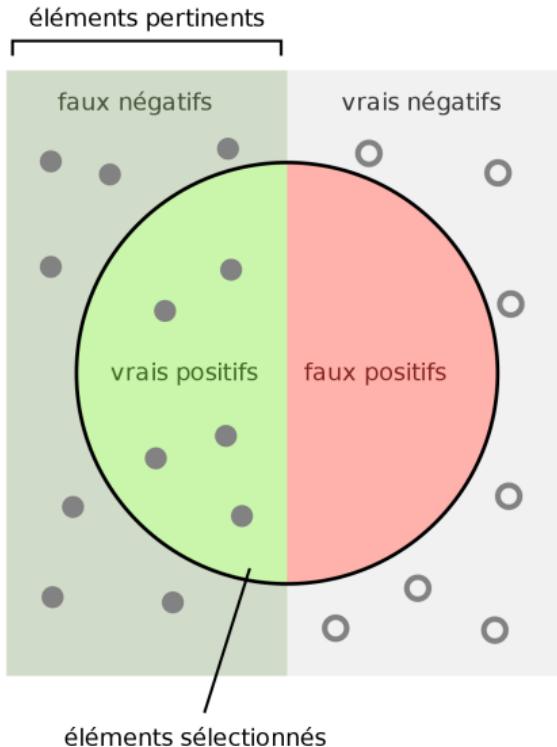
As a language model AI, I don't have real-time access to current dates. However, Barack



Obama was born on August 4, 1961, so if today's date is after August 4th, 2021, he would

MÉTRIQUE

# Précision et rappel



Combien de candidats sélectionnés sont pertinents ?

$$\text{Précision} = \frac{\text{vrais positifs}}{\text{éléments sélectionnés}}$$

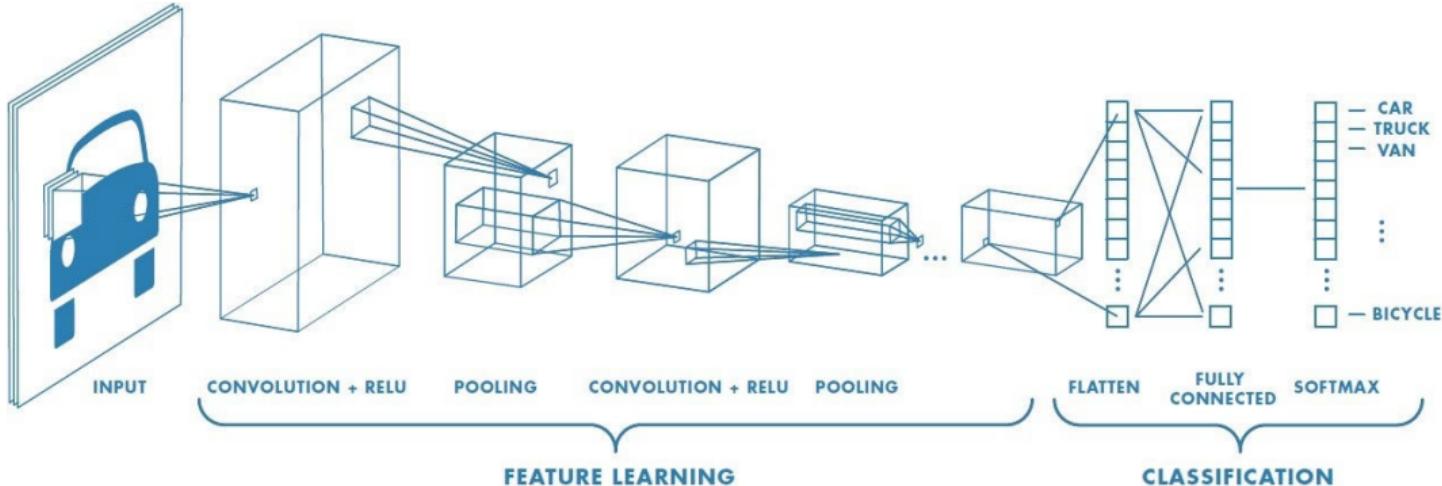
Combien d'éléments pertinents sont sélectionnés ?

$$\text{Rappel} = \frac{\text{vrais positifs}}{\text{éléments pertinents}}$$

- ▶ Précision ↗ = ne pas dire de bêtise
- ▶ Rappel ↗ = couvrir les informations pertinentes
- ▶ F1 = moyenne (harmonique) de la précision et du rappel

# GPT-4: MODÈLE MULTI-MODAUX

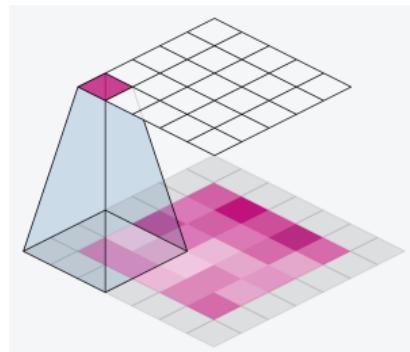
# Apport de l'architecture transformer en vision



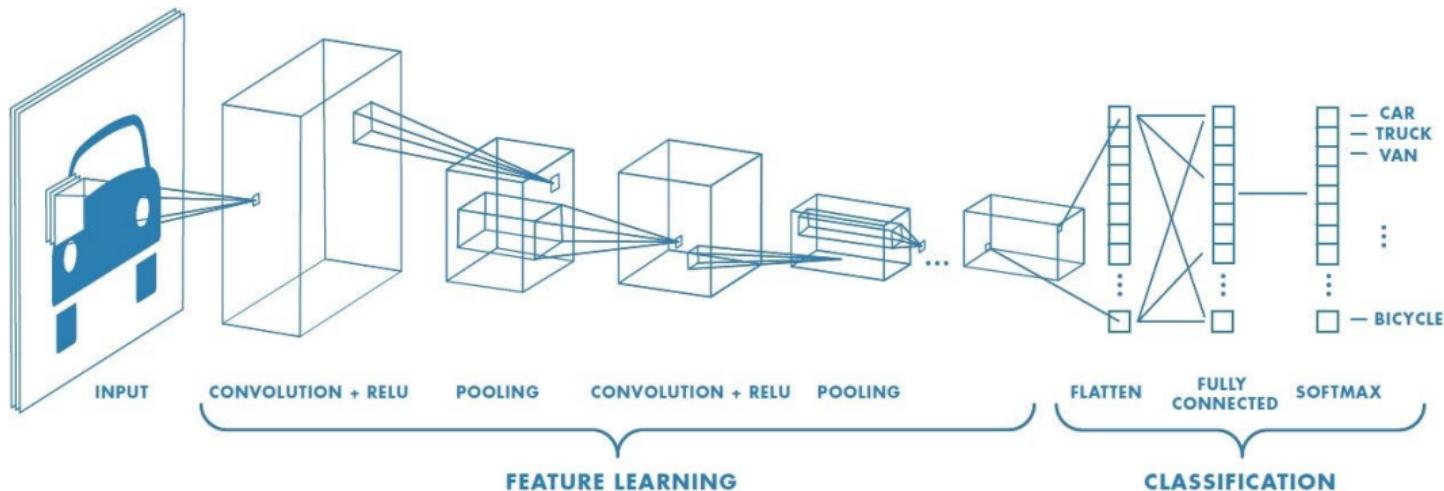
## 1. Convolution

Filtre d'analyse

- ▶ Peu de paramètres
- ▶ Apprentissage des motifs à extraire
- ▶ Agrégation progressive des échelles



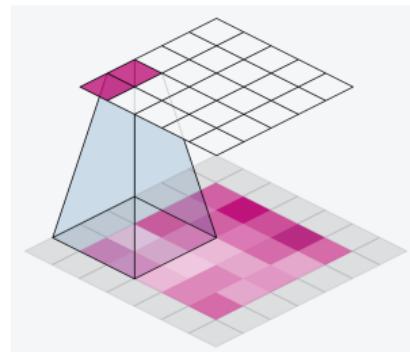
# Apport de l'architecture transformer en vision



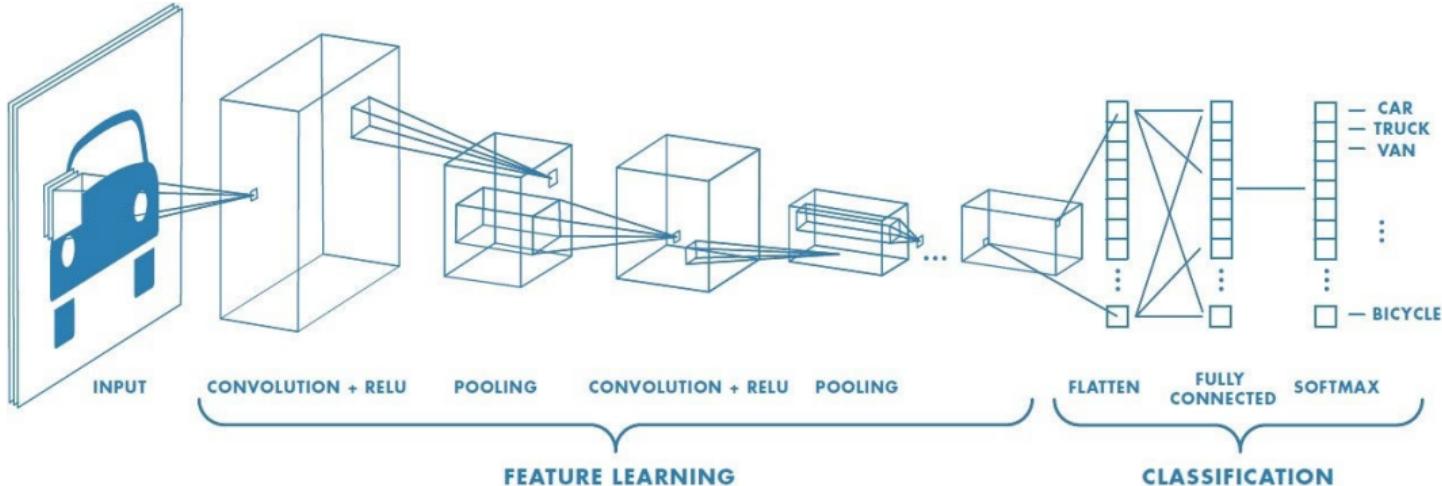
## 1. Convolution

Filtre d'analyse

- ▶ Peu de paramètres
- ▶ Apprentissage des motifs à extraire
- ▶ Agrégation progressive des échelles



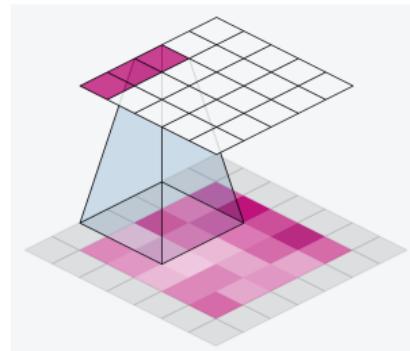
# Apport de l'architecture transformer en vision



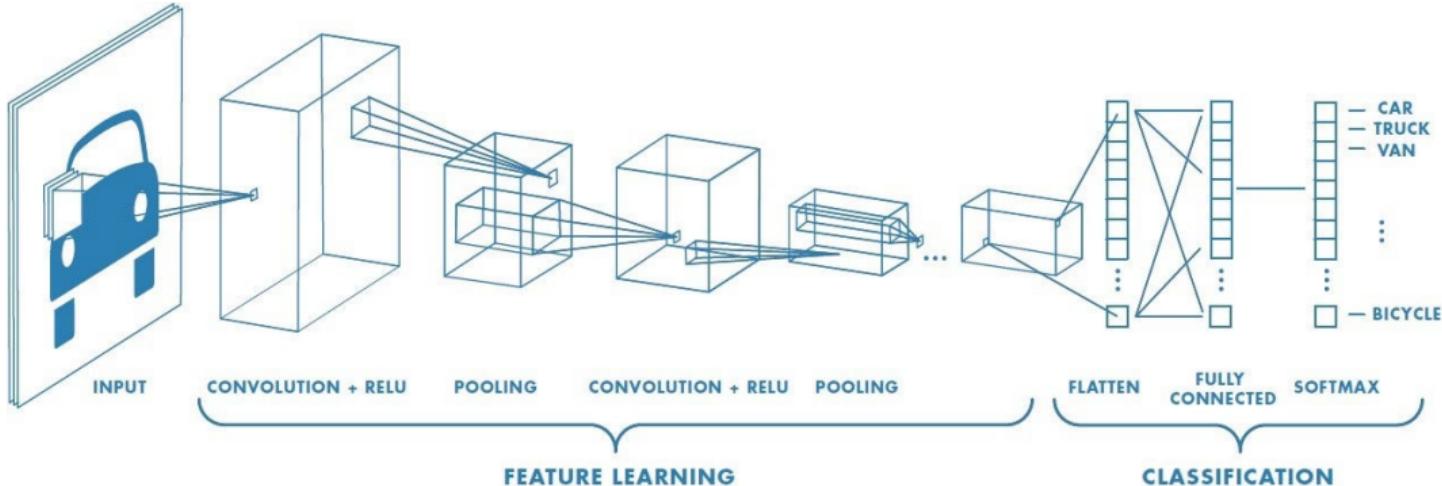
## 1. Convolution

Filtre d'analyse

- ▶ Peu de paramètres
- ▶ Apprentissage des motifs à extraire
- ▶ Agrégation progressive des échelles



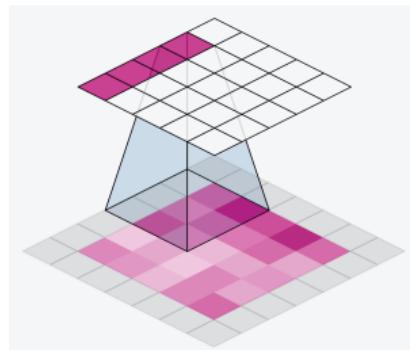
# Apport de l'architecture transformer en vision



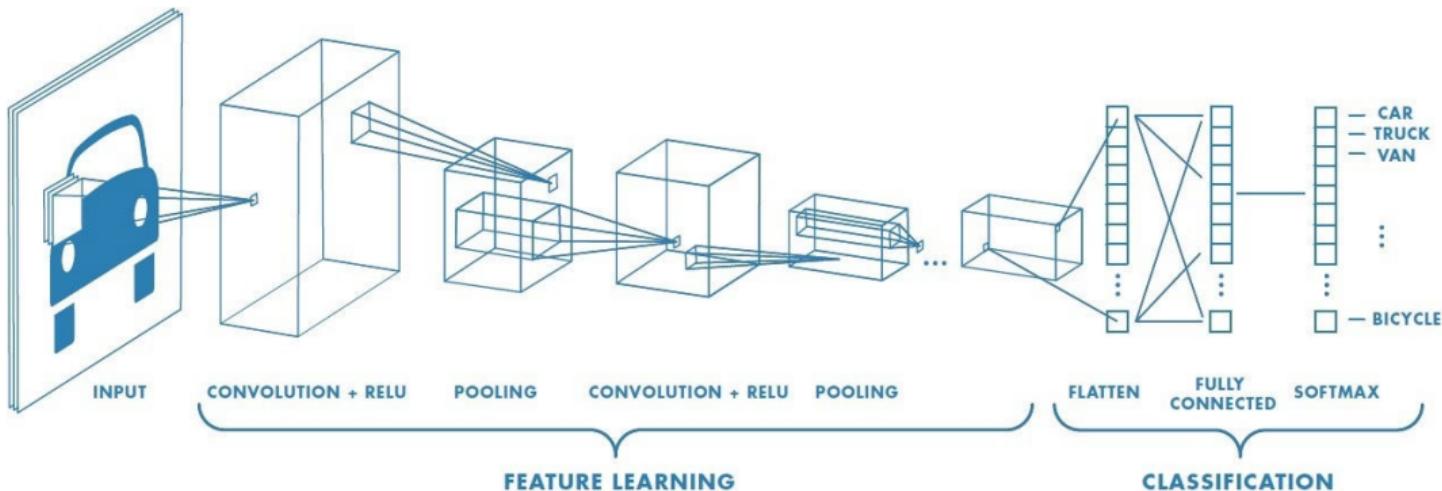
## 1. Convolution

Filtre d'analyse

- ▶ Peu de paramètres
- ▶ Apprentissage des motifs à extraire
- ▶ Agrégation progressive des échelles



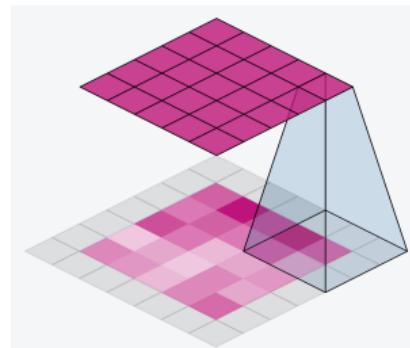
# Apport de l'architecture transformer en vision



## 1. Convolution

Filtre d'analyse

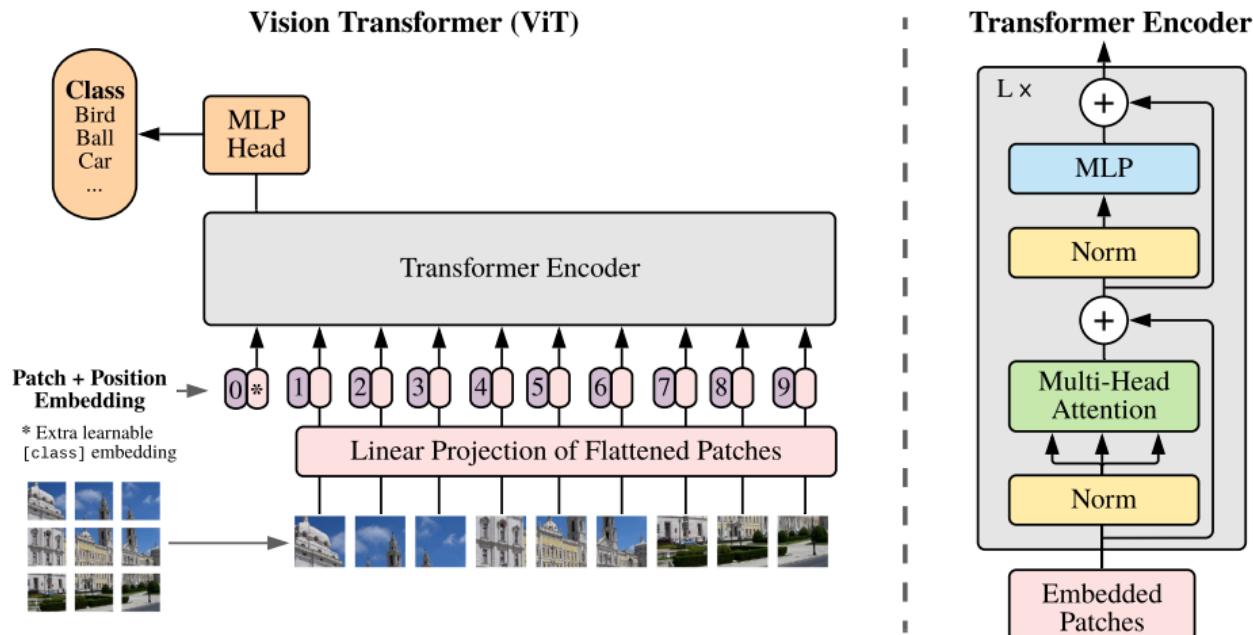
- ▶ Peu de paramètres
- ▶ Apprentissage des motifs à extraire
- ▶ Agrégation progressive des échelles



# Apport de l'architecture transformer en vision

## 2. Transformer

- ▶ Apprentissage des motifs (locaux) à extraire
- ▶ Analyse des interactions entre les portions d'images



# Mieux comprendre les images, mieux modéliser le texte

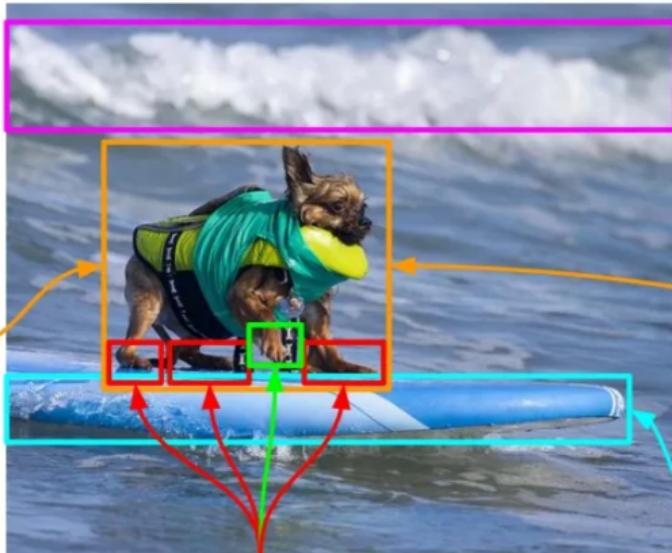
## Visual Grounding

Where does this scene take place?

- A) In the sea. ✓
- B) In the desert.
- C) In the forest.
- D) On a lawn.

What is the dog doing?

- A) Surfing. ✓
- B) Sleeping.
- C) Running.
- D) Eating.



Why is there foam?

- A) Because of a wave. ✓
- B) Because of a boat.
- C) Because of a fire.
- D) Because of a leak.

What is the dog standing on?

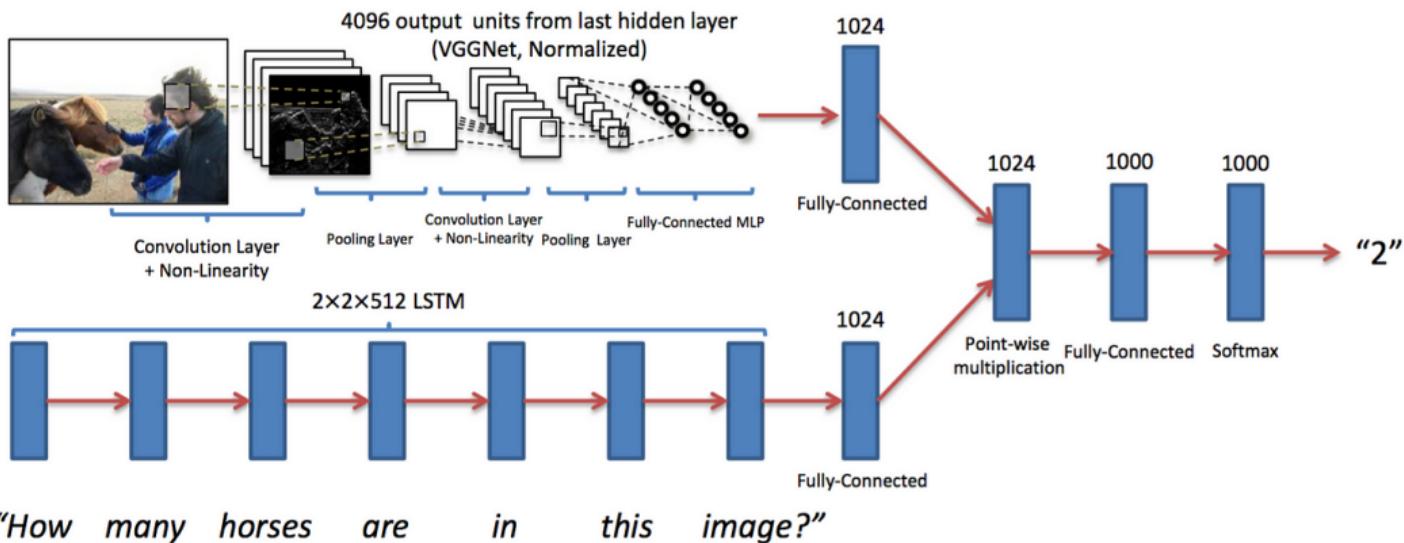
- A) On a surfboard. ✓
- B) On a table.
- C) On a garage.
- D) On a ball.

Comprendre les éléments d'une image  
Intégrer du bon sens (common sense reasoning)

Modéliser le contexte des mots

# Apprendre un système multimodal

L'exemple du VQA: visual question answering



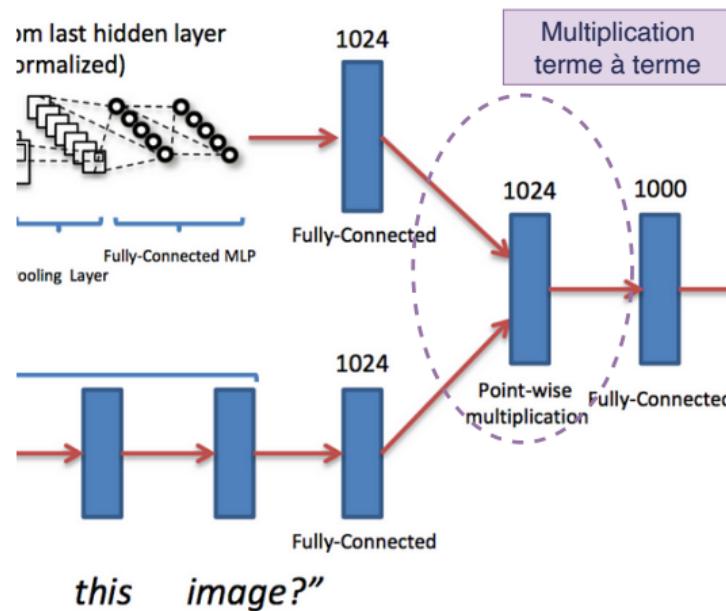
⇒ Rétro-propager l'erreur ⇒ modifier les représentations de mots + analyse de l'image

Compréhension globale / approfondie

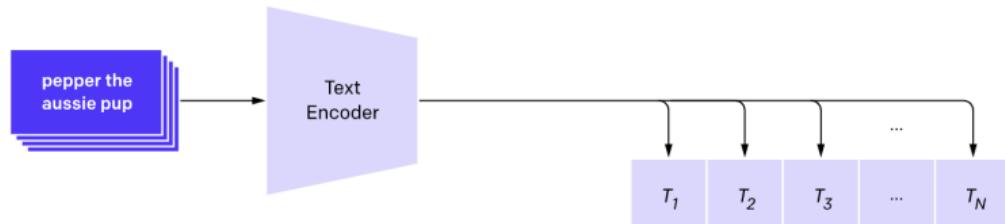
Risque de spécialisation /  
Oubli de la connaissance générale

# Processus de fusion multi-modale

- ▶ Simple multiplication terme à terme des représentations latentes
- ▶ Concaténation et mixture complexe



# Opportunité multi-modale



## Attention multi-modale:

Mot / Portion d'image

⇒ Modélisation fine

⇒ Distinction éléments /  
background

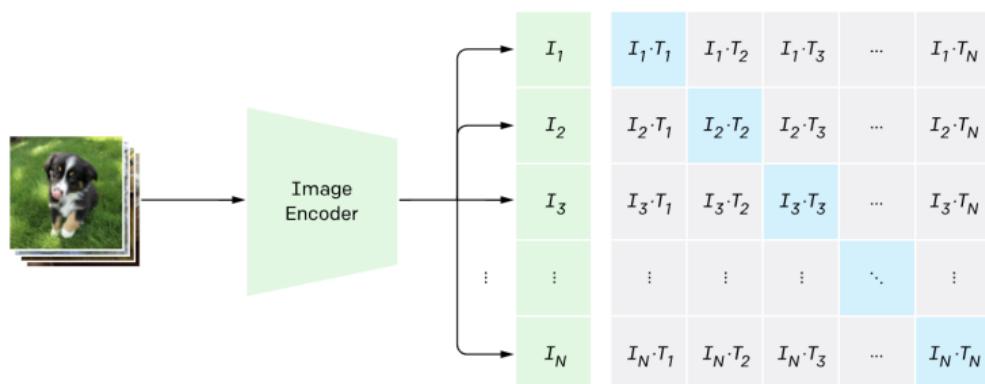


Illustration = encodeur seul (compréhension)

⇒ La même architecture permet d'obtenir du texte ou des images en sortie