

IA GÉNÉRATIVE: FONCTIONNEMENT ET USAGES EN PÉDAGOGIE

13 Juin 2024
Journée IA & pratiques enseignantes

Vincent Guigue
vincent.guigue@agroparistech.fr
<https://vguigue.github.io>

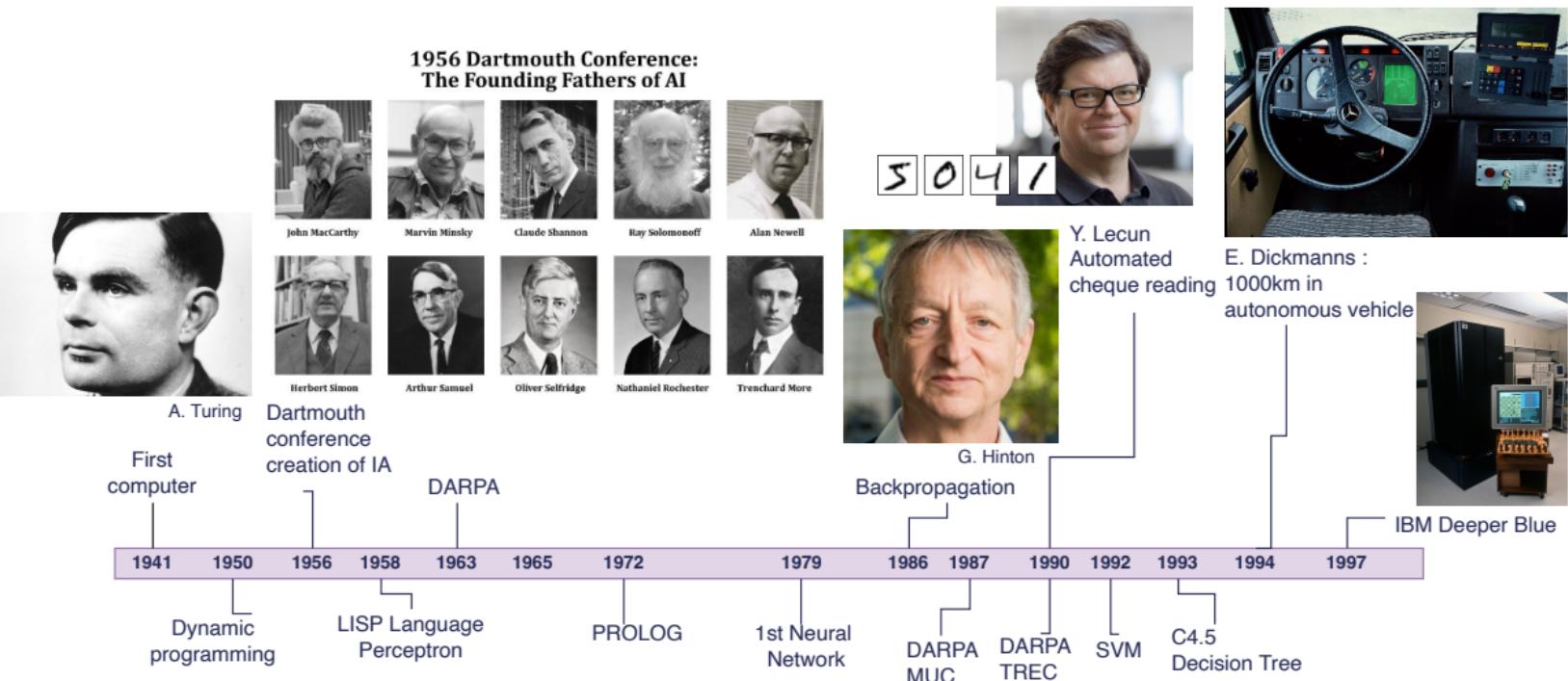


DE L'IA AU MACHINE-LEARNING[★]

[^{*} APPRENTISSAGE STATISTIQUE/APP. AUTOMATIQUE]

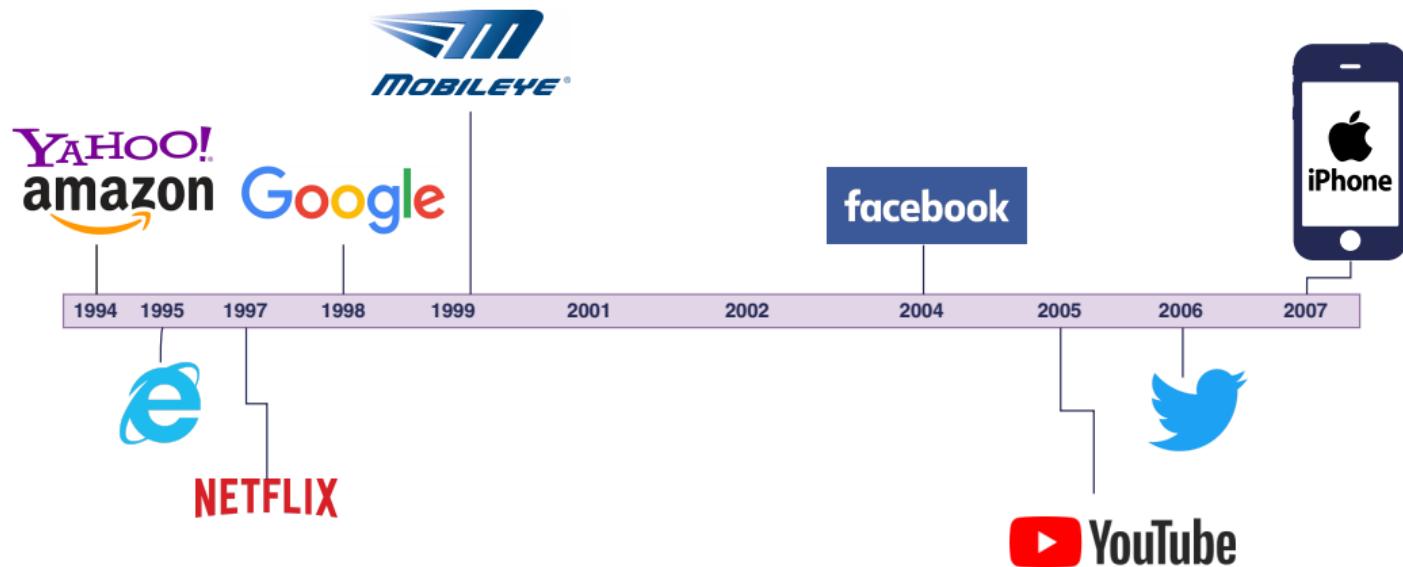
Historique rapide de l'IA

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



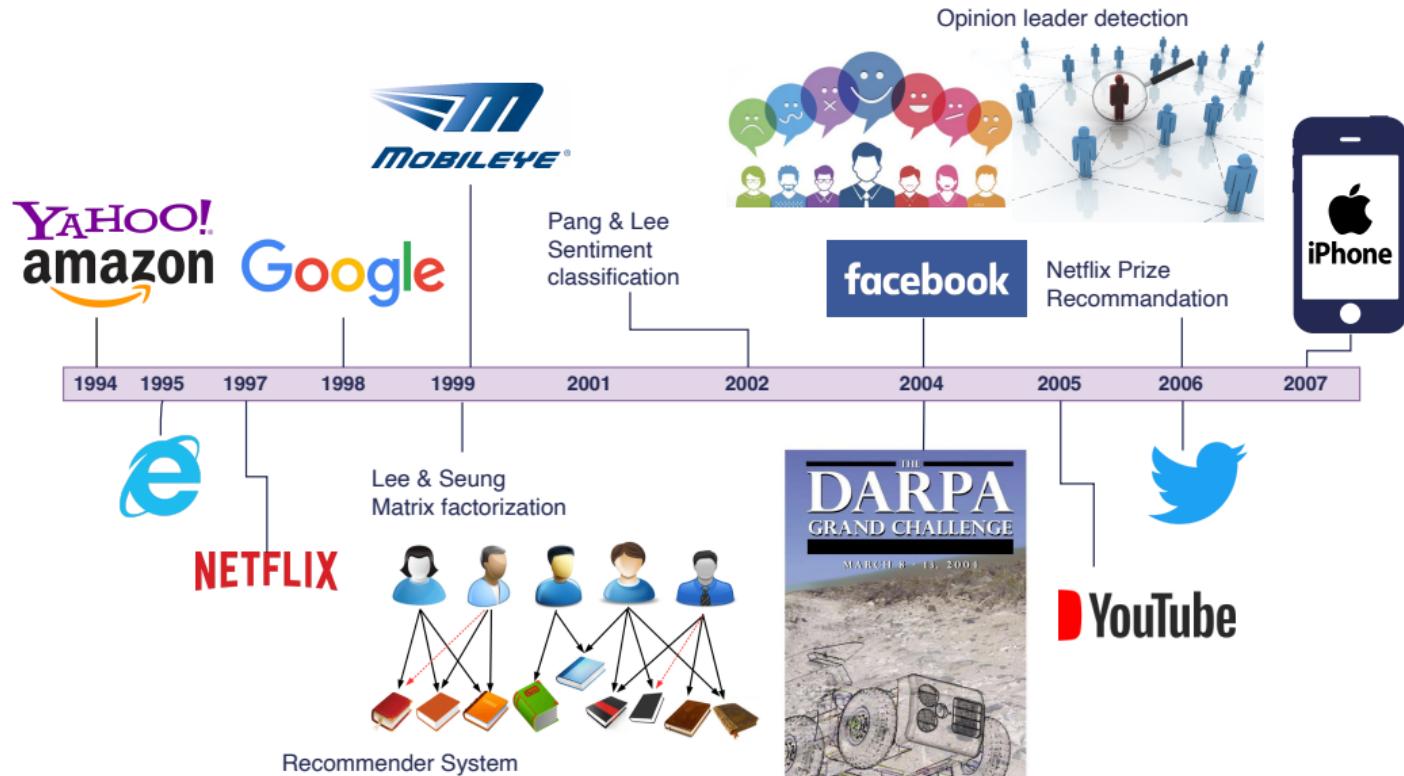
Historique rapide de l'IA

Emergence (ou refondation) des GAFAM/GAMMA



Historique rapide de l'IA

Emergence (ou refondation) des GAFAM/GAMMA

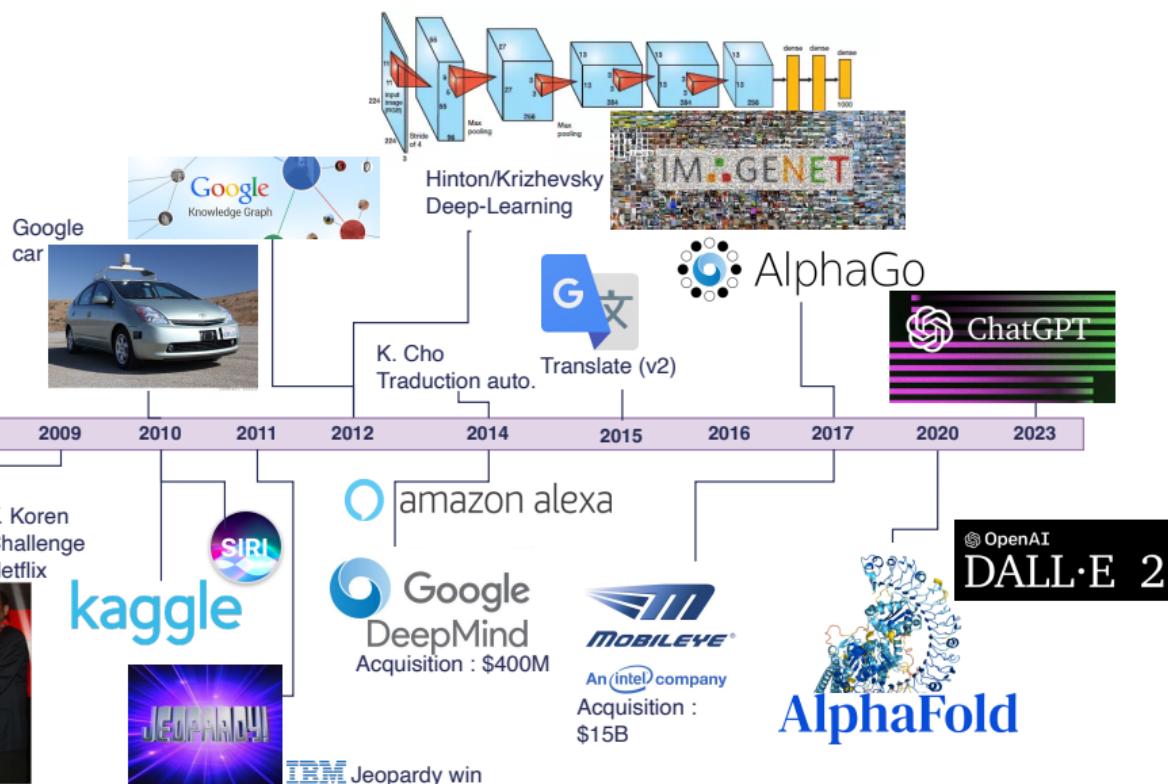


Historique rapide de l'IA

Formation d'une vague de l'Intelligence Artificielle



Sebastian Thrun:
DARPA Grand Challenge
victory



Ingrédients de l'Intelligence Artificielle



Données
Capteurs



Cadre
logiciel



Modèles



Stockage &
Calcul

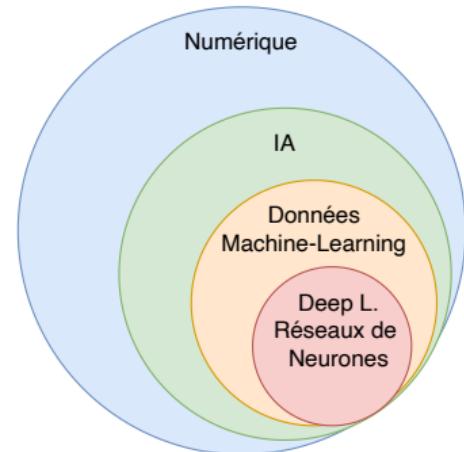


Google Cloud



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.

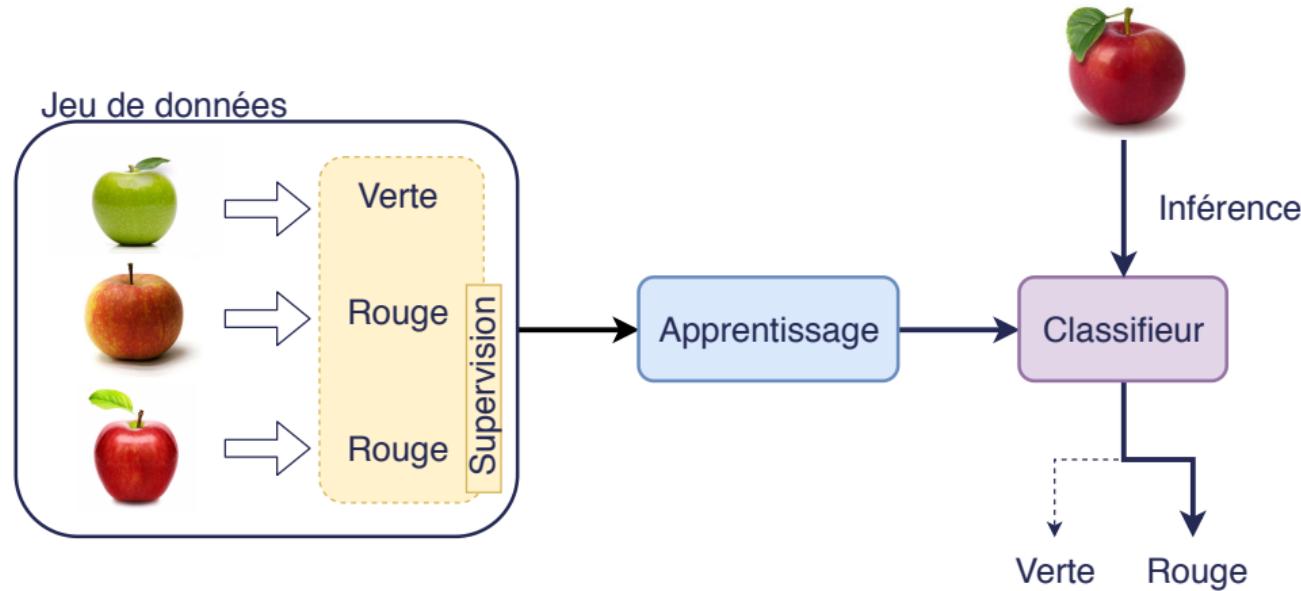
Marvin Lee Minsky, 1956

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

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

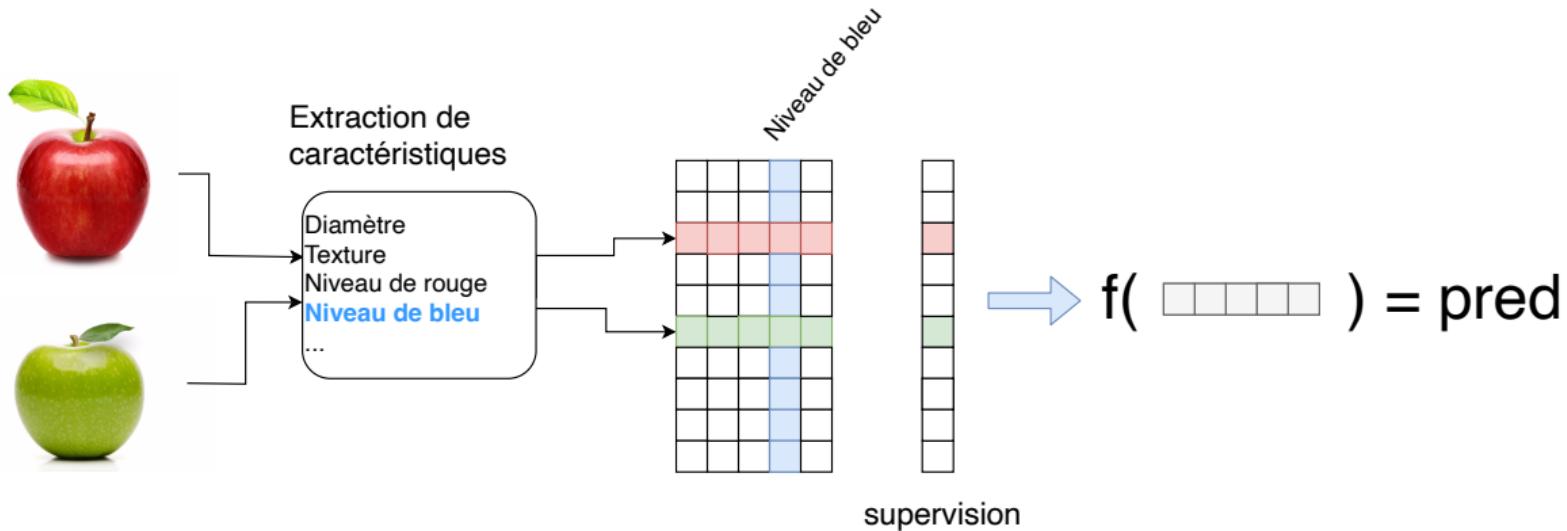
Andrew Ng, 2015

Chaine de traitements supervisée & modèles

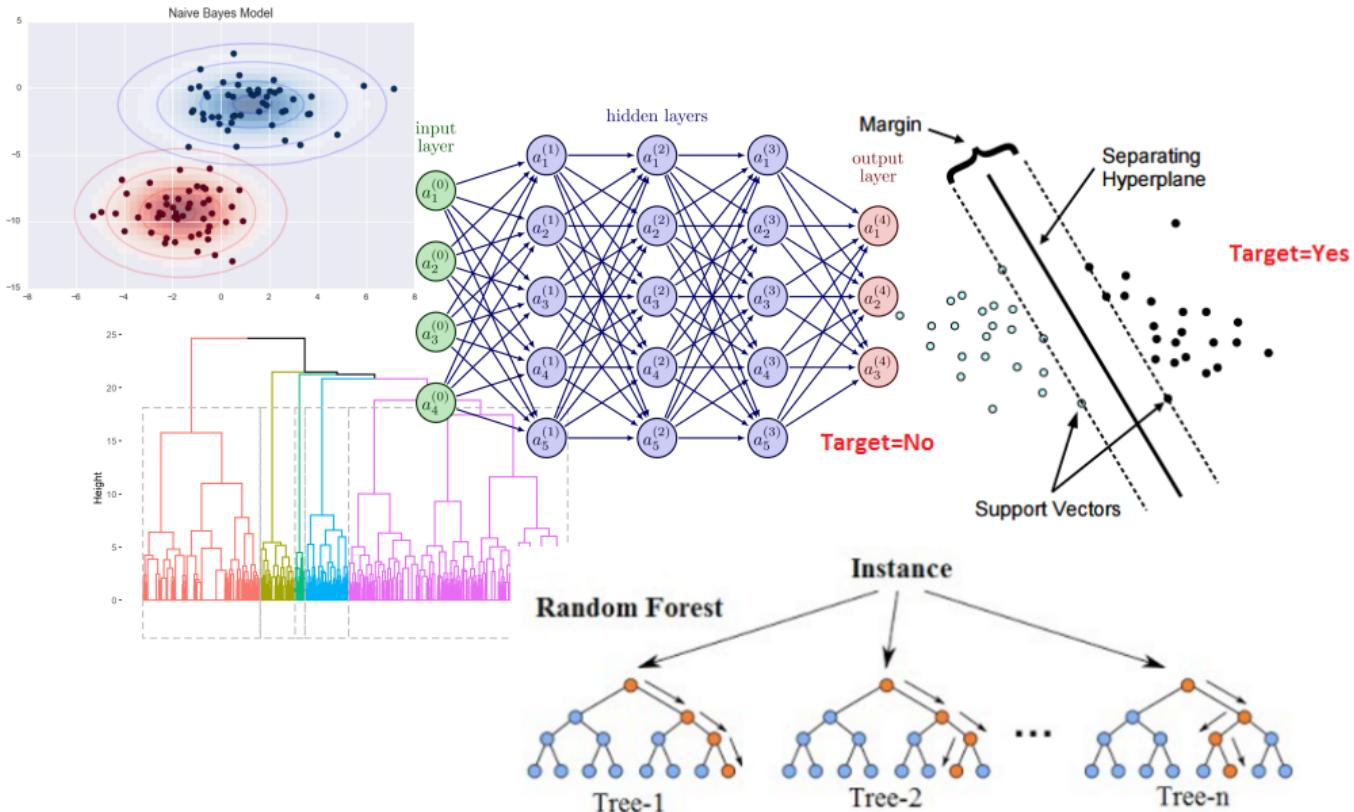


- ▶ Promesse = construire un modèle *uniquement* à partir des observations

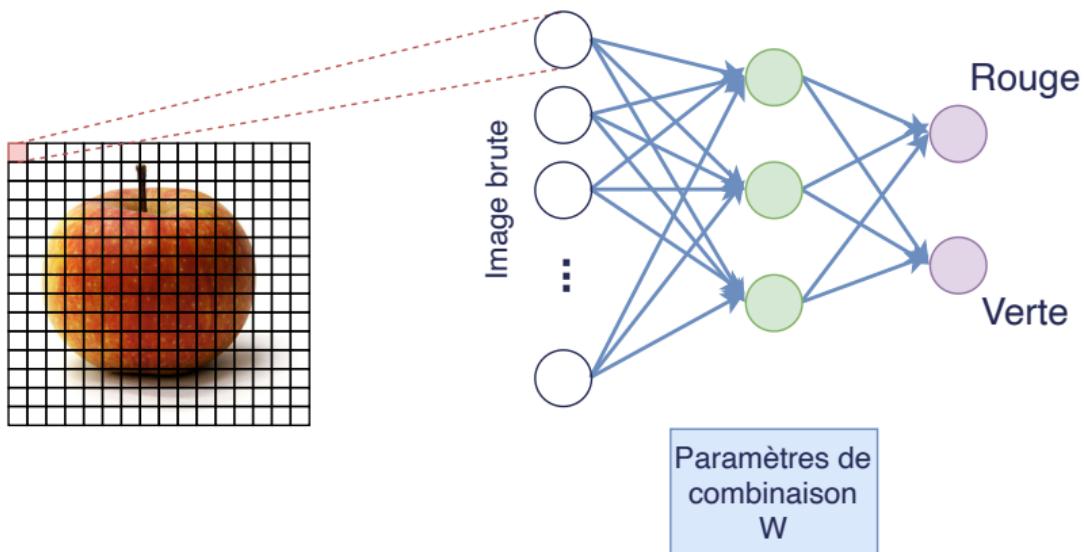
Chaine de traitements supervisée & modèles



Chaine de traitements supervisée & modèles



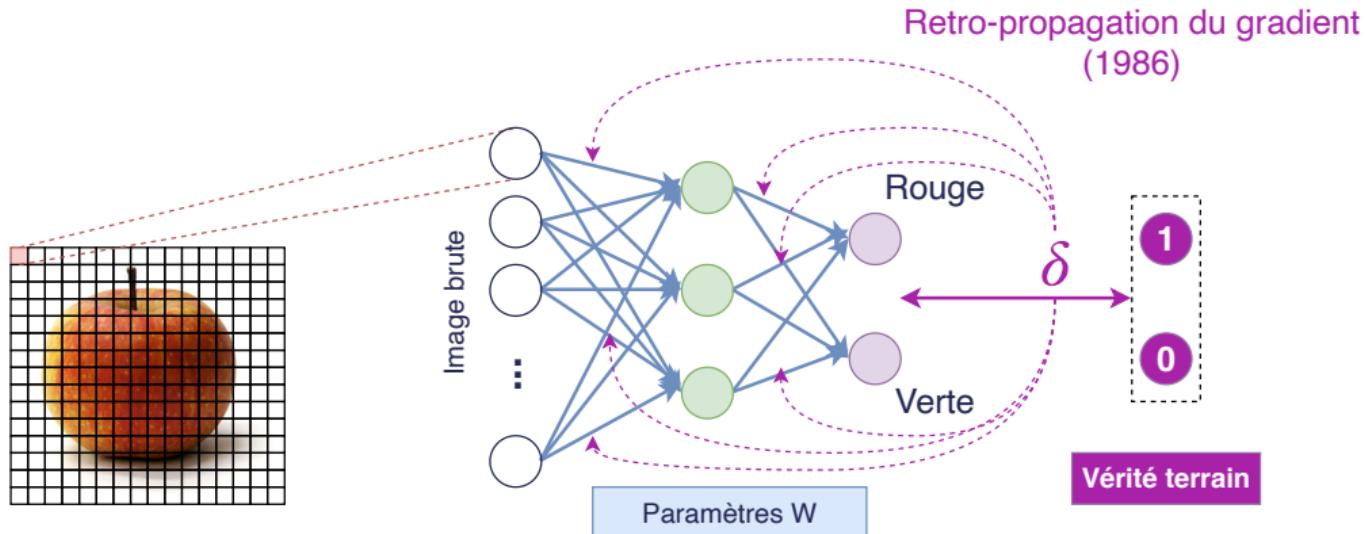
Chaine de traitements supervisée & modèles



► Initialisation aléatoire...

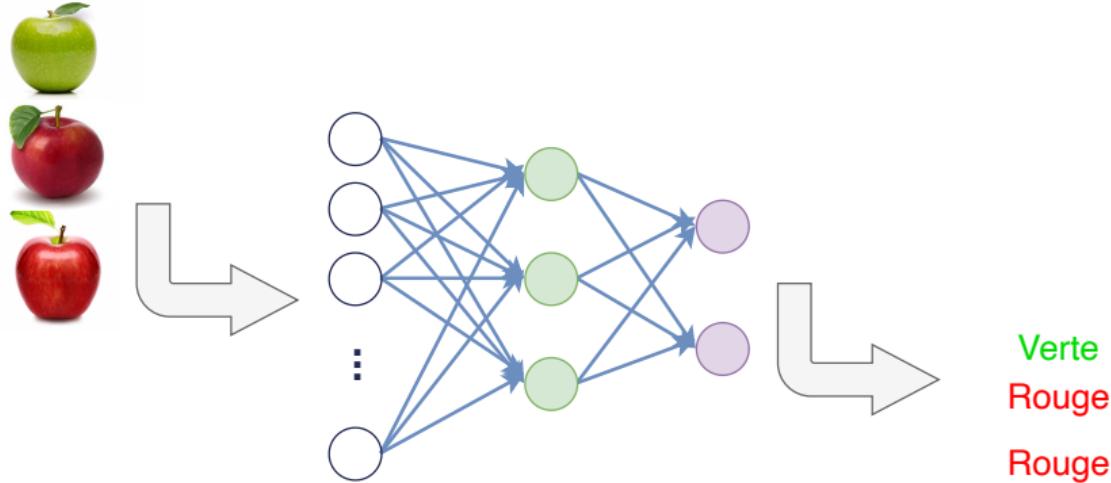
Et décision aléatoire (au début!)

Chaine de traitements supervisée & modèles



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

Chaine de traitements supervisée & modèles

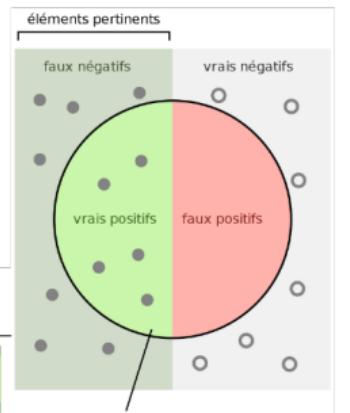
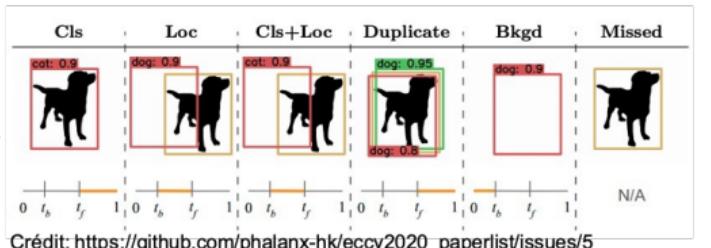
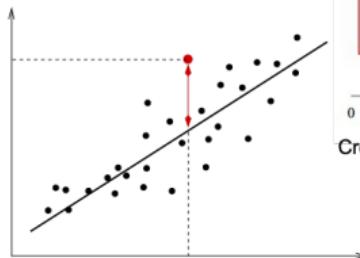


- ▶ Apprentissage lent et couteux
- ▶ Inférence (beaucoup plus) rapide

Mesurer les performances

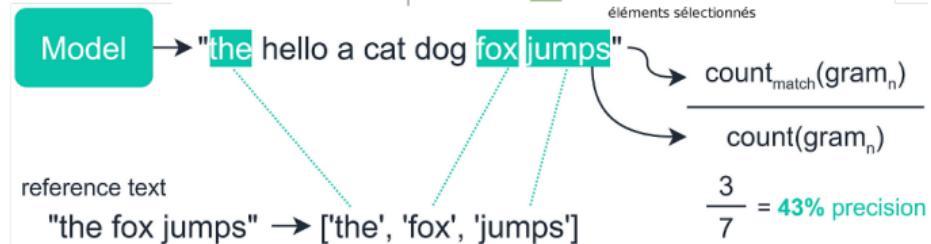
Estimer les performances (en généralisation)...

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



1	2	3	4	5	
<u>Relevance</u>					
Position	1	2	3	4	5

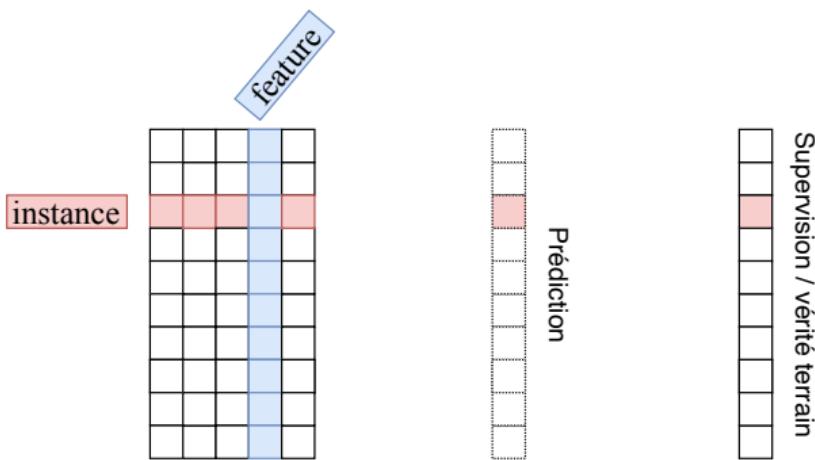
Recall@3 = $2/(2+1) = 2/3 = 0.67$



Mesurer les performances

Estimer les performances (en généralisation)...

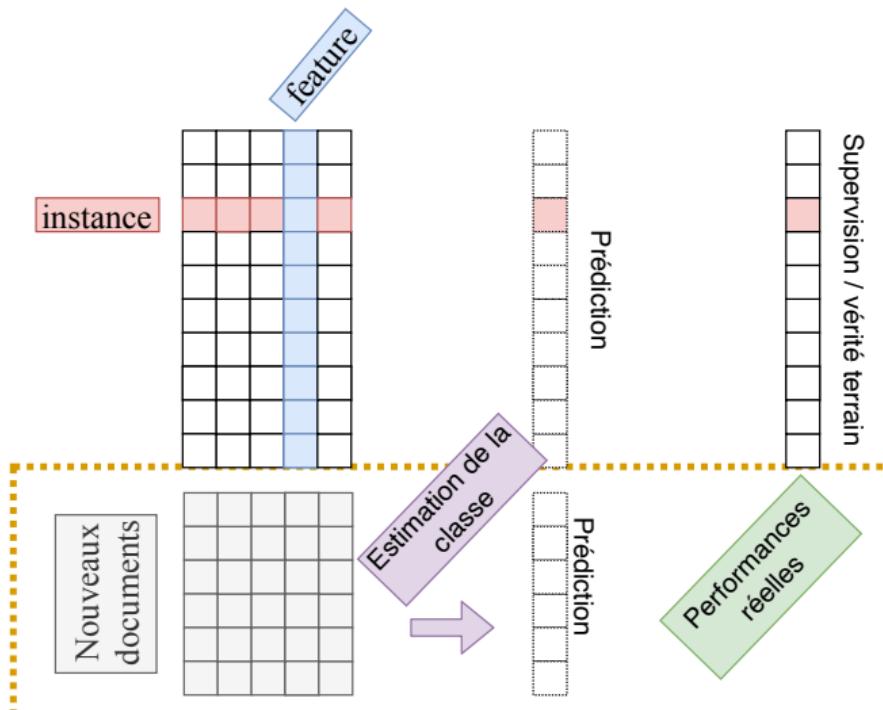
Est aussi important que l'apprentissage du modèle lui-même!



Mesurer les performances

Estimer les performances (en généralisation)...

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



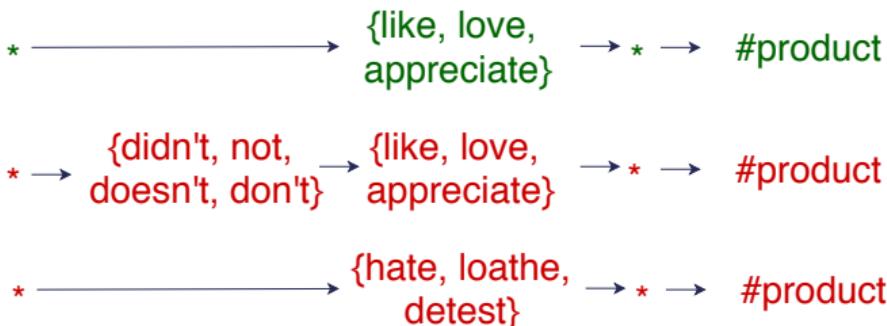
DEEP LEARNING & APPRENTISSAGE DE REPRÉSENTATION

L'IA et les données textuelles

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

Linguisitique [1960-2010]

Systèmes à base de règles:

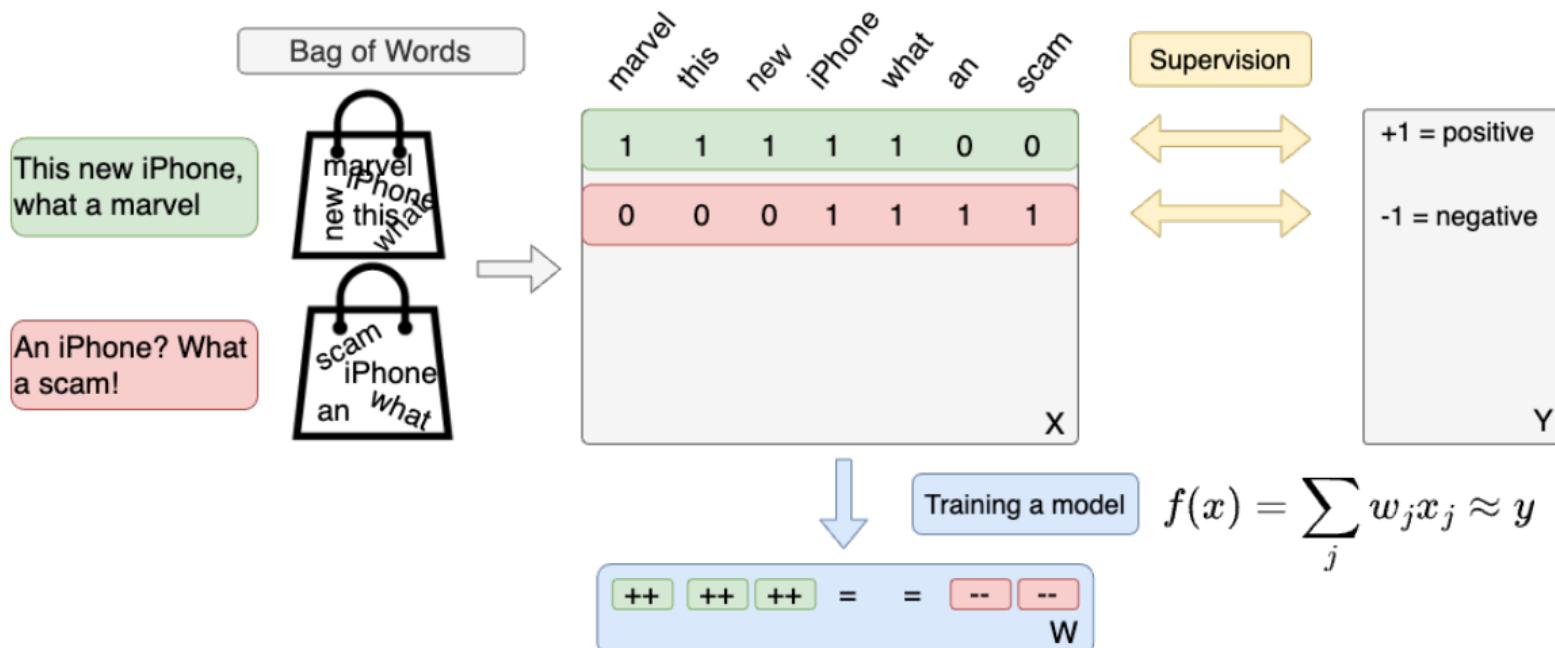


- Besoin de connaissance expert
- Extraction des règles ⇔ données très propres
- Très bonne précision
- Faible rappel
- Système interprétable

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 ⇔ données très propres
- + Système interprétable
- + Très bonne précision
- Faible rappel

Machine Learning [1990-2015]

- ▶ Peu de connaissance expert
- ▶ Extraction statistique ⇔ résiste aux données bruitées
- ≈ 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

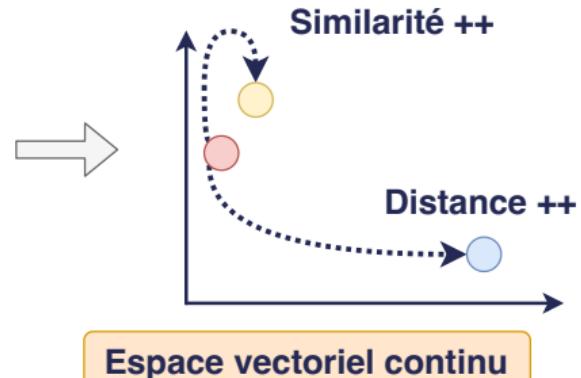
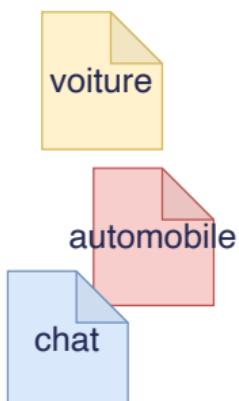
[2008, 2013, 2016]

Corpus en sac de mots

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

mot₁ ... voiture ... automobile ... chat ... mot_D

Mêmes distances



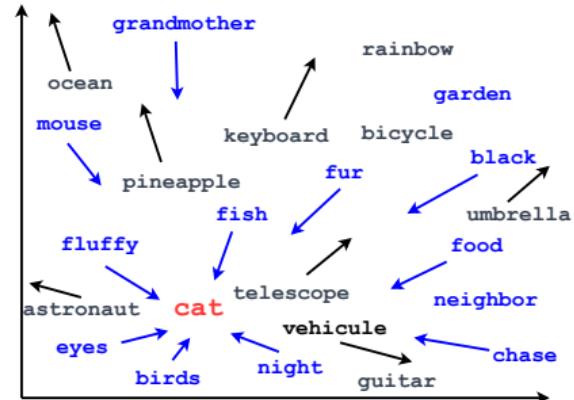
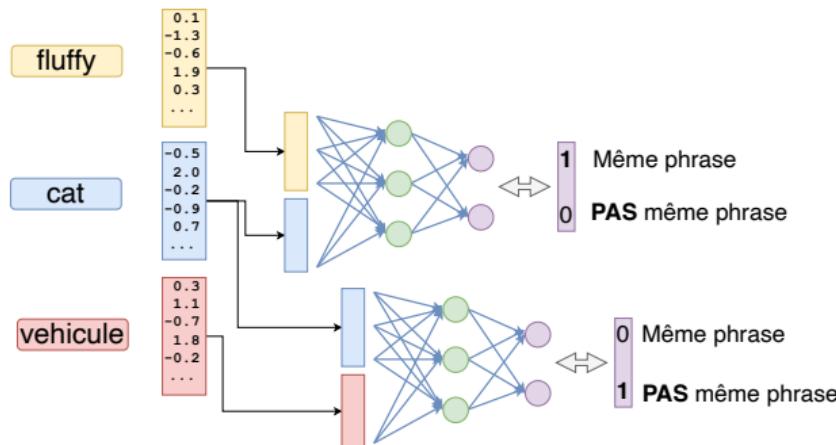
Deep learning et données textuelles: révolutionS 2010'

1. Des sacs de mots aux représentations vectorielles

[2008, 2013, 2016]

The fluffy cat napped lazily in the sunbeam.
I adopted a stray cat from the shelter last week.
My cat loves to chase after toy mice.
The black cat stealthily crept through the dark alley.
I often find my cat perched on the windowsill, watching birds.
She gently stroked her cat's fur as it purred contentedly.
Our neighbor's cat frequently visits our backyard.
The playful cat swatted at the dangling string with its paw.
My cat has a preference for fish flavored cat food.
The cat stealthily stalked a mouse in the garden.
My grandmother has a collection of porcelain cat figurines.

Corpus

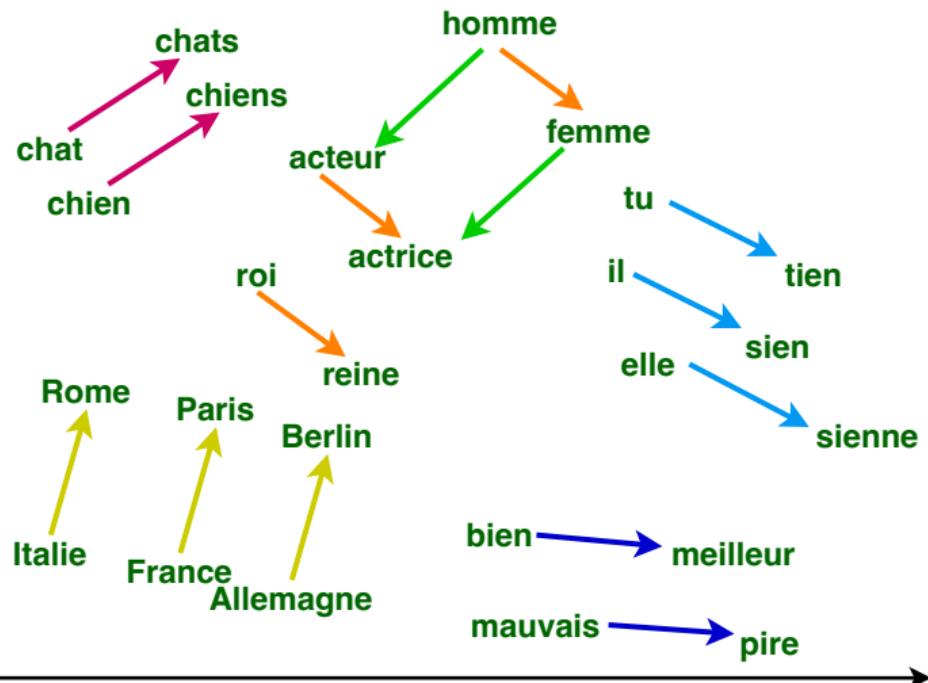




Deep learning et données textuelles: révolutionS 2010'

1. Des sacs de mots aux représentations vectorielles

[2008, 2013, 2016]



- ▶ Espace sémantique :
signification proche
 \Leftrightarrow
position proche
- ▶ Espace structuré :
régularités grammaticales,
connaissances élémentaires,
...

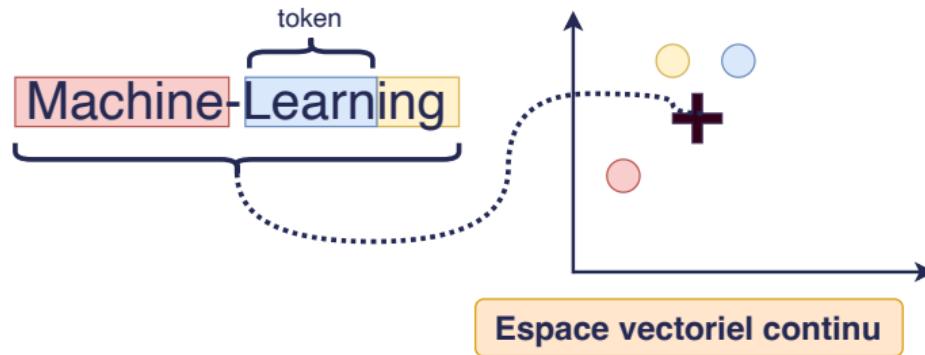
Deep learning et données textuelles: révolutionS 2010'

1. Des sacs de mots aux représentations vectorielles

[2008, 2013, 2016]

Des mots aux tokens

Décomposition en groupes
de lettres fréquents



- ▶ Représentation des mots inconnus
- ▶ Résistance aux fautes d'orthographe

Deep learning et données textuelles: révolutionS 2010'

2. Agrégation des représentations de mots

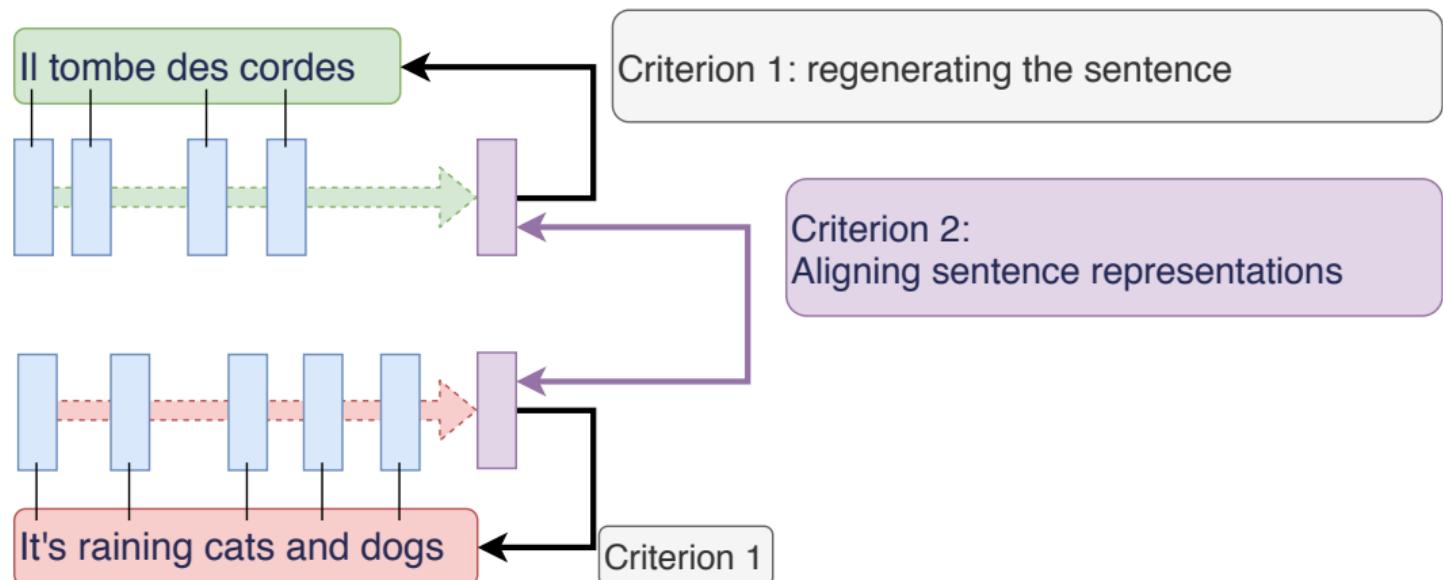
[2014, 2017]



Deep learning et données textuelles: révolutionS 2010'

2. Agrégation des représentations de mots

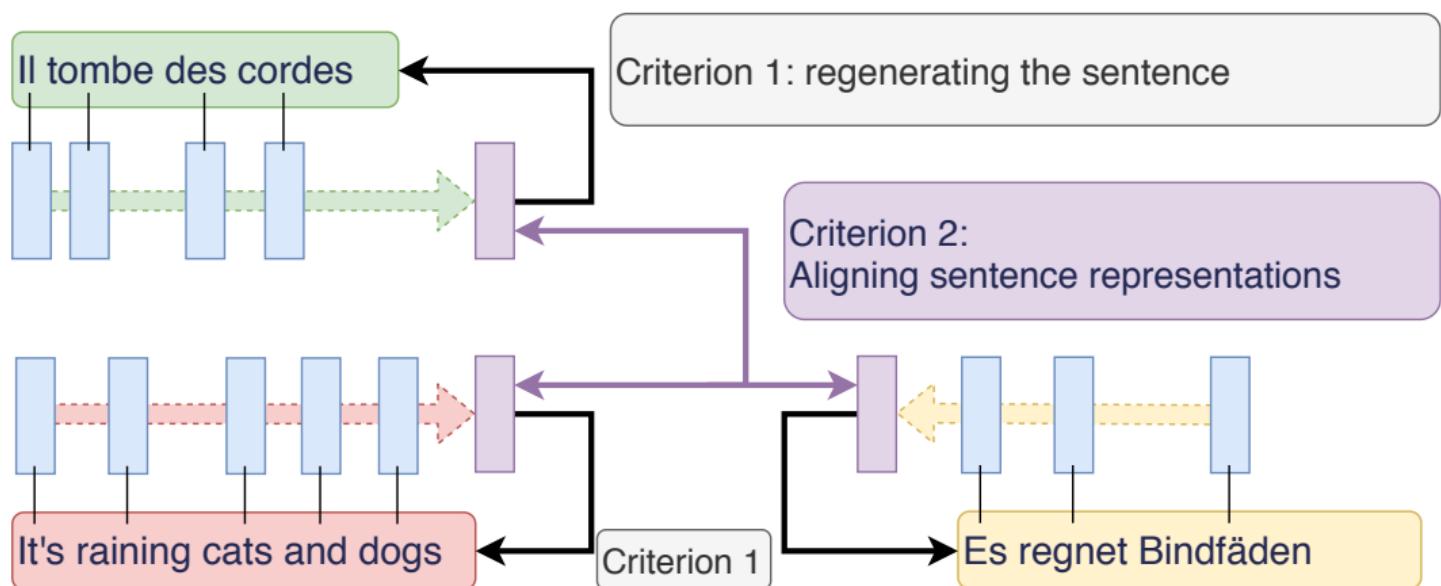
[2014, 2017]



Deep learning et données textuelles: révolutionS 2010'

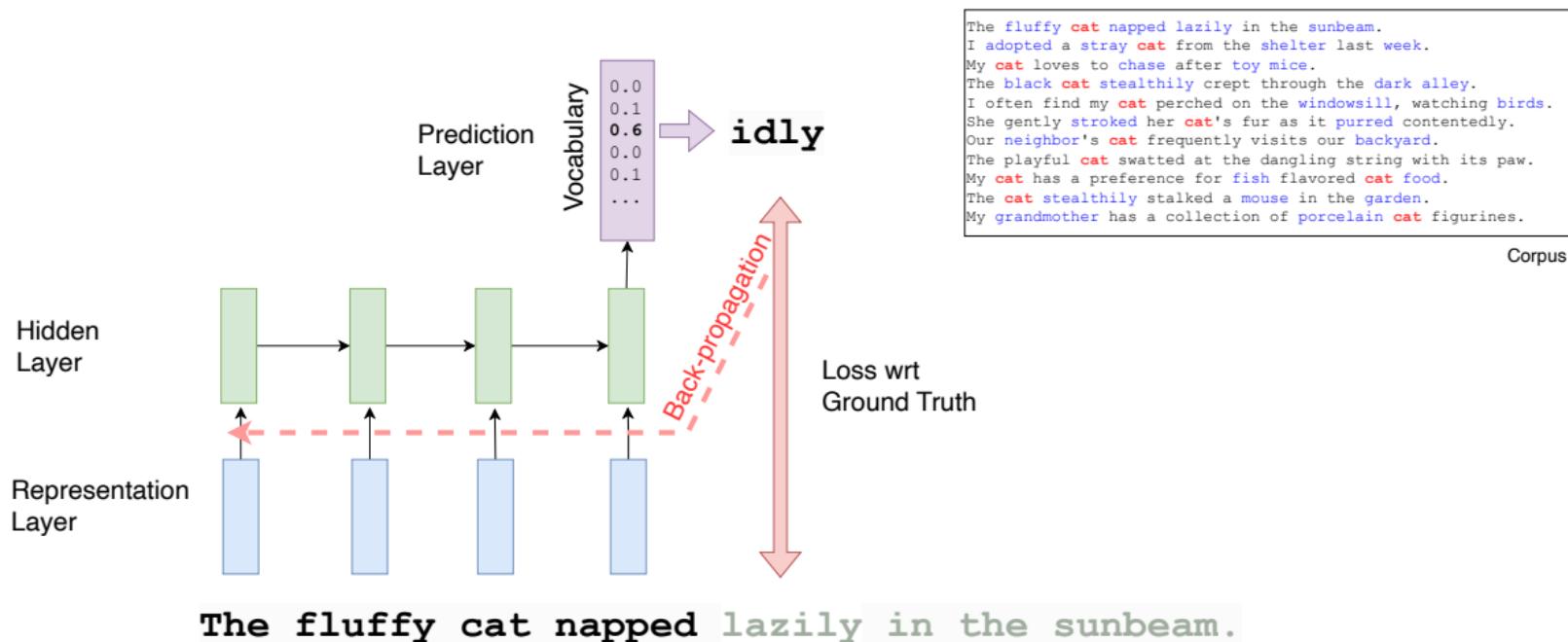
2. Agrégation des représentations de mots

[2014, 2017]



Deep learning et données textuelles: révolutionS 2010'

- Génération & Représentation
- Nouvelle manière d'apprendre les positions des mots



Sequence to sequence learning with neural networksy, Sutskever et al. NeurIPS 2014

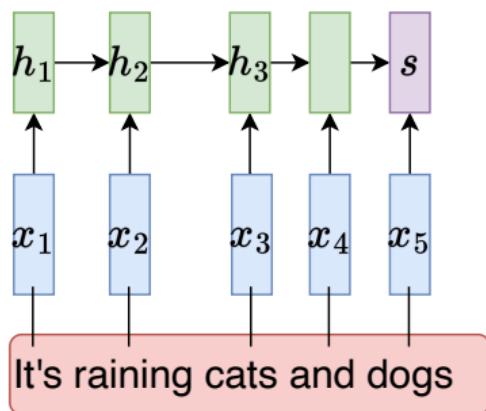
Deep learning et données textuelles: révolutionS 2010'

2. Agrégation des représentations de mots... V2

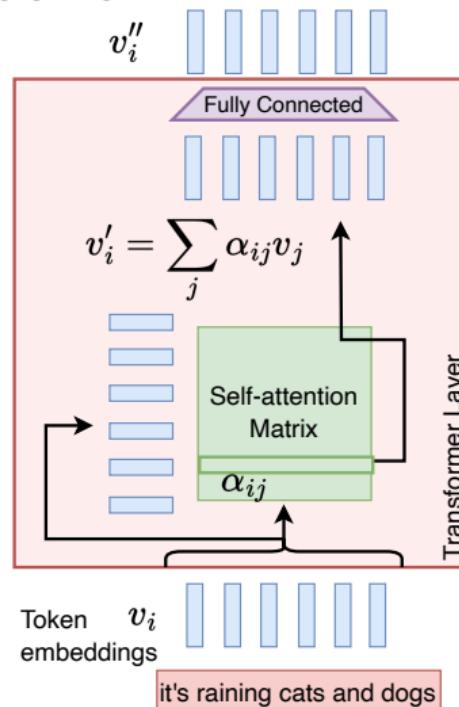
[2000, 2017-2023]

Recurrent Neural Network:

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



Transformer:

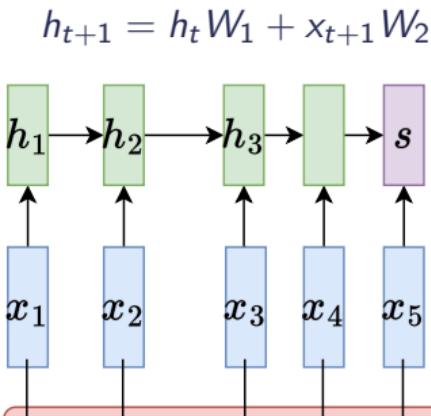


Deep learning et données textuelles: révolutionS 2010'

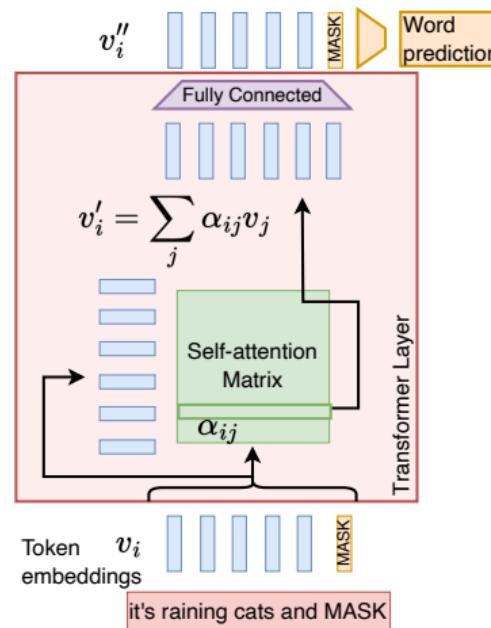
2. Agrégation des représentations de mots... V2

[2000, 2017-2023]

Recurrent Neural Network:



Transformer:

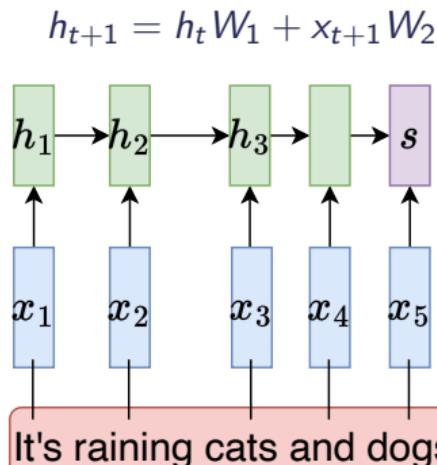


Deep learning et données textuelles: révolutionS 2010'

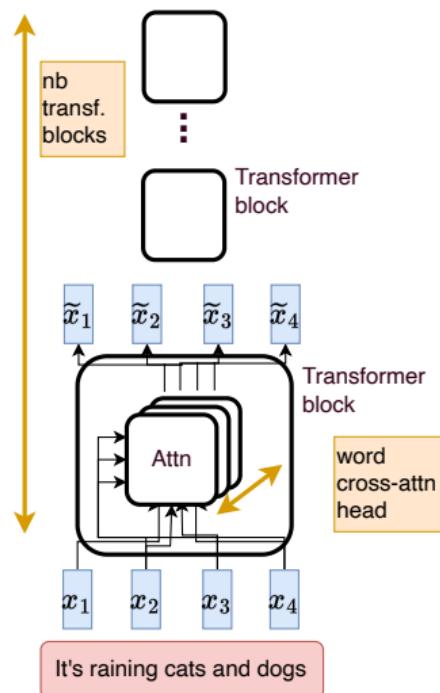
2. Agrégation des représentations de mots... V2

[2000, 2017-2023]

Recurrent Neural Network:



Transformer:

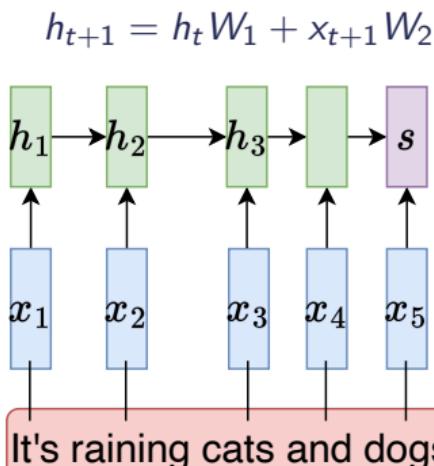


Deep learning et données textuelles: révolutionS 2010'

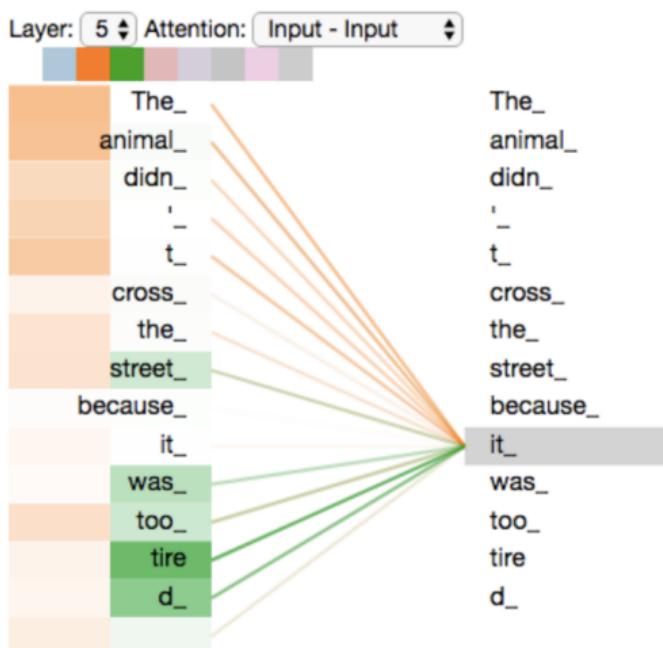
2. Agrégation des représentations de mots... V2

[2000, 2017-2023]

Recurrent Neural Network:



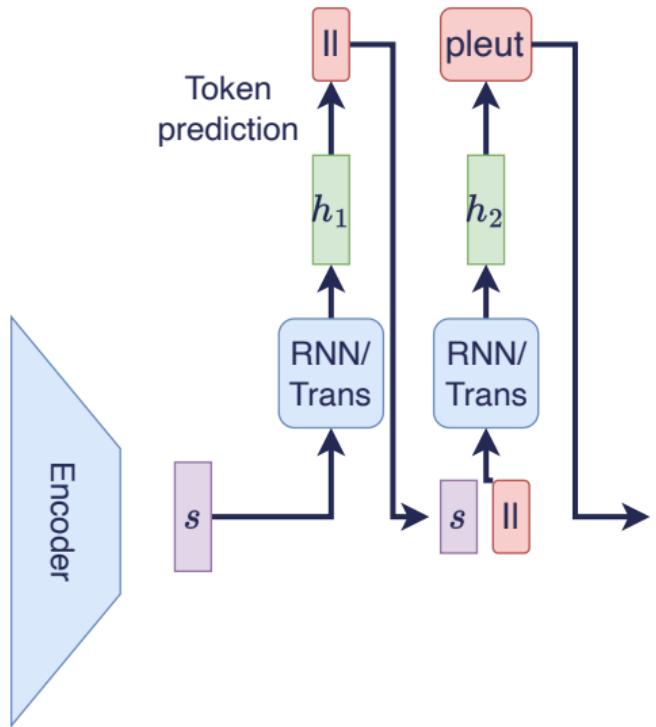
Transformer:



Deep learning et données textuelles: révolutionS 2010'

3. Des modèles génératifs

[2014, 2022]

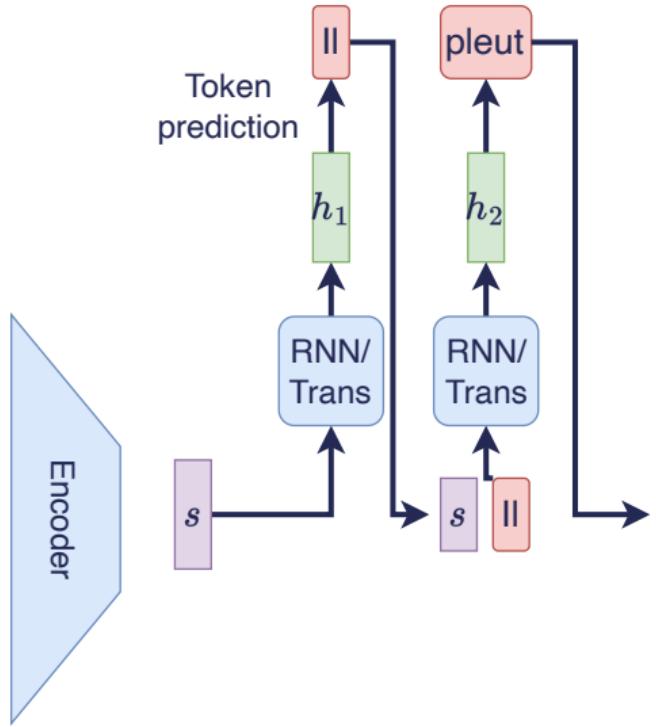


- ▶ Coût élevé (+beam search)
 - ▶ 1 appel / token
- ▶ Génération au sens du maximum de vraisemblance
- ▶ Principales tâches de NLP \Leftrightarrow reformulation en mode génératif

Deep learning et données textuelles: révolutionS 2010'

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

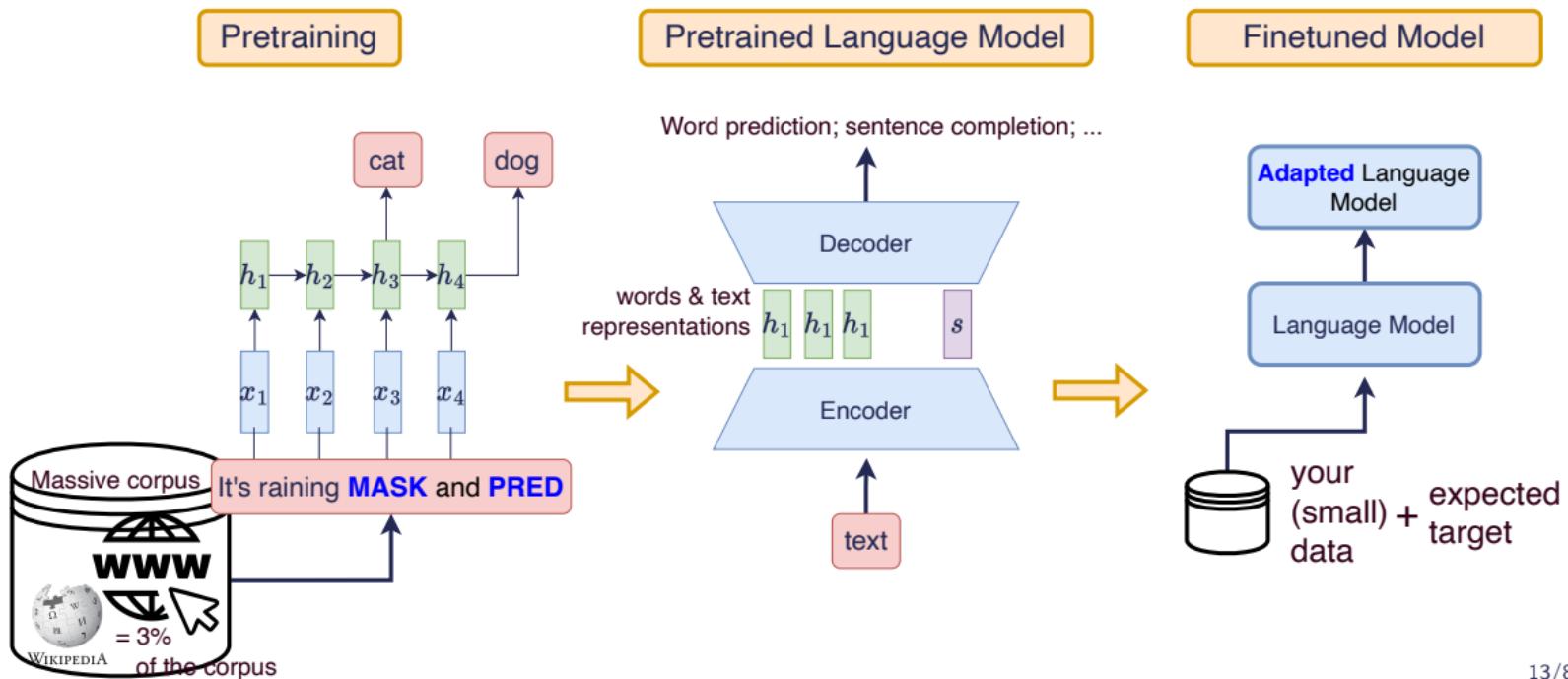


Deep learning et données textuelles: révolutionS 2010'

4. Transfert & fine-tuning

[2008, 2012, 2018]

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



CHATGPT

30 NOVEMBRE 2022

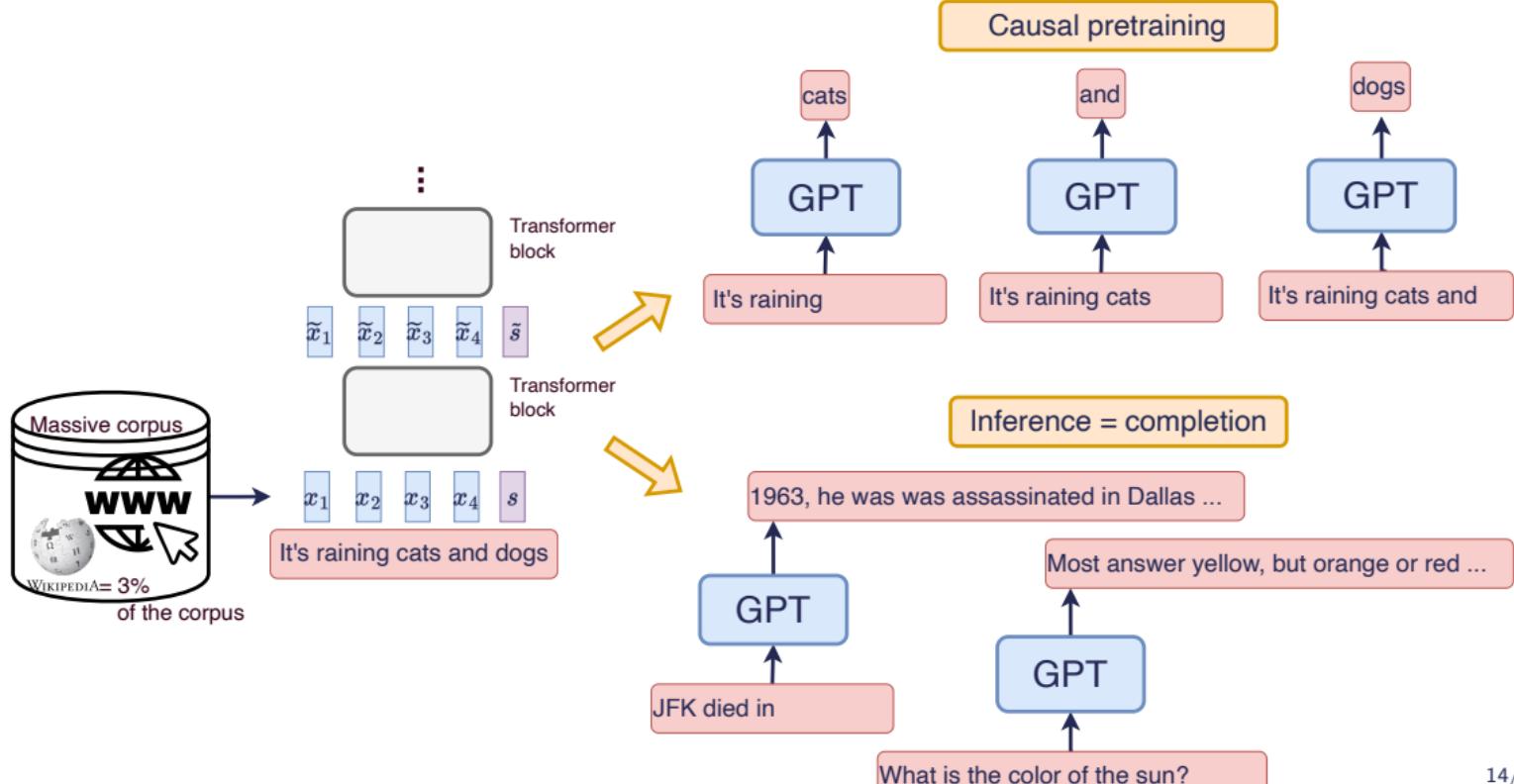
1 MILLION D'UTILISATEURS EN 5 JOURS

100 MILLIONS FIN JANVIER 2023

1.16 MILLIARDS EN MARS 2023

Les ingrédients de chatGPT

0. Transformer + données massives (GPT)



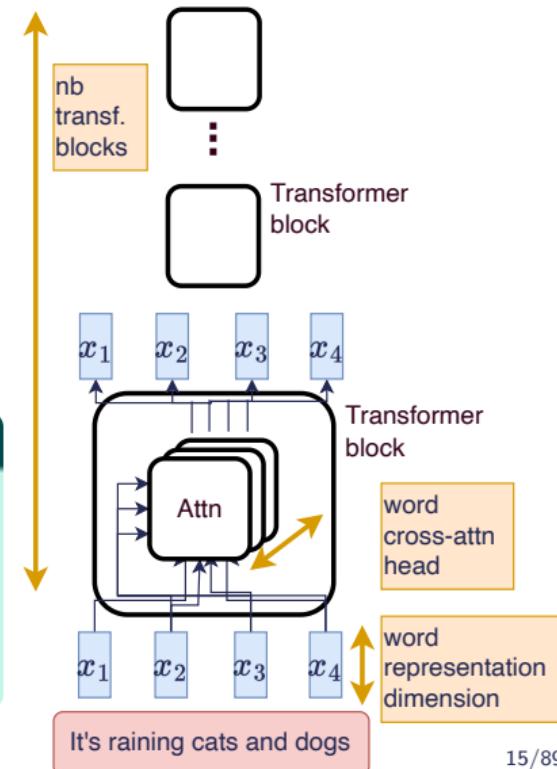
Les ingrédients de chatGPT

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?

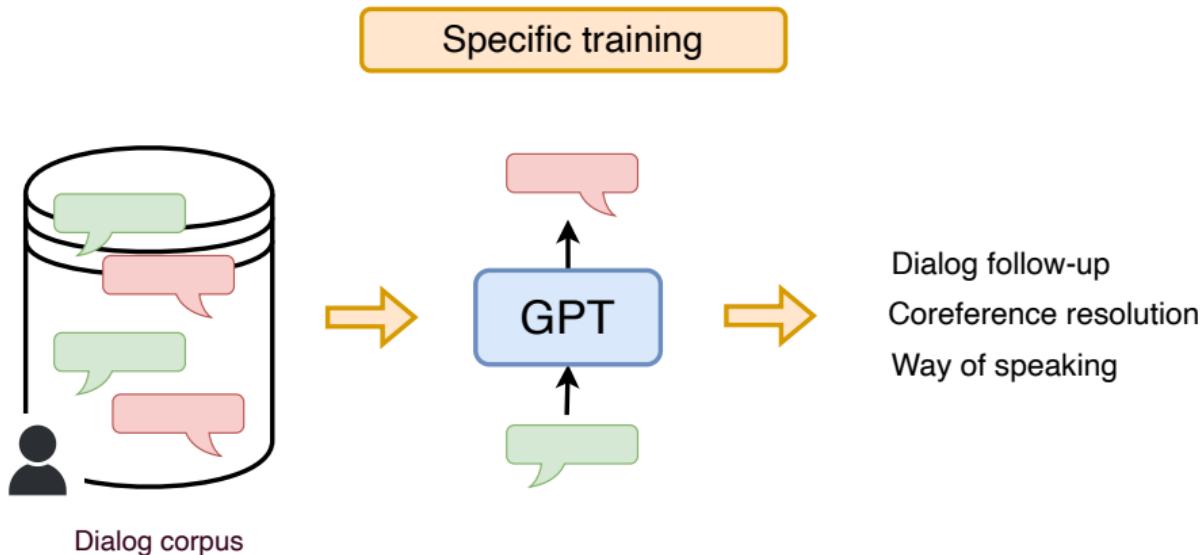
- ▶ $1.75 \cdot 10^{11} \Rightarrow 300 \text{ Go} + 100 \text{ Go}$ (stockage des données en inférence) $\approx 400 \text{ Go}$
- ▶ GPU NVidia A100 = 80Go de mémoire (=20k€)
- ▶ Coût pour (1) entraînement: 4.6 Millions d'€



It's raining cats and dogs

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

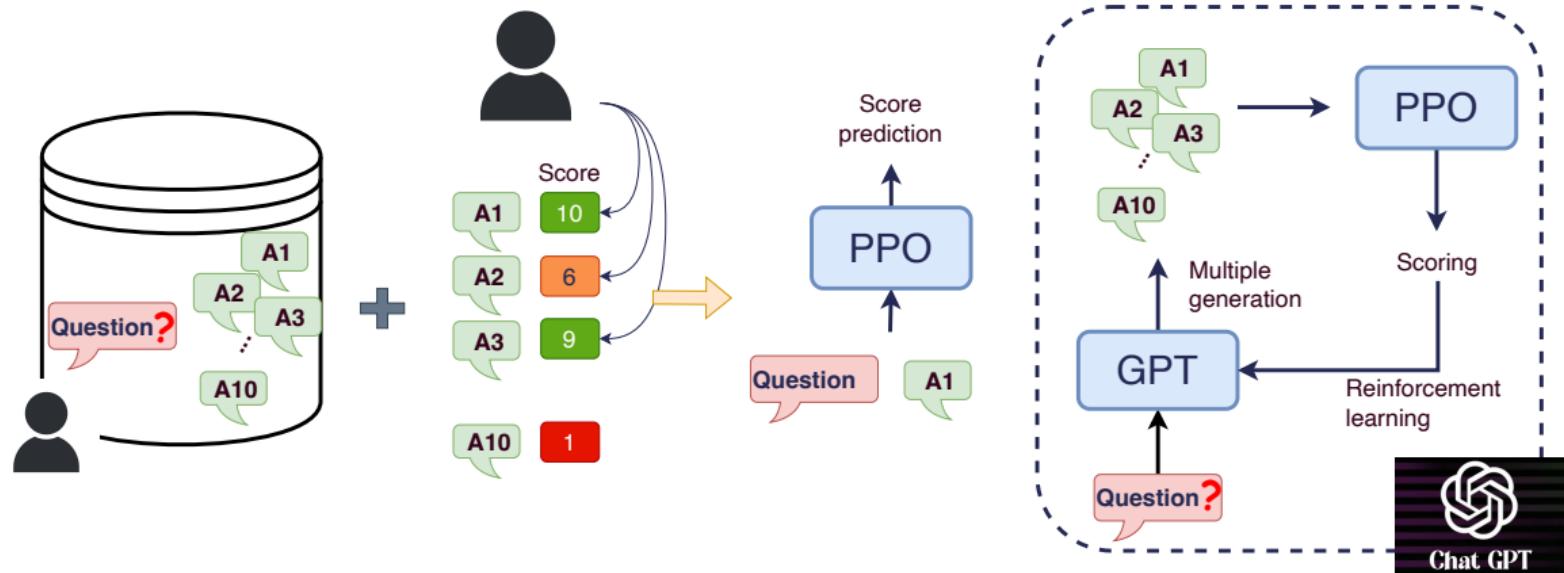
-320.4F

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

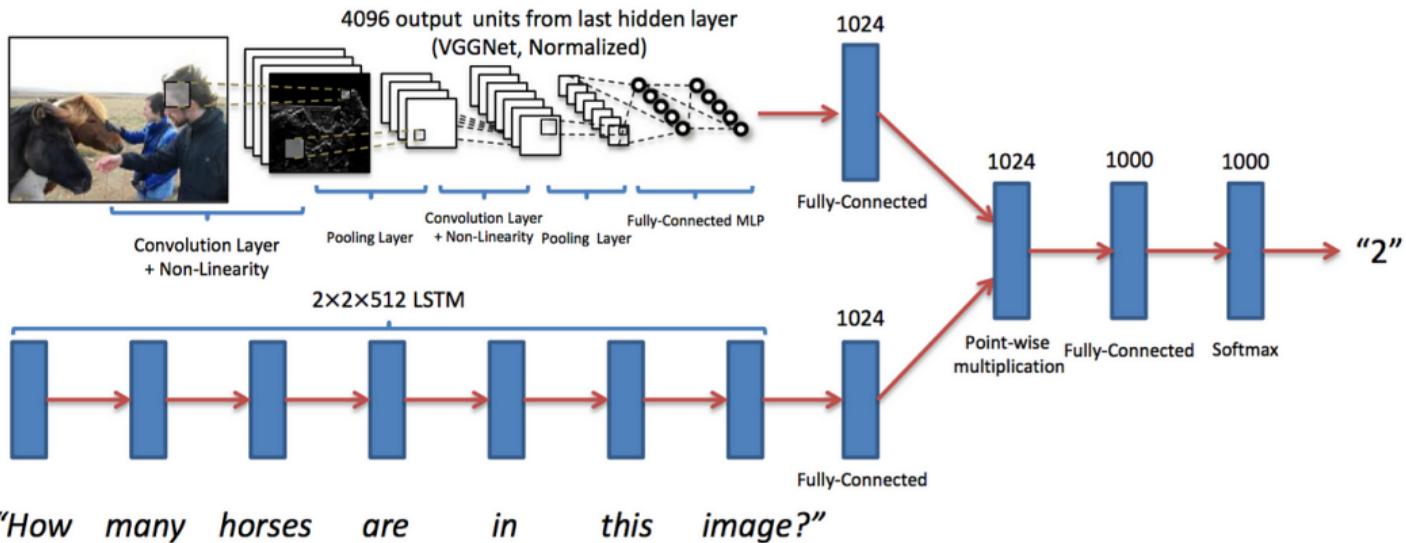


- BD faite par des humains
- Amélioration des réponses
- ... Aussi une manière d'éviter les sujets critiques

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



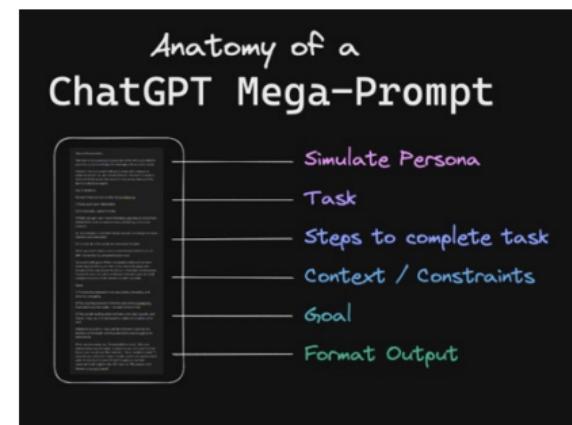
"How many horses are in this image?"

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

→ + de détails

Usage de chatGPT & Prompting

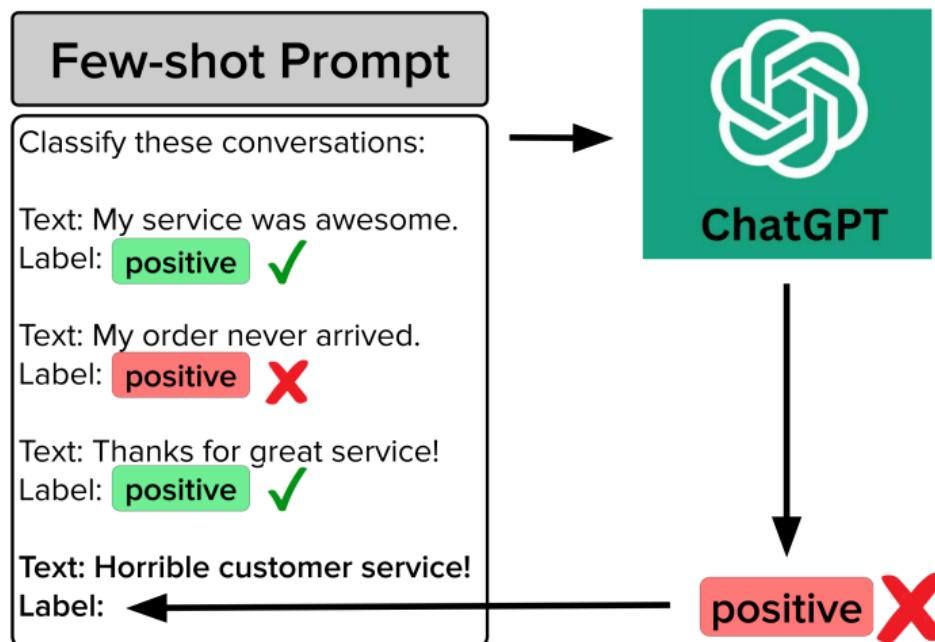
- ▶ Interroger chatGPT... Ca s'apprend! = *prompting*
 - ▶ Bien poser une question: ... *en détails*, ... *step by step*
 - ▶ Spécifier nb elts e.g. : *3 qualités pour ...*, *5 éléments pour...*
 - ▶ Poser un contexte : *cellule* pour un biologiste / assistant juridique
 - ▶ Ne pas s'arrêter à la première question
 - ▶ Détailler des points particuliers
 - ▶ Réorienter la recherche
 - ▶ Dialoguer
 - ▶ Reformulation
 - ▶ Explain like I'm 5, plus formel, à la manière d'un article scientifique, bro style, ...
 - ▶ Résumer, étendre
 - ▶ Ajouter des fautes (!)
- ⇒ Besoin de **pratique** [1 à 2 heures], discuter entre collègues



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

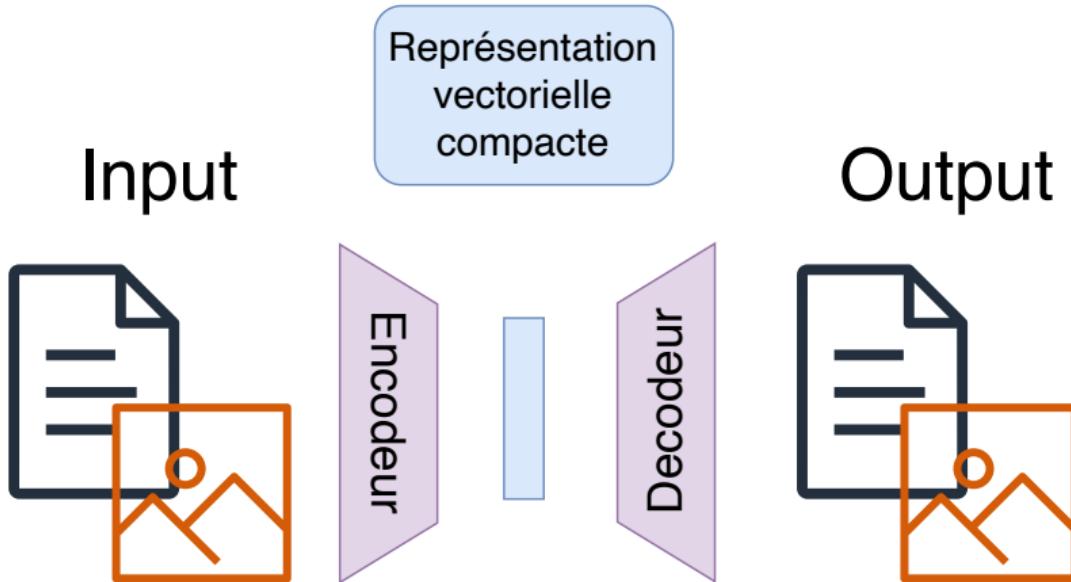
Vers du *few-shot learning*

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



IAs GÉNÉRATIVES

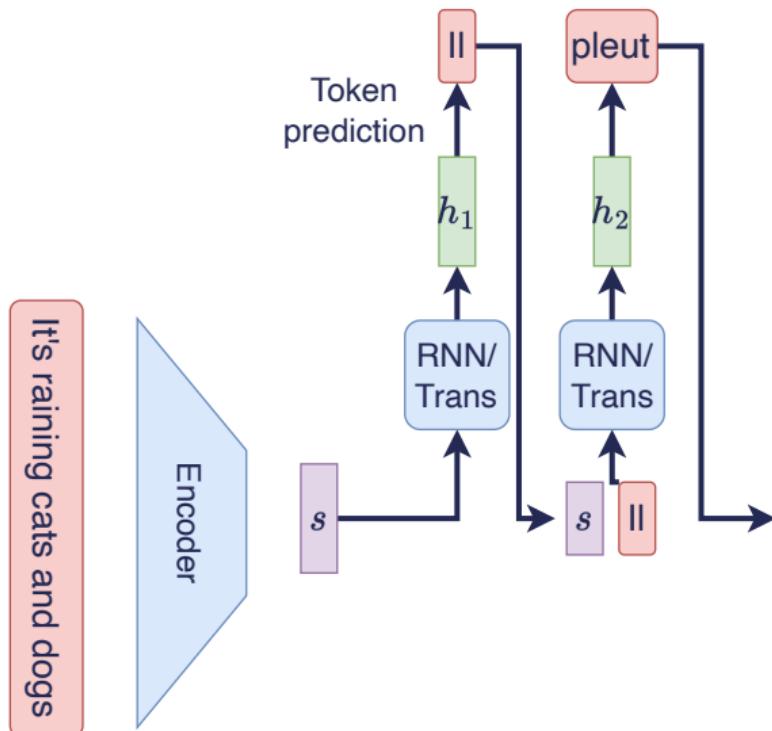
Différentes formes d'IA génératives



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

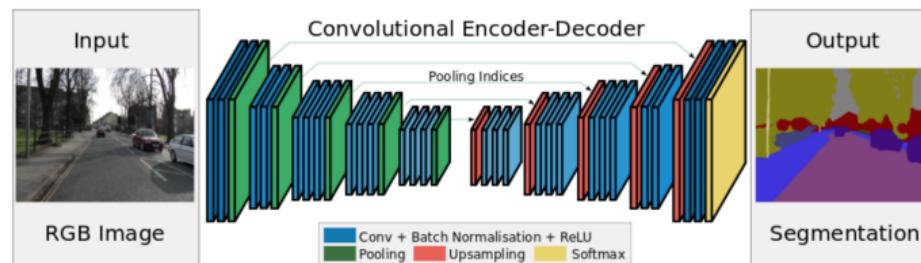
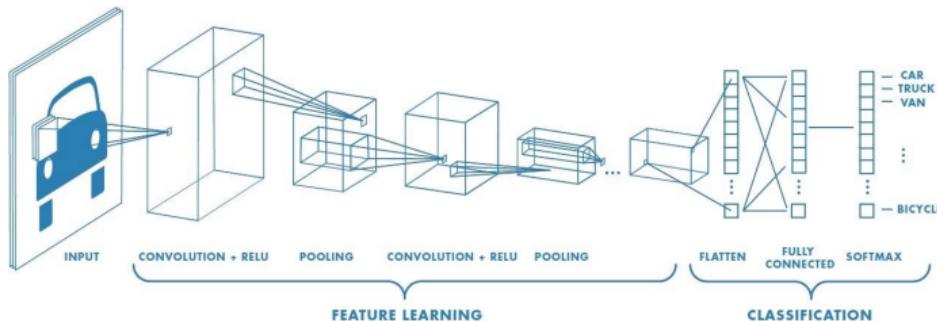
Différents media / différentes architectures

- ▶ Textes : problème classification



Différents media / différentes architectures

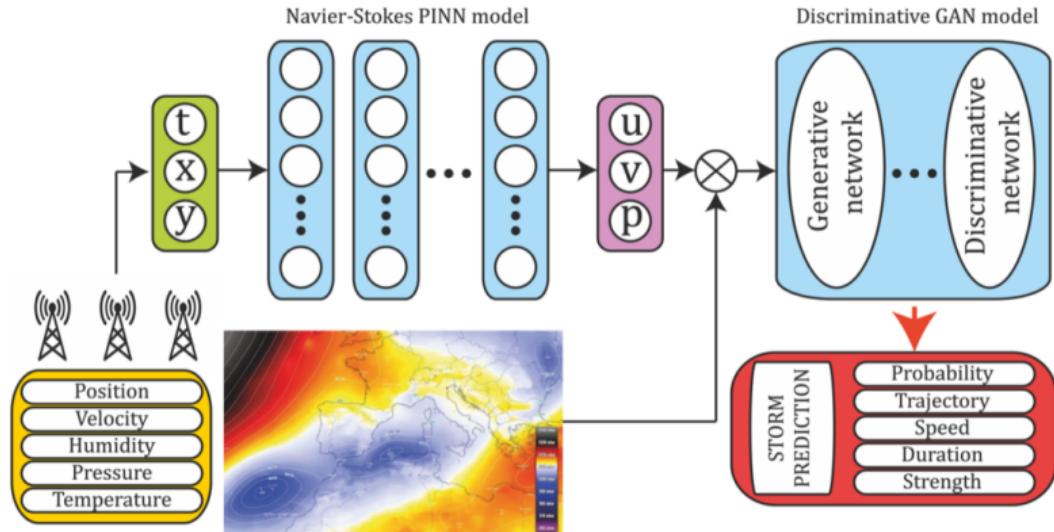
- ▶ Textes : problème classification
- ▶ Images : problème de régression multi-variées



Différents media / différentes architectures

- ▶ Textes : problème classification
 - ▶ Images : problème de régression multi-variées
 - ▶ Processus physique
-
- ▶ Mixer les approches mécanistes et *data-driven*

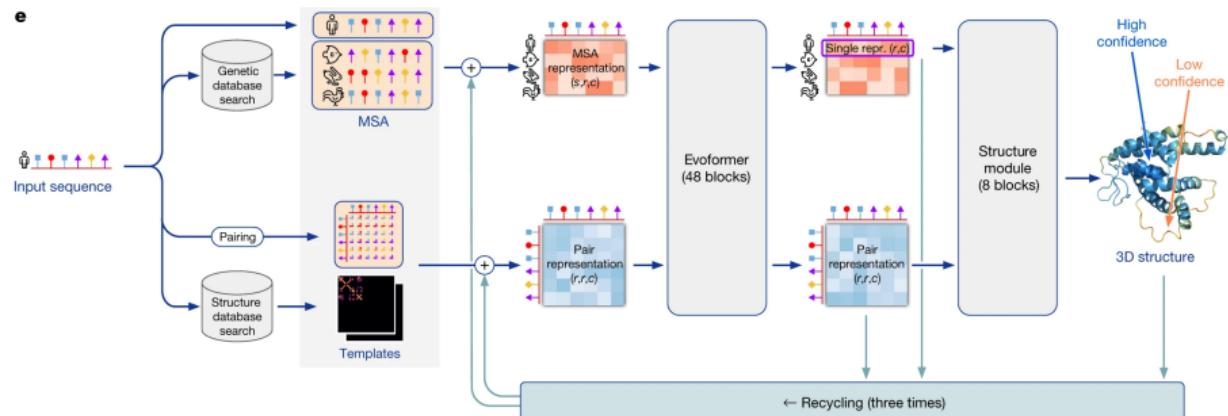
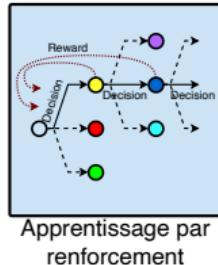
e.g. Modéliser des équations diff. dans un réseau de neurones



Différents media / différentes architectures

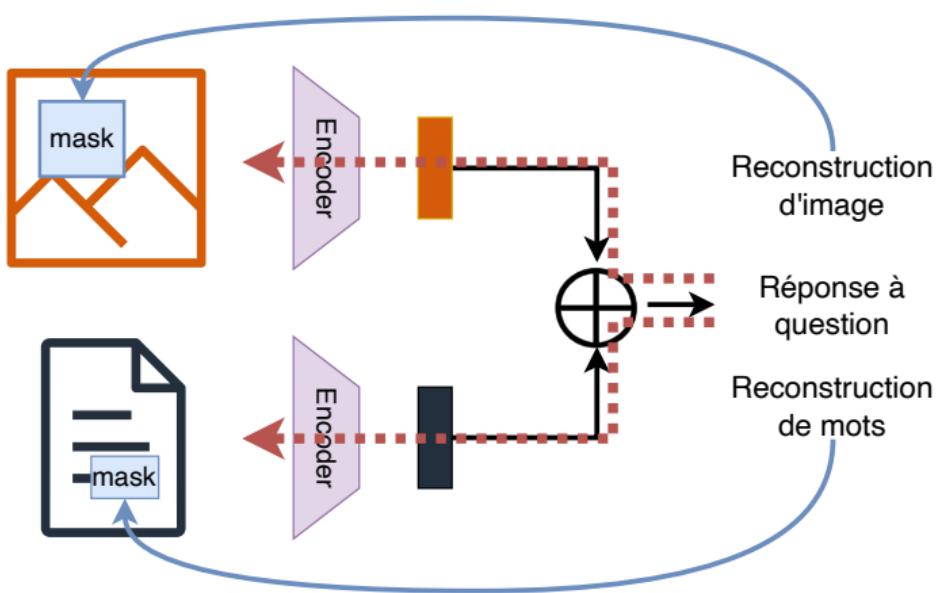
- ▶ Textes : problème classification
- ▶ Images : problème de régression multi-variées
- ▶ Processus physique
- ▶ Structures complexes / 3D / graphes : problème séquentiel

- ▶ Apprentissage par renforcement: action/récompense



Multi-modalité

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

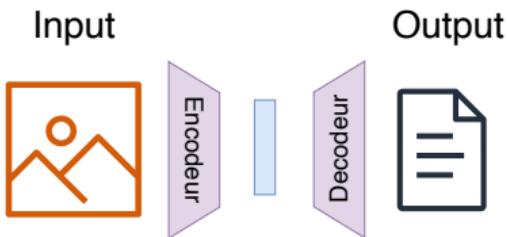


Alignement des espaces de représentations

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, ...*

Input



Encodeur



Decodeur



Output

TEXT DESCRIPTION

An astronaut Teddy bears A bowl
of soup

riding a horse lounging in a tropical
resort in space playing basketball
with cats in space

in a photorealistic style in the style
of Andy Warhol as a pencil drawing

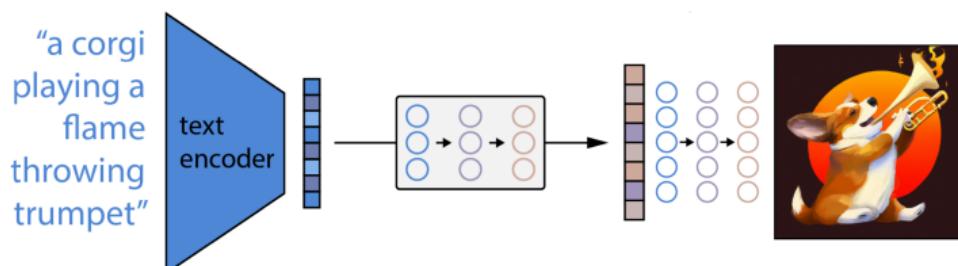
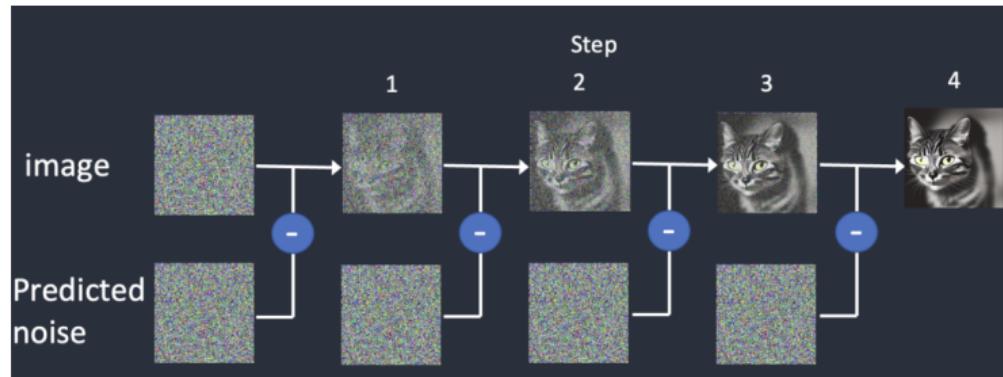


DALL·E 2



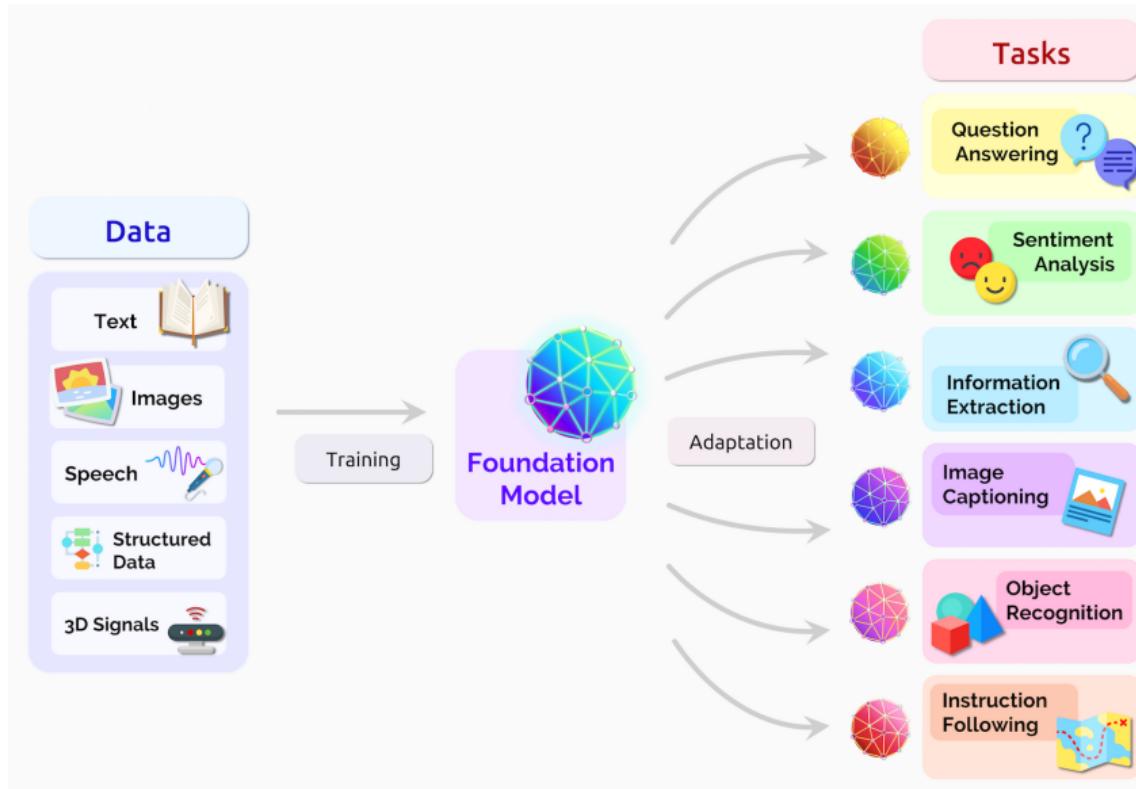
Multi-modalité

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



Vers des modèles de fondation plus larges?

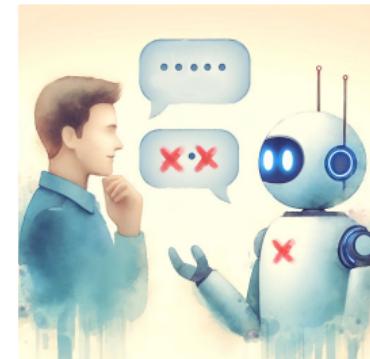
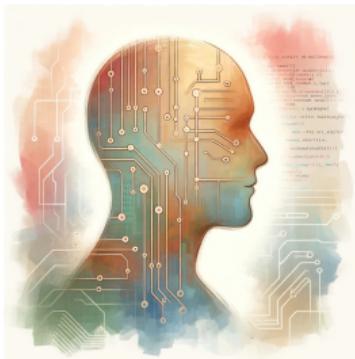
- Que les modalités s'enrichissent les unes avec les autres



CONCLUSION

Pourquoi tant de controverses?

- ▶ Nouvel outil [Décembre 2022]
- ▶ + Rapidité d'adoption jamais vue [1M user en 5 jours]
- ▶ Des forces et des faiblesses... Mal cernées par les utilisateurs
 - ▶ Des usages étonnantes / parfois absurdes
- ▶ Des retours d'expérience mal interprétés
 - ▶ Anthropomorphisation de l'algorithme et de ses erreurs
- ▶ Coût prohibitif : quel modèle économique? écologique? sociétal?



LIMITES

chatGPT et le rapport à la vérité

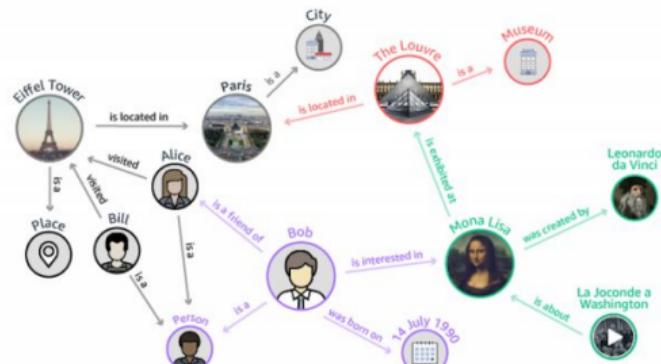
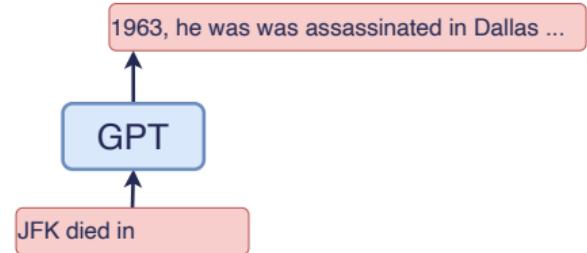
- 1 Vraisemblance = grammaire, accords,
concordance des temps,
enchaînements logiques...
⇒ Connaissances répétées
≈ grammaire

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

- 3 Fonctionnement **hors-ligne**

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

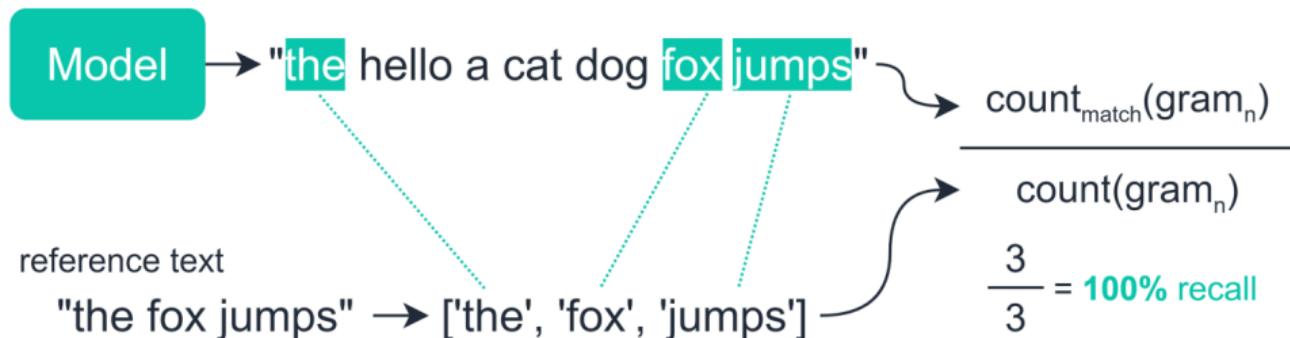
- 5 Des réponses brillantes...
Et des erreurs bêtes!
+ on ne sait pas prédire les erreurs



IA Génératives: comment évaluer les performances?

Le point critique aujourd'hui

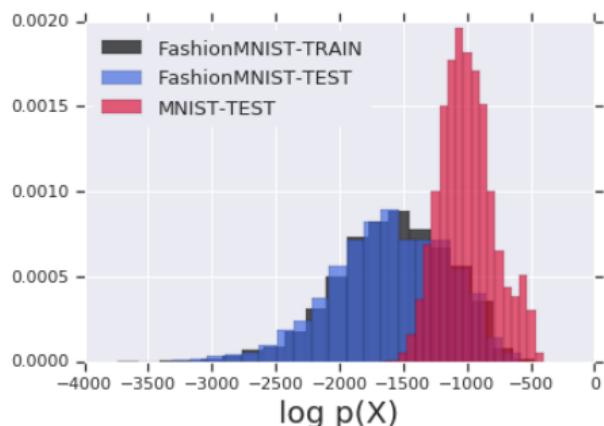
- ▶ Comment s'évaluer par rapport à une vérité terrain?
- ▶ Comment évaluer la confiance du système / vraisemblance de la génération ?



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

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/prédictibilité

- ▶ 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.



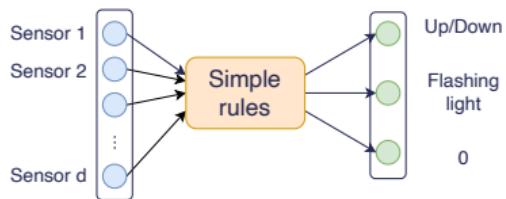
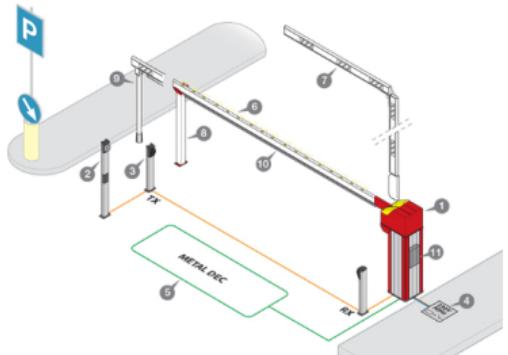
V and today?



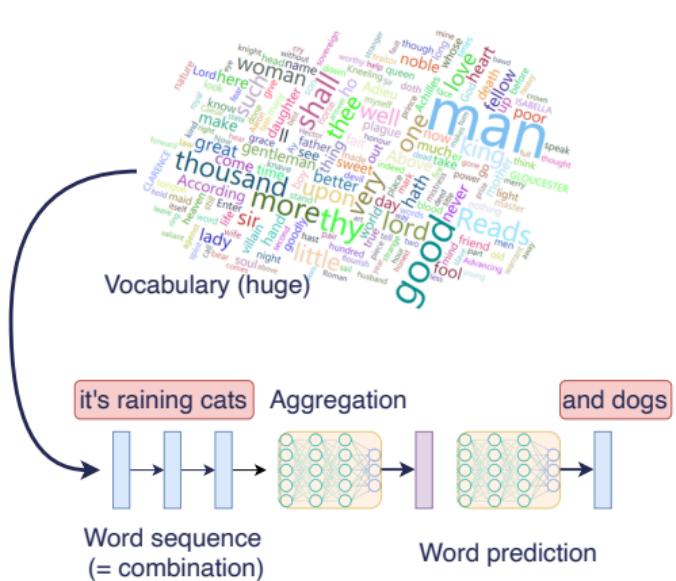
As a language model AI, I don't have real-time access to current dates. However, Barack



Stabilité, explicabilité... Et complexité



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



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

Transparence

- ▶ Les poids du modèle (*open-weight*)... ⇒ mais pas que les poids
- ▶ Les données d'entraînement (*BLOOM*) + distribution + instructions
- ▶ Techniques d'apprentissage
- ▶ Evaluation

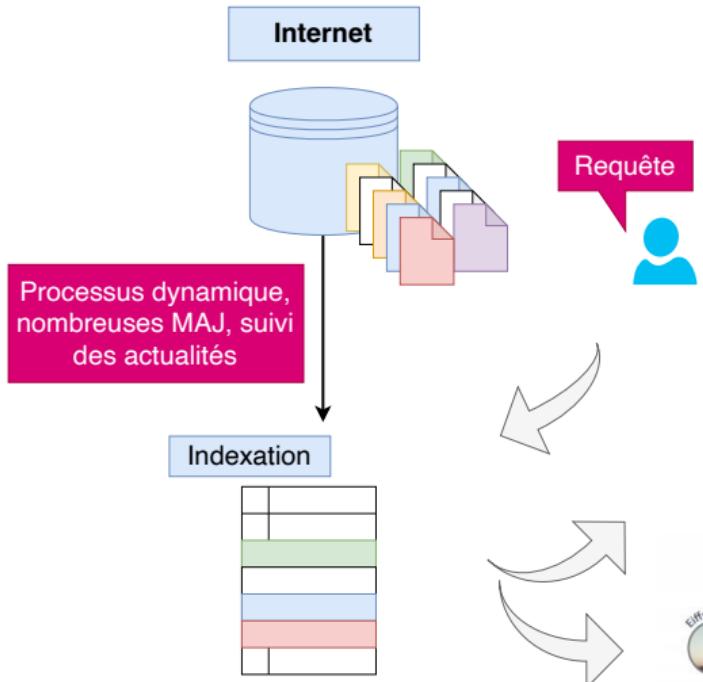
Foundation Model Transparency Index Scores by Major Dimensions of Transparency, 2023

Source: 2023 Foundation Model Transparency Index

		Meta	BigScience	OpenAI	stability.ai	Google	ANTHROPIC	cohere	AI21labs	Inflection	amazon	Average
		Llama 2	BLOOMZ	GPT-4	Stable Diffusion 2	PaLM 2	Claude 2	Command	Jurassic-2	Inflection-1	Titan Text	
Major Dimensions of Transparency	Data	40%	60%	20%	40%	20%	0%	20%	0%	0%	0%	20%
	Labor	29%	86%	14%	14%	0%	29%	0%	0%	0%	0%	17%
	Compute	57%	14%	14%	57%	14%	0%	14%	0%	0%	0%	17%
	Methods	75%	100%	50%	100%	75%	75%	0%	0%	0%	0%	48%
	Model Basics	100%	100%	50%	83%	67%	67%	50%	33%	50%	33%	63%
	Model Access	100%	100%	67%	100%	33%	33%	67%	33%	0%	33%	57%
	Capabilities	60%	80%	100%	40%	80%	80%	60%	60%	40%	20%	62%
	Risks	57%	0%	57%	14%	29%	29%	29%	29%	0%	0%	24%
	Mitigations	60%	0%	60%	0%	40%	40%	20%	0%	20%	20%	26%
	Distribution	71%	71%	57%	71%	71%	57%	57%	43%	43%	43%	59%
Usage Policy		40%	20%	80%	40%	60%	60%	40%	20%	60%	20%	44%
Feedback		33%	33%	33%	33%	33%	33%	33%	33%	33%	0%	30%
Impact		14%	14%	14%	14%	14%	0%	14%	14%	14%	0%	11%
Average		57%	52%	47%	47%	41%	39%	31%	20%	20%	13%	

USAGES

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
- francetvinfo
- Franceinfo - Actualités en temps réel et info en direct
- 20 Minutes

Information, informations - Dictionnaire de français Larousse

À propos des extraits optimisés • Commentaires

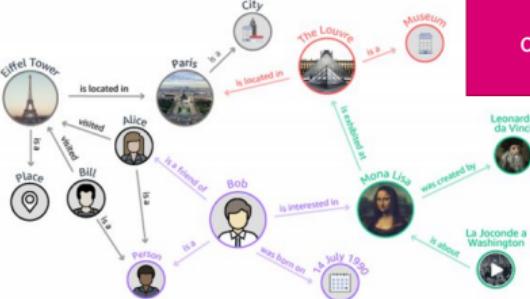
Francetvinfo

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

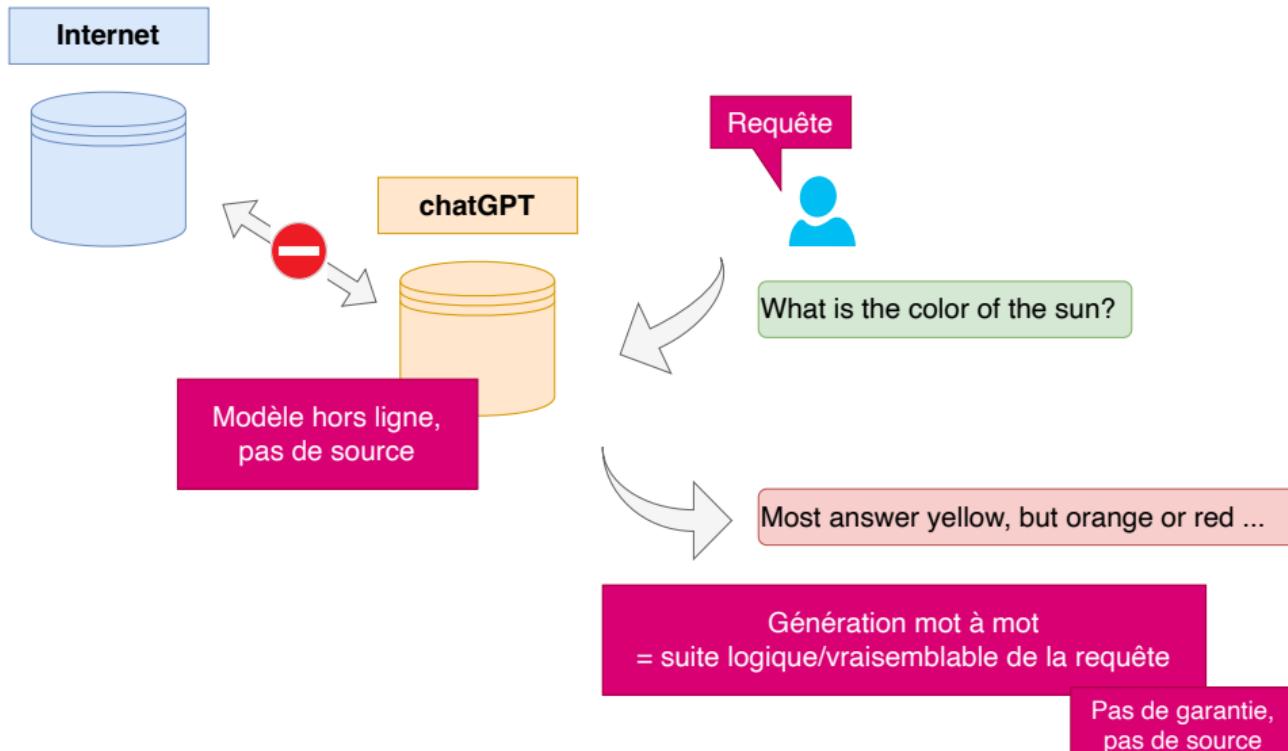
Résultats sourced



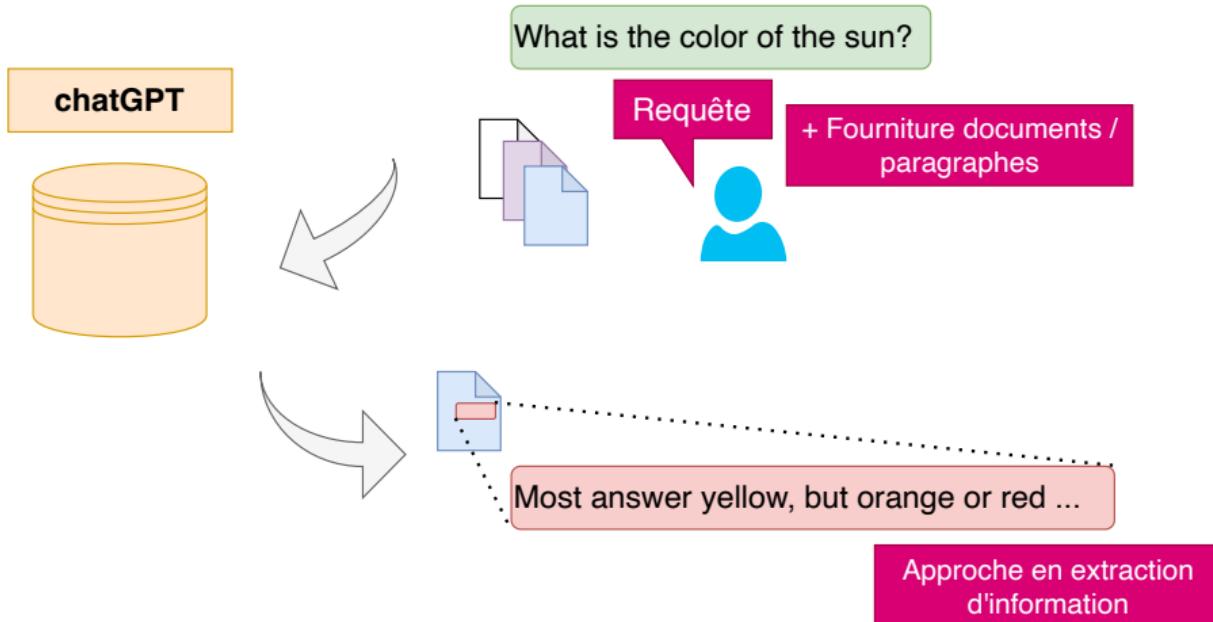
Graphe de connaissances vérifiées

Usage en accès à l'information

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



Usage en accès à l'information

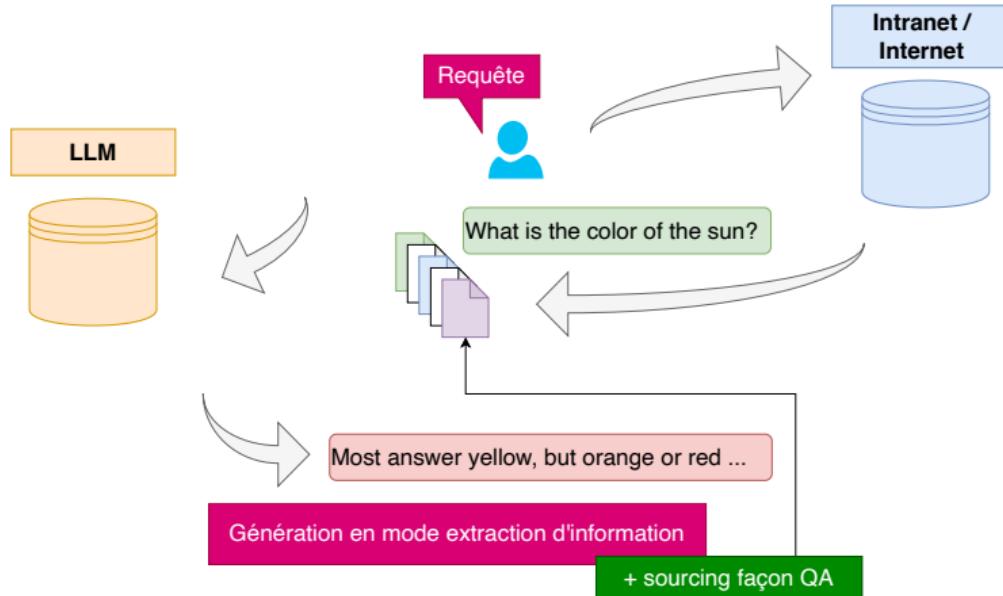


- Requête web + analyse, résumé automatique, reformulation, compte-rendus de réunion...
- Limite (actuelle) sur la taille des entrées (2k puis 32k tokens)

Usage en accès à l'information

- Possibilité d'ajouter des *post-traitements*

e.g. Bing



- RAG: Retrieval Augmented Generation
- Limite (actuelle) sur la taille des entrées (2k puis 32k tokens)

Les autres usages des IA Génératives

Un formidable outil de
mise en forme

Idée importante
Prise de notes
Elements de CV...



- ▶ Assistant personnel
 - ▶ Courrier standard, lettres de recommandation, de motivation, de résiliation, ...
- ▶ Compte-rendu de réunion
 - ▶ Mise en forme des notes
- ▶ Rédaction d'article scientifique
 - ▶ Rédaction des idées, en français, en anglais

Les autres usages des IA Génératives

Et un outil de **réflexion** !

- ▶ Brainstorming
 - ▶ Développement argumentaire, Recherche de contradiction
- ▶ Assistant pour le développement informatique
 - ▶ Génération de code, recherche d'erreurs, ...
 - ▶ Documentation
- ▶ Assistant pédagogique
 - ▶ Wikipédia ++, proposition de plan pour des dissertations,
 - ▶ Explication de code / proposition de corrections
- ▶ Analyse de document
 - ▶ Extraction d'information, question-réponse, ...



RISQUES

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

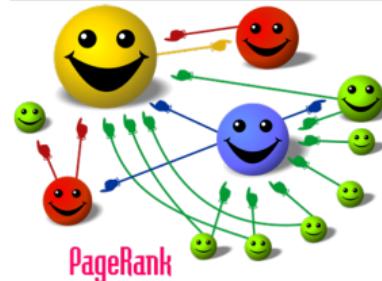
Harms that arise from environmental or downstream economic impacts of the language model.

Accès à l'information

- ▶ Accès aux informations dangereuses / interdites
- ▶ Autorités des informations
 - ▶ Nature: inconsciemment, image = vérité
 - ▶ Source: journaux, réseaux sociaux, ...
 - ▶ Masse: nombre de variantes, de citations (pagerank)
- ▶ Génération de texte: harcèlement...
- ▶ Risque d'anthropomorphisation de l'algorithme
 - ▶ Distinguer l'homme de la machine



DROIT À L'OUBLI



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

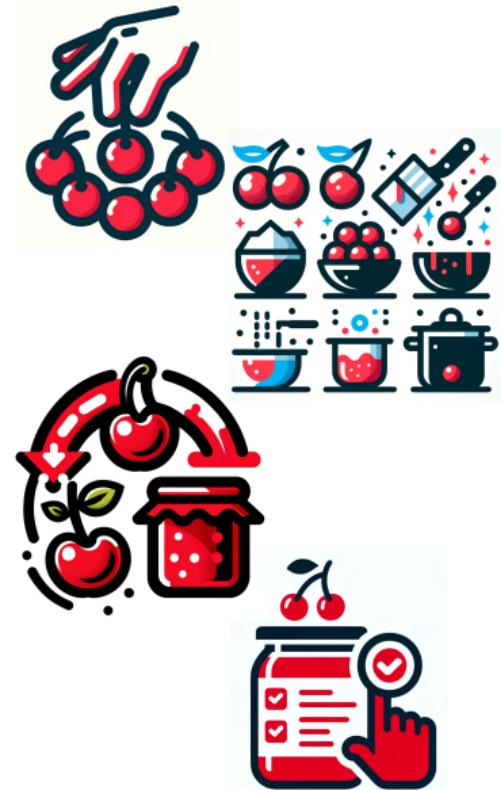
⇒ Quelle légitimité? Quelle transparence? Quelle efficacité?



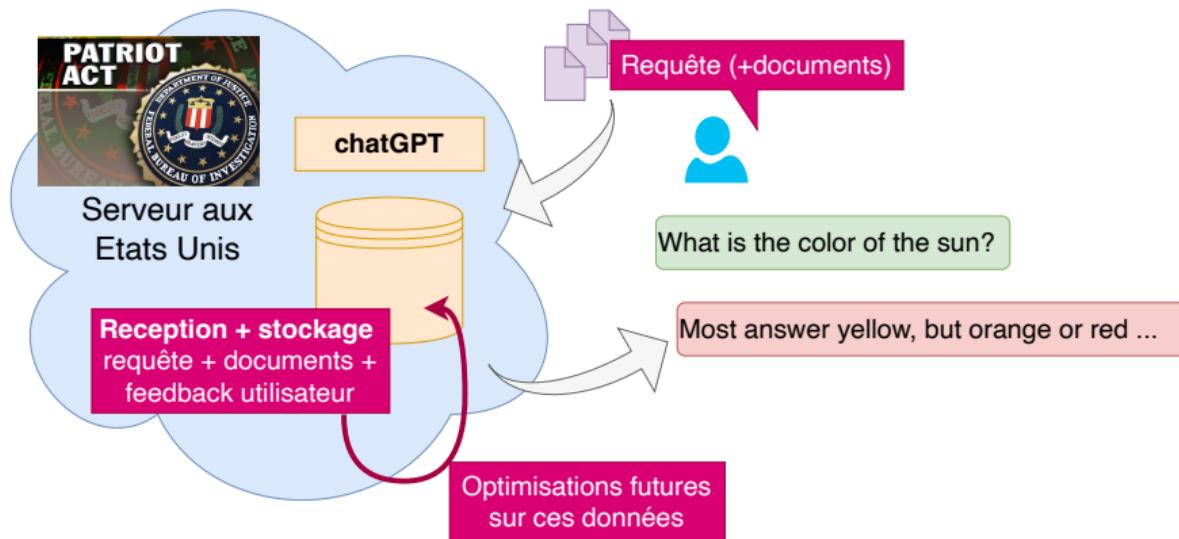
Le machine learning n'est jamais neutre

- 1 Sélection des données
- 2 Transformation des données
- 3 A priori & choix des opérateurs
- 4 Filtrage des sorties

⇒ Des choix qui orientent les résultats de l'algorithme



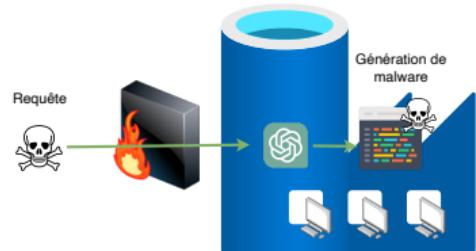
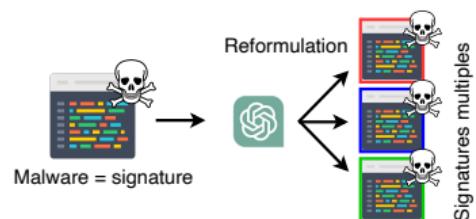
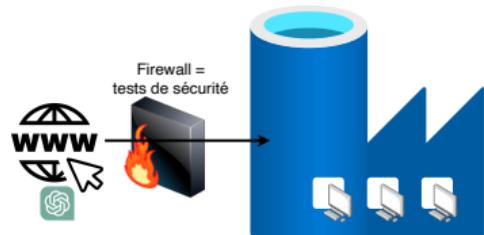
Fuite(s) de données



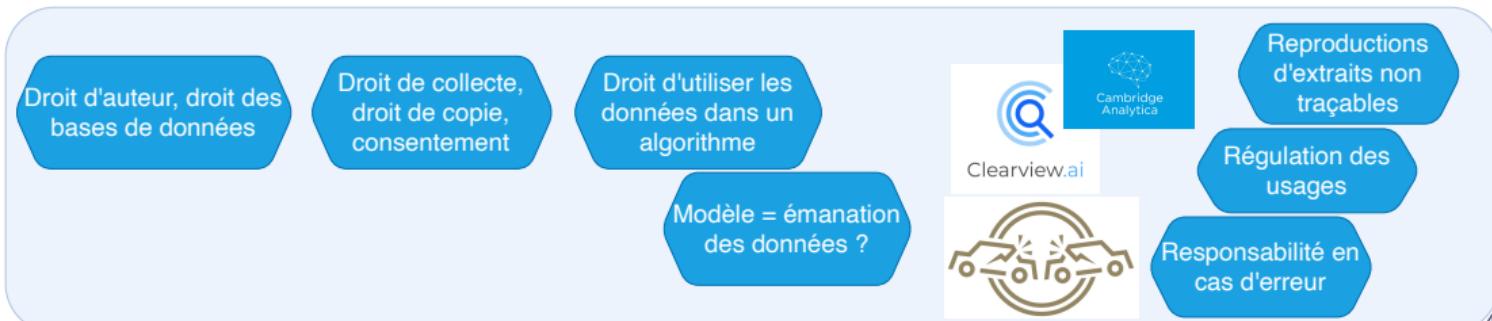
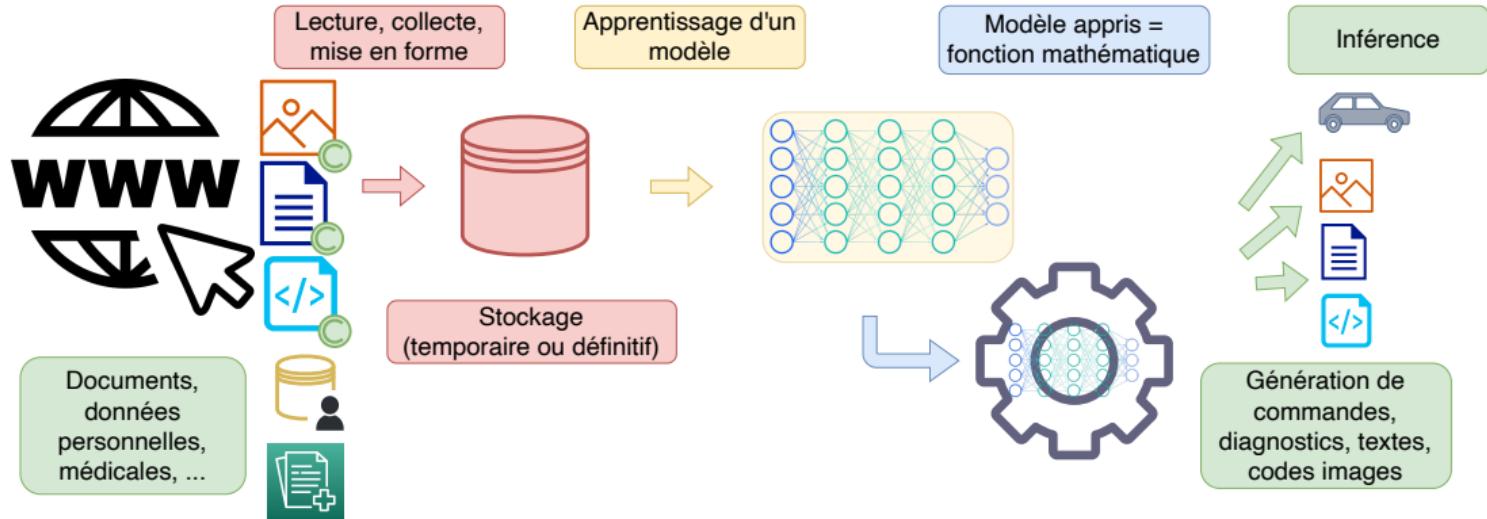
- ▶ Transfert de données sensibles
- ▶ Exploitation des données par OpenAI (ou d'autres)
- ▶ Fuite des données dans les modèles futurs

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



Risques/Questions juridiques



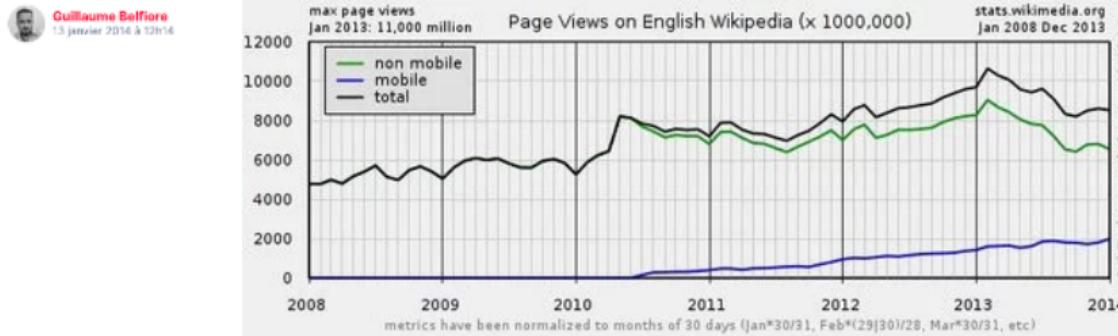
Questions économiques

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



Comment aborder la question de l'éthique?

Médecine

- 1 **Autonomie** : le patient doit pouvoir prendre des décisions éclairées.
- 2 **Bienfaisance** : obligation de faire le bien, dans l'intérêt des patients.
- 3 **Non-malfaisance** : éviter de causer du tort, évaluer les risques et les bénéfices.
- 4 **Justice** : équité dans la distribution des ressources de santé et des soins.
- 5 **Confidentialité** : confidentialité des informations des patients.
- 6 **Vérité et transparence** : fournir des informations honnêtes, complètes et compréhensibles.
- 7 **Consentement éclairé** : obtenir le consentement libre et éclairé des patients.
- 8 **Respect de la dignité humaine** : traiter tous les patients avec respect et dignité.

Intelligence Artificielle

- 1 **Autonomie** : l'Humain contrôle le processus
- 2 **Bienfaisance** : en incluant l'environnement?
- 3 **Non-malfaisance** : Humain + environnement / soutenabilité / usages malveillants
- 4 **Justice** : accès à l'IA et égalité des chances
- 5 **Confidentialité** : quid du modèle économique Google/Facebook ?
- 6 **Vérité et transparence** : le drame de l'IA moderne
- 7 **Consentement éclairé** : des cookies aux algorithmes, savoir quand on est au contact d'une IA
- 8 **Respect de la dignité humaine** :

Comment aborder la question de l'éthique?

Médecine

- 1 **Autonomie** : le patient doit pouvoir prendre des décisions éclairées.
- 2 **Bienfaisance** : obligation de faire le bien, dans l'intérêt des patients.
- 3 **Non-malfaisance** : éviter de causer du tort, évaluer les risques et les bénéfices.
- 4 **Justice** : équité dans la distribution des ressources de santé et des soins.
- 5 **Confidentialité** : confidentialité des informations des patients.
- 6 **Vérité et transparence** : fournir des informations honnêtes, complètes et compréhensibles.
- 7 **Consentement éclairé** : obtenir le consentement libre et éclairé des patients.
- 8 **Respect de la dignité humaine** : traiter tous les patients avec respect et dignité.

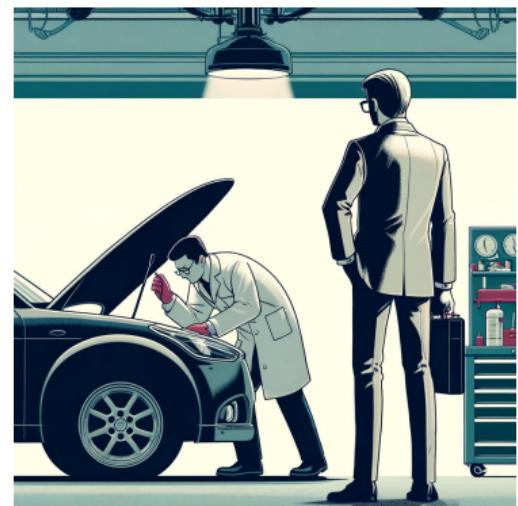
Intelligence Artificielle

- 1 **Autonomie** : l'Humain contrôle le processus
- 2 **Bienfaisance** : en incluant l'environnement?
- 3 **Non-malfaisance** : Humain + environnement / soutenabilité / usages malveillants
- 4 **Justice** : accès à l'IA et égalité des chances
- 5 **Confidentialité** : quid du modèle économique Google/Facebook ?
- 6 **Vérité et transparence** : le drame de l'IA moderne
- 7 **Consentement éclairé** : des cookies aux algorithmes, savoir quand on est au contact d'une IA
- 8 **Respect de la dignité humaine** :

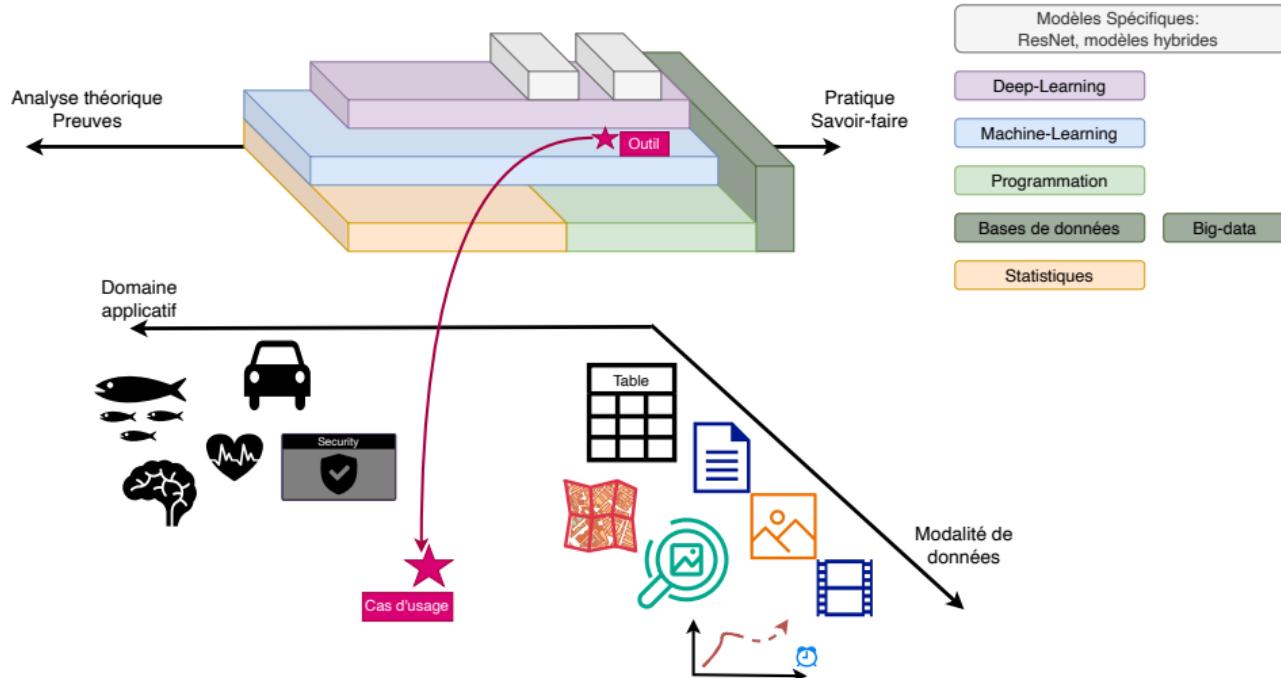
IA & PÉDAGOGIE

Enseignement de l'IA vs enseignement avec l'IA

- ▶ Enseignement **de** l'IA = arbre de compétences à empiler + spécialiser
 - ▶ Programmation / Statistiques
 - ▶ Machine-Learning (modèles de base + protocoles & métriques)
 - ▶ Deep-learning
 - ▶ Modalités spécifiques : image, textes, séries temporelles
 - ▶ Domaines spécifique : Biologie, médecine, finance, mécanique, ...
 - ▶ Approfondissement théorique : optimisation, intervalles de confiances, convergences, ...
- ▶ Enseignement **avec** l'IA
 - ▶ Nouvelles opportunités
 - ▶ Nouveaux risques
 - ▶ Nouvelles contraintes



Enseignement de l'IA



- ▶ Différents niveaux d'accès (acculturation, outil, développement)
- ▶ Différents domaines applicatifs

LLM : un outil pédagogique?

chatGPT fait-il les devoirs à votre place?

- ▶ LLM = mémoire (partielle) d'internet
 - ▶ Champion de la reformulation
 - ▶ Capacité à comprendre/traduire/générer du code informatique
 - ▶ Entrainer à répondre à de nombreux types de questions
- ⇒ Oui, ils vont répondre à beaucoup de choses...
En faisant régulièrement des petites/grosses **fautes**
- ⇒ Dans tous les cas, **grosse production de textes**, plutôt bien écrit



Paradigme de la calculatrice:
*s'il existe une machine,
pourquoi apprendre les
tables de multiplication?*

Questions pédagogiques: que faire ou ne pas faire?

Quels usages vertueux?

[niveau de base] Un prof présent 24/7

- ▶ **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** les lettres de motivation / égalité sociale (?)
 - ▶ Quelles sont les qualités pour...

Questionnement?

- ▶ Est-il possible de résister à la tentation de demander la réponse?
- ▶ Comment trier des lettres de motivation issues de chatGPT?



Questions pédagogiques: que faire ou ne pas faire?

Quels usages vertueux?

[niveau avancé] Un assistant pour aller plus loin

- ▶ Se concentrer sur le fond, les idées, l'architecture générale
- ▶ Laisser le robot aider sur la forme, la rédaction, l'écriture du code
- ▶ Plans d'exposé / vérifier potentiels oubli

[Prof]

- ▶ Braistorming & plan de cours
- ▶ Proposer des quizz / questions de cours
- ▶ Générer des illustrations



Questions pédagogiques: que faire ou ne pas faire?

Quels usages frauduleux?

- ▶ Ecrire une rédaction
- ▶ Générer un devoir d'histoire
- ▶ Traiter un sujet présent dans wikipedia
- ▶ Produire une analyse de document(s)



- ▶ Connaissances générales (LLM compétent)
 - ▶ Devoir orienté sur la mise en forme
(LLM très compétent)
 - ▶ Analyse de documents fournis (LLM assez compétent)
- ⇒ **LLM = cœur du sujet ⇔ Compétences perso** ↴



Questions pédagogiques: que faire ou ne pas faire?

Quels usages dangereux?

- ▶ Poser des questions pointues / sur des connaissances rares / bibliographies / textes de loi

⇒ **Hallucinations garanties !**

- ▶ Générer du code... + ne pas le tester en profondeur

⇒ **Des bugs bien gratinés qui vont coûter cher!**



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 :

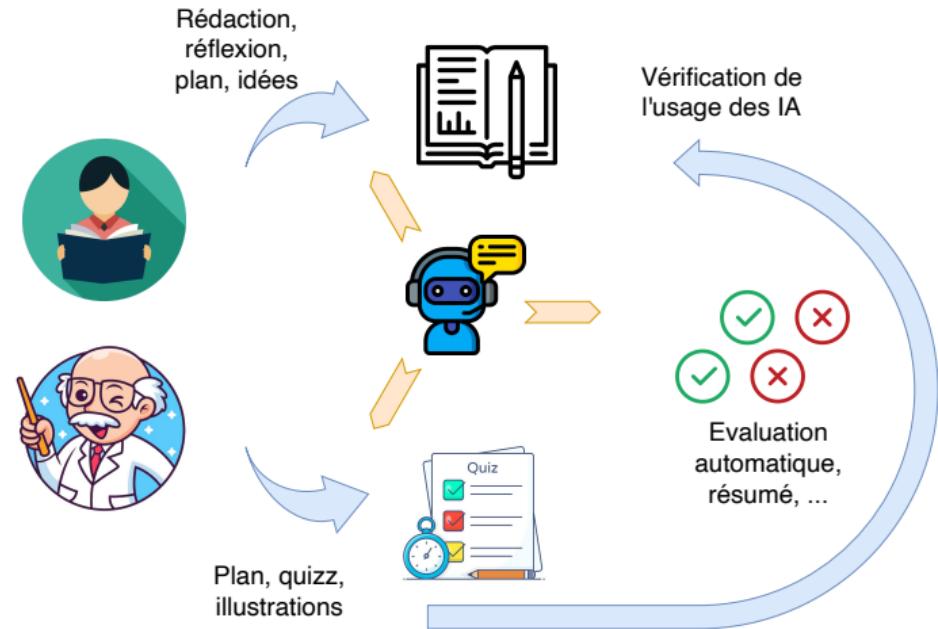
1. "Variational Autoencoder for Deep Learning of Images, Labels and Sequences" par Zhao et al. (2017) - Cette étude présente une approche VAE pour la modélisation de séquences temporelles en utilisant une architecture de réseaux de neurones récurrents.
2. "Variational Recurrent Autoencoders" par Chung et al. (2016) - Cette étude introduit une approche VAE pour la modélisation de séquences temporelles en utilisant des réseaux de neurones récurrents et une variational recurrent auto-encoders
3. "Generative Modeling for Time Series: Variational Recurrent Auto-Encoder (VRAE)" by 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 ...
☆ Enseignement, RD Citer, CITA 202 fois, Autres articles, Les 2 versions, ▾

Risques de généralisation de l'IA

IA partout = perte de sens?

- ▶ Transposition aux RH
- ▶ Aux systèmes de financement par projet



Détection des textes générés par chatGPT

The screenshot shows a Microsoft Word document titled "Plagiat Def 2 -- Saved to my Mac". The document contains a paragraph about externalities and a note at the bottom: "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 conscientement 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 Detection Report:

Document	Date	Mots	Ko	Plagiat (%)
Plagiat Def 2 #4483eb	07/01/2023 19:18	122 mots	19,47 ko	0%
Plagiat Def 1 #f90ff3	07/01/2023 19:16	135 mots	16,78 ko	100%

Définition de Wikipedia

Crédit: S.
Pajak

Détection des textes générés par chatGPT

GPTZero

Detect AI Plagiarism. Accurately



ORIGINALITY.AI

Chat GPT



AI Detector

- ▶ **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**)
- ▶ **δ-vraisemblance** sur des textes perturbés ([DetectGPT](#))
- ▶ **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*)

Et au-delà de chatGPT? les outils pédagogique

Les EIAH –Environnements Informatiques pour l'Apprentissage Humain– n'ont pas attendu chatGPT pour proposer des solutions !

- ▶ Plateforme + outils (e.g. moodle)
 - ▶ Pédagogie
 - ▶ + Gestion
- ▶ Manuels scolaires augmentés
- ▶ Réseaux sociaux (pédagogiques) / Slack
- ▶ Outil de publication vidéo (pédagogie inversée)
- ▶ Sondage en direct, avec/sans téléphone



moodle

Explain
Everything
wooclap

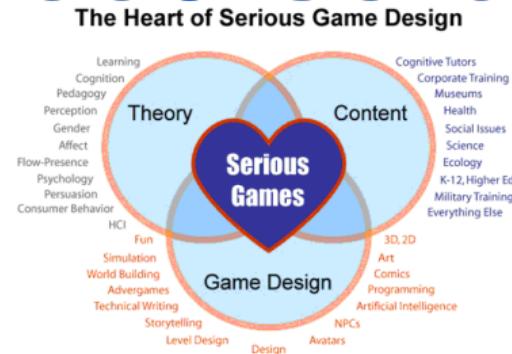
Direct
Poll

Ne pas confondre *numérique* & IA

Et au-delà de chatGPT? les outils pédagogique (suite)

Les EIAH –Environnements Informatiques pour l'Apprentissage Humain– n'ont pas attendu chatGPT pour proposer des solutions !

- ▶ Gamification des enseignements / serious game
 - ▶ Place du numérique
 - ▶ Pour quelles connaissances?
- ▶ MOOC / SPOC
 - ▶ FUN –France Université Numérique—



Réflexions par rapport à la crise du COVID...
... Et par rapport à la crise universitaire !

Ne pas confondre *numérique* & *IA*



Et au-delà de chatGPT? La correction automatique

- QCM + scanner (+numéro d'anonymat) ⇒ correction automatique

+ 2/188+

QCM — Nombres relatifs — Somme

NOM : AVOGADRO PRÉNOM : AMEDEO CLASSE : 401 DATE : LUNDI 2 SEPTEMBRE 2021

Pour chaque des questions, une seule réponse est possible. Vous devez marquer la case qui correspond à la bonne réponse.

QUESTION 1 : Calculer $(+13) + (-7) + (-3)$

(-2) (+2) 0 (-1) (+3) (-3)

QUESTION 2 : Calculer $(-7) + (+11) + (-3)$

(+3) 0 (-1) (-2) (+2) (+1)

QUESTION 3 : Calculer $(-11) + (-3) + (+3) + (+7)$

(+3) (+2) (-2) (+4) (-4) (-3)

QUESTION 4 : Calculer $(-5) + (-3) + (+9)$

(+2) 0 (+3) (-2) (-1) (+1)

QUESTION 5 : Calculer $(-6) + (+8) + (-11) + (-7) + (+3)$

(-13) (+13) (+11) (-12) (-11) (-9)

QUESTION 6 : Calculer $(+3) + (-7) + (+8) + (-5)$

(+1) (+23) (+3) (-3) (-1) (-2)

QUESTION 7 : Calculer $(-5) + (-3) + (+7) + (-3) + (+11)$

(+5) (-7) (+6) (-5) (-6) (+7)

QUESTION 8 : Calculer $(+3,76) + (-7,4) + (+8,112) + (-3,76) + (-8,112)$

(-11,780) 0 (-7,4) (+3,76) (-8,112) (+3,567)

QUESTION 9 : Calculer $(-1,1) + (-3,1) + (+2,6) + (-1,3)$

(+2,9) (+1,2) (+2,4) 0 (-1,2) (-2,9)

QUESTION 10 : Calculer $(-7) + (-8) + (+15) + (+7) + (-15)$

(-4) (-8) (+7) (-15) (+15) (-7)



Et au-delà de chatGPT? La correction automatique

- ▶ QCM + scanner (+numéro d'anonymat) ⇒ correction automatique
- ▶ Test unitaire (un mode de développement informatique à part entière)

The screenshot shows a developer's workspace with three main windows:

- Test Explorer:** Shows a list of tests. Under "Failed Tests" (1), there is one entry: "Test_MultiplyMethod" with a duration of 99 ms. Below it, under "Passed Tests" (3), are three entries: "Test_AddMethod" (5 ms), "Test_DivideMethod" (< 1 ms), and "Test_SubtractMethod" (< 1 ms). The status bar at the bottom indicates an elapsed time of 99 ms.
- UvITest1.cs:** A C# file containing a single method: `Test_MultiplyMethod`. It includes a failing assertion: `Message.Assert.AreEqual failed. Expected:<20>, Actual:<100>`.
- BasicMath.cs:** A C# file containing four methods: `Add`, `Subtract`, `Divide`, and `Multiply`. The `Multiply` method has a note: `// To trace error while testing, writing + operator instead of * operator.`

Three green checkmarks are overlaid on the code editor, indicating successful validation or completion of specific parts of the code.

Ne pas confondre *numérique* & IA'

Et au-delà de chatGPT? La correction automatique

- ▶ QCM + scanner (+numéro d'anonymat) ⇒ correction automatique
- ▶ Test unitaire (un mode de développement informatique à part entière)
- ▶ Analyse du texte: comprendre si une réponse *contient* la vraie réponse ou pas

P^a	A senior is waiting at the window of a restaurant that serves sandwiches.	Relationship
H^b	A person waits to be served his food.	Entailment
	A man is looking to order a grilled cheese sandwich.	Neutral
	A man is waiting in line for the bus.	Contradiction
^a P, Premise.		
^b H, Hypothesis.		

Ne pas confondre *numérique* & *IA'*

Et au-delà de chatGPT? La correction automatique

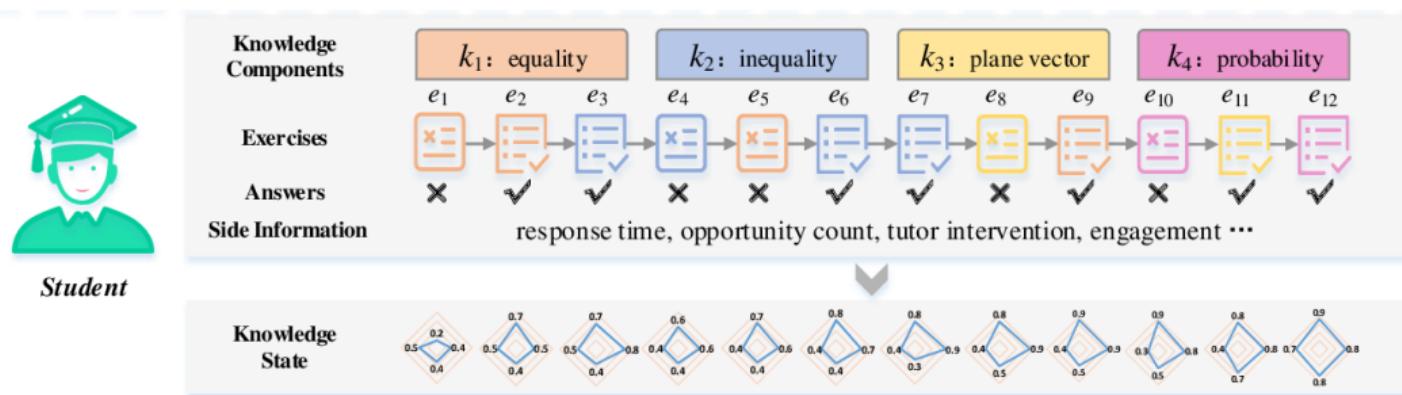
- ▶ QCM + scanner (+numéro d'anonymat) ⇒ correction automatique
- ▶ Test unitaire (un mode de développement informatique à part entière)
- ▶ Analyse du texte: comprendre si une réponse *contient* la vraie réponse ou pas
- ▶ Résumé automatique
⇒ on revient à chatGPT



Ne pas confondre *numérique* & *IA*'

Et au-delà de chatGPT? Knowledge Tracing

- ▶ Système de recommandation... Pour l'apprenant
 - ▶ Vers une **pédagogie adaptative...**
 - ▶ ... ⇔ modéliser tous les concepts + niveaux de difficultés
 - ▶ Quel coût pour modéliser l'ensemble d'un cours?
 - ▶ Quel apport réel pour l'étudiant?
- PB mesure de performance d'un système de reco

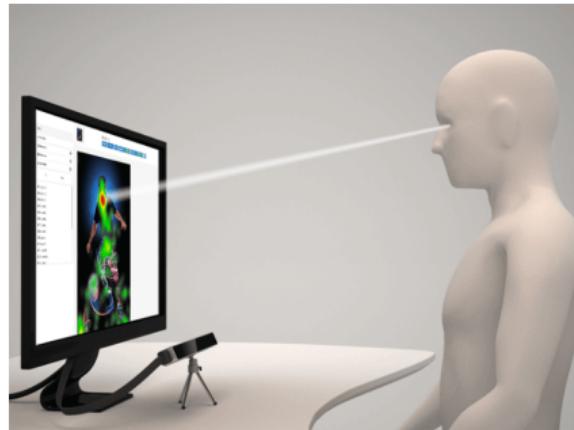


Et au-delà de chatGPT? Modélisation utilisateur / eye tracking

Il est néanmoins possible d'aller assez loin dans la modélisation de l'apprenant

- ▶ Internet / cookies
- ▶ Eye tracking
- ⇒ Comprendre comment l'apprenant fonctionne pour mieux l'aider

- ▶ Analyse de la lecture, étude de la concentration



DISCUSSION AVEC LES ÉTUDIANTS

Trans. numériques : de l'ordinateur à la programmation

1950-60 Ordinateur à carte [IBM 610, \$55k]

1970 Datapoint 2200 (écran/clavier/2ko mémoire) [\$5k ≈ \$38k (2023)]

fin 1970 Apple II (1Mhz, 4KB ram) [\$1.3k (≈ \$6.5k (2023))]

2000 Généralisation (750MHz, 128Mo Ram, 20Go DD, GPU) [<\$1k]

1970 100-200k Développeurs en informatique

2000 2-3M Développeurs

2020 Tous les étudiants sont-ils des développeurs ?

Le cadre des possibles

- ▶ Matériel disponible
- ▶ Environnement convivial
(Fortran⇒python/VSC/Jupyter)
- ▶ Bibliothèques à tous les niveaux
- ▶ Formation généralisée

Trans. Numériques : révolution logistique

1908 Travail à la chaîne / Fordisme

1950-70 Flux tendus / Juste à temps / Kanban

1980-90 Optimisation chaîne de fournisseurs

1980-... Robotisation

2010 Robotisation de la logistique

1987 Mappy (minitel !)

2005 Google map

2005 Amazon Web Service

2010 Révolution des entrepôts Amazon

Optimisation de tous les processus:

- ▶ une clé technique de la mondialisation
- ▶ qui vous guide aussi au quotidien

Trans. Numériques : la place des réseaux

1969	ARPANET	1995	Amazon — Altavista
1989	World Wide Web (WWW)/ hypertexte	1998	Google
2000	Métier le plus coté: <i>ingénieur réseau</i>	2001	Wikipedia
>2005	Les années Internet : Google, Amazon, Facebook, Apple, Netflix	2003	Skype
		2004	Gmail
		2006	Google Docs

Transformation des usages: distances réduites, instantanéité

- ▶ Recherche d'informations
- ▶ Mail, Travail collaboratif,
- ▶ Télé-travail, shopping

Trans. numériques : les promesses des modèles mécanistes

- 1940 Premier modèles à éléments finis
- 1992 1er Super-calculateur chez Météo-France
- 1996 Fin des essais nucléaires français
- 1998 Catia V5 disponible sur toutes les plateformes

>2010 Dassault Systèmes = 1er éditeur Français de logiciel
(2ème derrière SAP en Europe)

2020 Tout le monde peut développer des modèles mécanistes?

Impact sociétal important

- ▶ Optimisation de structure,
- ▶ Crash test,
- ▶ Prévision météo, climatologie, ...

Trans. Numériques : la place de l'IA

- 1956 Invention du terme *Intelligence Artificielle*
- 1986 Rétropropagation du gradient / réseaux de neurones
- 2001 Création du terme *Big Data*
- 2020 Multiplication des modèles hybrides (mécaniste/symbolique/numérique)
- 1997 IBM Deep Blue bat Kasparov
- 2012 Victoire de l'architecture *Deep-Learning* à ImageNet
- 2016 AlphaGo bat Lee Sedol
- 2022 Lancement de chatGPT

Un outil omniprésent:

- ▶ Reconnaissance vocale/écrite
- ▶ Moteur de recherche / recommandation
- ▶ Imagerie/diagnostic Médical
- ▶ Véhicule intelligent
- ▶ Usine 4.0
- ▶ Robotique

Des outils et des questions

Des outils nouveaux:

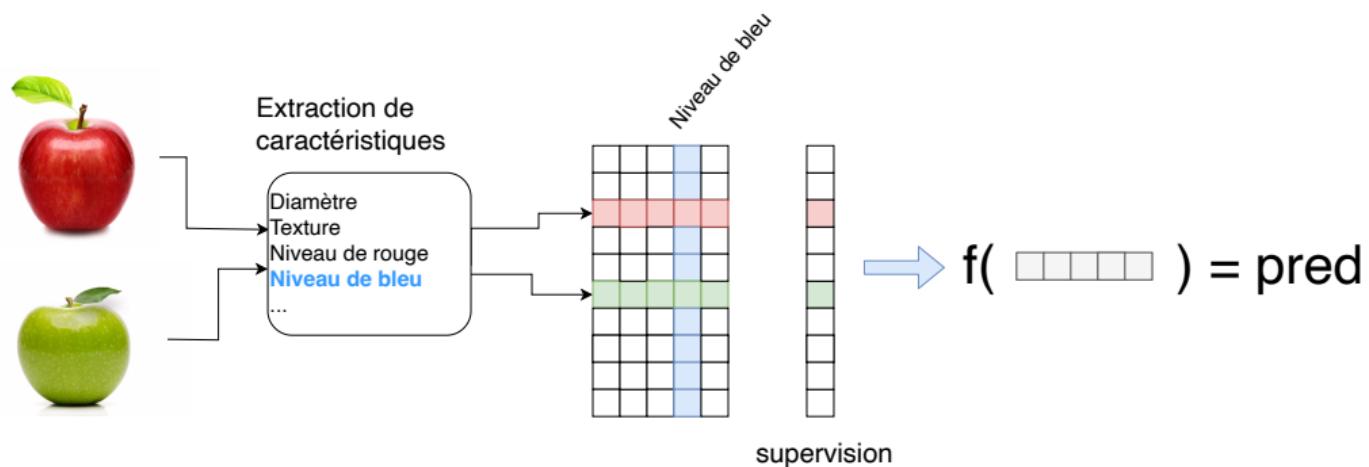
- ▶ Nouvelles manières de traiter des problèmes existants
- ▶ Aborder de nouveaux problèmes
- ▶ ... Mais évidemment, ça ne marche pas à tous les coups !
- ▶ L'IA fait souvent des erreurs (assistant *vs* remplaçant)

Apprendre à utiliser un système d'IA

- ▶ IA non adaptée à beaucoup de problème
- ▶ IA = une part du problème (+interface, exploitation, acceptation...)

Le défi des données

- 1 Identifier les données qui font un problème
- 2 Disponibilité & collecte des données
 - Souvent le plus cher/délicat dans un projet IA
 - Sciences participatives
 - Collecte des données: numérique, réseau et interface
 - Amazon Mechanical Turk et ses alternatives



Maturité des outils & environnements

Des outils (plus) matures

- ▶ **Environnements:** Jupyter, Visual Studio Code, ...
 - ▶ **Machine-Learning** Scikit-Learn : des briques à assembler
 - ▶ Formation : 1 semaine
 - ▶ Réalisation d'un projet : quelques heures à quelques jours
 - ▶ **Deep-Learning** pytorch, tensorflow : des briques... mais plus complexes
 - ▶ Formation : 2-5 semaines
 - ▶ Réalisation d'un projet : quelques jours à quelques mois
 - ▶ Obligatoire pour le texte et l'image
-
- ▶ Un projet sur les données = 10 ou 100 fois moins de temps / 2005
 - ▶ Développer un projet est **accessible aux non-informaticiens**

Niveaux d'accès à l'Intelligence Artificielle

1 Utilisateur via une interface: *chatGPT*

- ▶ **ATTENTION:** il faut quand même une petite formation (2-4h)

2 Utiliser des bibliothèques en python

- ▶ Bases sur les protocoles
- ▶ Chaines de traitements standard
- ▶ Formation: 1 semaine-3 mois (ML/DL)

3 Développeur d'outils

- ▶ Adapter les outils à un cas spécifique
- ▶ Intégrer des contraintes métiers
- ▶ Construire des systèmes hybrides (mécaniste/symbolique)
- ▶ Méler du texte et des images
- ▶ Formation: ≥ 1 an

Retours d'expériences

[Journée du 5 juin 2024 / Univ. Paris-Saclay]

- ▶ Comment former les étudiants?
 - ▶ Mise en place de cours à tous les niveaux / base obligatoire
- ▶ Confronter les étudiants à l'outils
 - ▶ Ecrire un texte, travailler sur le design, ...
 - ▶ Avec/sans outils: quel rapport à la tâche?
 - ▶ Désinhibition? Amélioration? Déterioration? Plaisir?
 - ▶ Formation aux nouveaux outils et productivité
 - ▶ Le rapport à l'information
- ▶ Comment évaluer les étudiants?
 - ▶ Réflexion/prise de recul sur les contenus ± importants
- ▶ Mettre à disposition des outils
 - ▶ LLM à l'université (?)
 - ▶ Traduction
- ▶ Acceptabilité de l'IA
 - ▶ Ethique, règlementation, accès aux outils

Les questions aux étudiants

- ▶ Quel ressenti?
 - ▶ Opportunité lié à de nouveaux outils?
 - ▶ Contraintes/obligation de regarder un objet désagréable?
 - ▶ Indifférence: ce n'est pas mon métier !
- ▶ Quels besoins? [actuel]
 - ▶ Comment voyez-vous la formation en IA?
 - ▶ De quoi estimez-vous avoir besoin?
- ▶ Quelle projection?
 - ▶ Opportunité / nouveaux métiers: une place en or pour les nouvelles compétences
 - ▶ Dystopique: le monde du travail au bord de l'effondrement?

INTEL. ARTIFICIELLE & RESSOURCES HUMAINES

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

Les défis en Ressources Humaines

Défis internes aux ressources humaines

1 Trier les candidatures / proposer des postes

- ▶ Système de recommandation / profiling
- ▶ Matching entre pré-requis et compétences
- ▶ Risque de biais

2 Filtrer les candidats avec des chatbots

Analyser automatiquement les lettres de motivation / CV

- ▶ Quels critères sont réellement mis en avant?
- ▶ Perte de sens si les candidats automatisent la rédaction des lettres et les réponses !

3 Gestion des mouvements et agilité

- ▶ Analyse massive des documents produits (quid du coût/fiabilité)
- ▶ Recenser les compétences en interne + matching de projet
 - ▶ concurrence de LinkedIn

4 Gestion des carrières

- ▶ Multiplication des indicateurs / dashboards / etc...
- ▶ Quels critères sont réellement mis en avant?

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

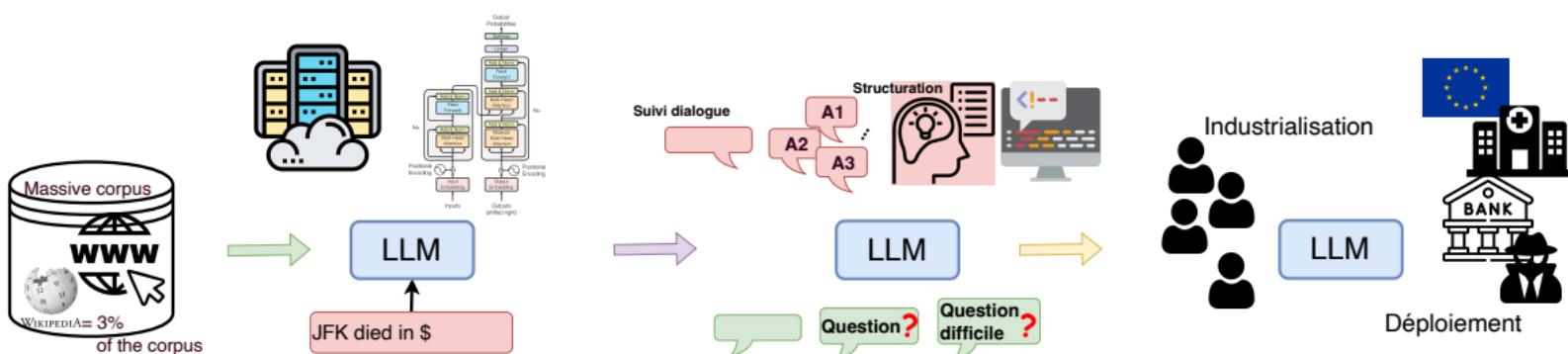
- Puissance machine (milliers de GPU)
- Architecture/recherche ML

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

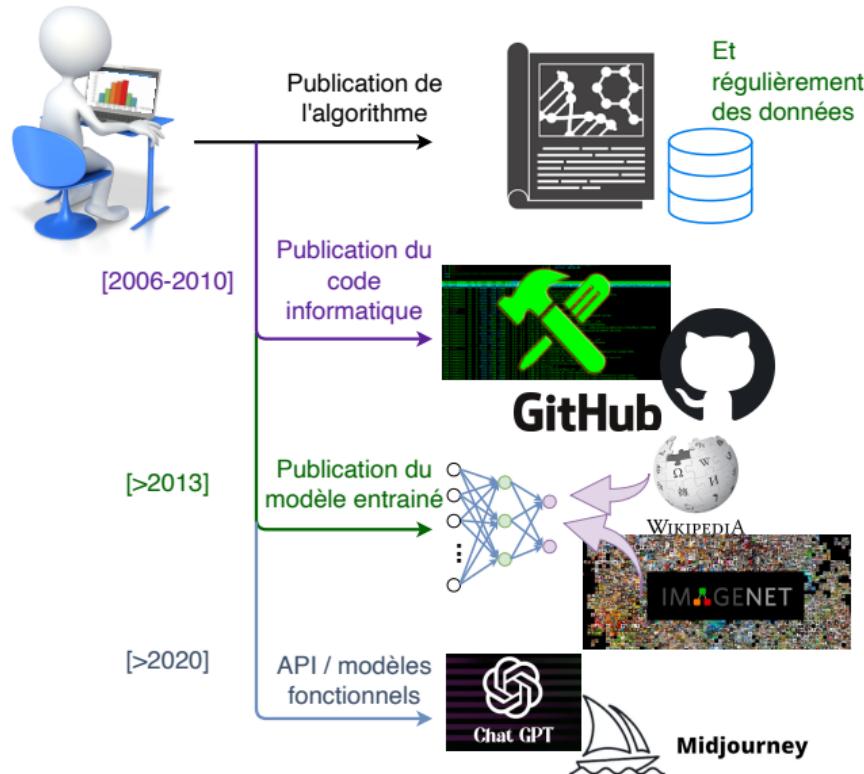
Ingénieur Machine Learning

- 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 :

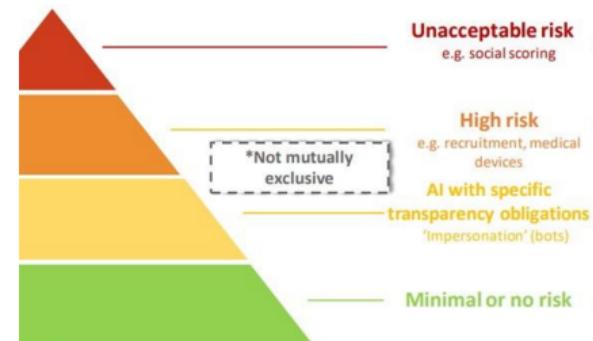
- ▶ Performances (facteur d'acceptation de chatGPT)
- ▶ Fiabilité / auto-évaluation

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

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 niveaux d'équilibre dans les marchés.

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.

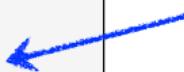
Overall, the phrase "soaring spending" refers to the high level of investment that the U.S. is making in chips, but it also implies that this level of investment is not unlimited.



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 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 conscientement 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

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 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]

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 votre temps et attention.

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.

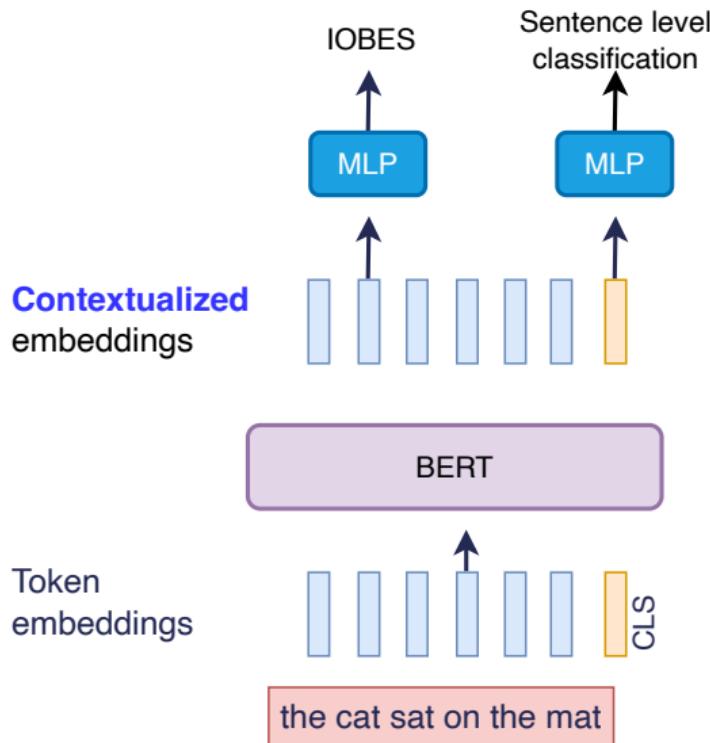


```
@article{boring2017gender,  
title={Gender Biases in Student Evaluations of Teaching},  
author={Boring, Anne},  
journal={Journal of Public Economics},  
volume={145},  
pages={27-41},  
year={2017}  
}
```



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 femme⁴, 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. thumb up thumb down

V how old is obama?

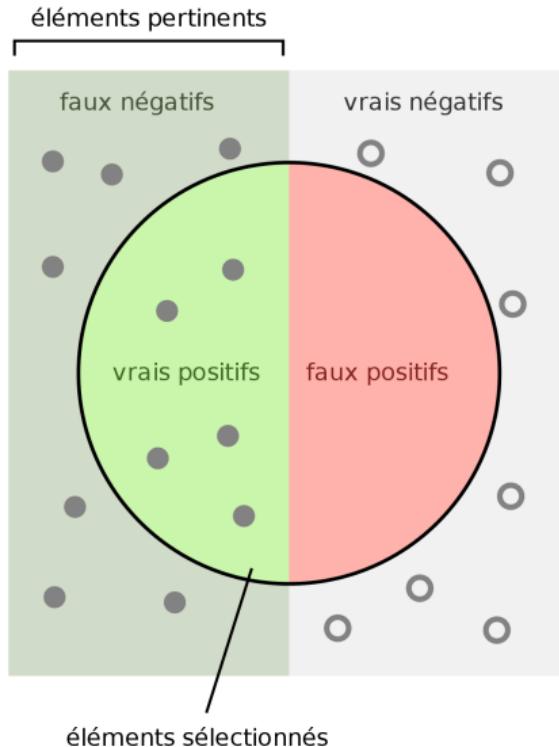
 As of 2021, Barack Obama was born on August 4, 1961, so he is 60 years old. thumb up thumb down

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 thumb up thumb down

MÉTRIQUE

Précision et rappel



Combien de candidats sélectionnés sont pertinents ?

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

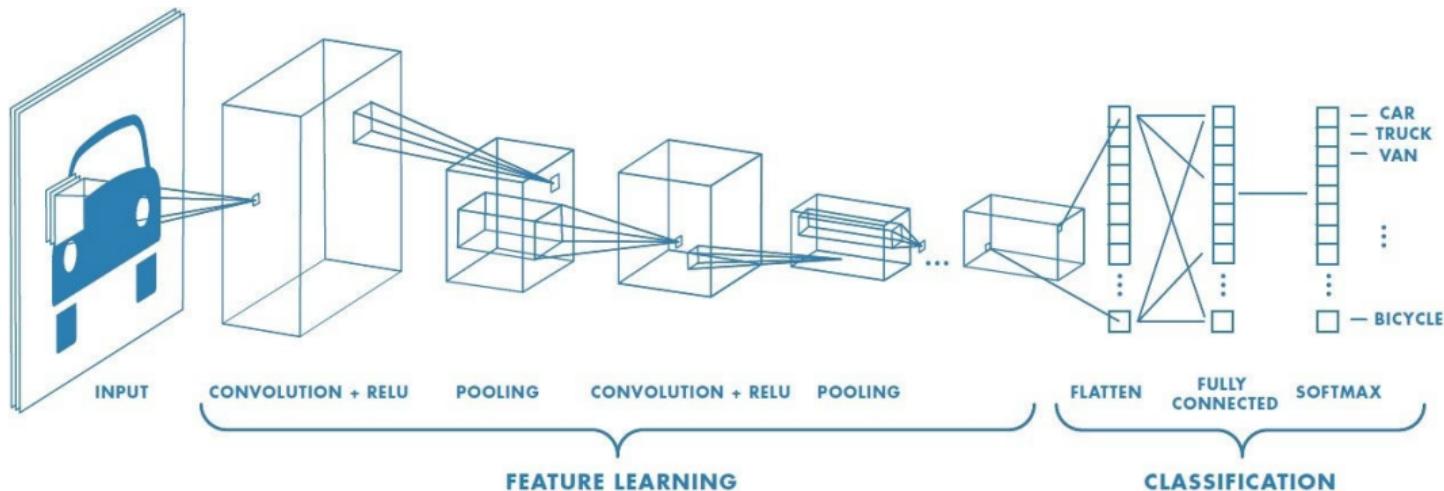
Combien d'éléments pertinents sont sélectionnés ?

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

- ▶ 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

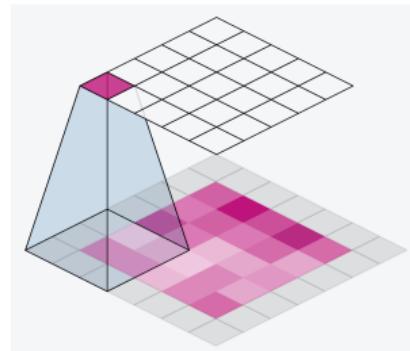
Apport de l'architecture transformer en vision



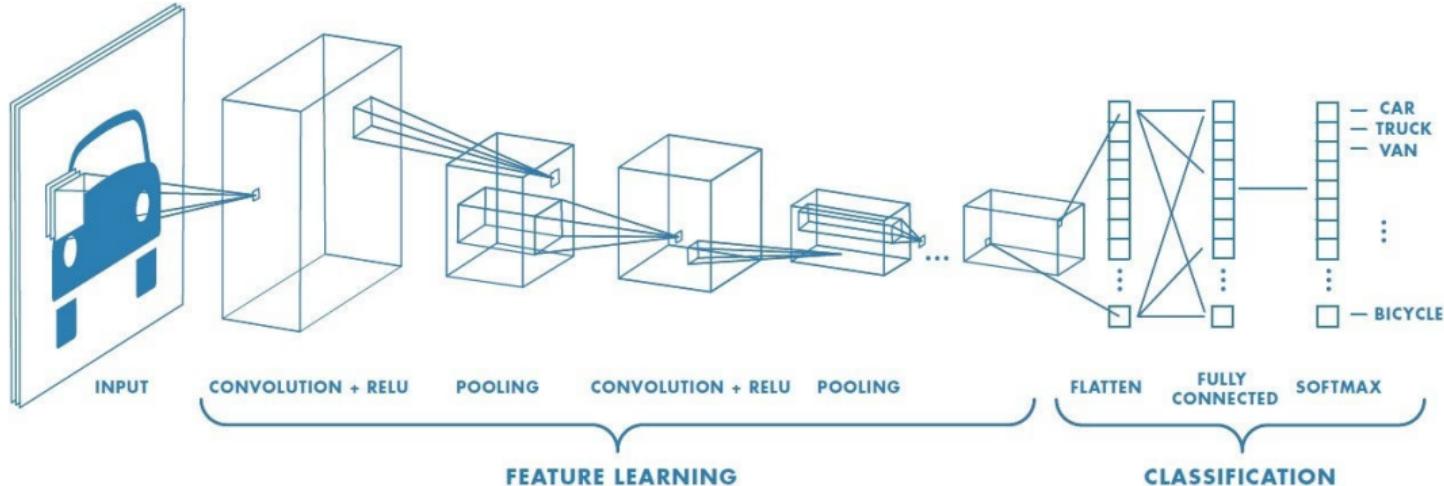
1. Convolution

Filtre d'analyse

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



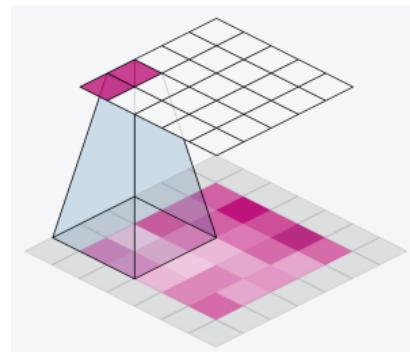
Apport de l'architecture transformer en vision



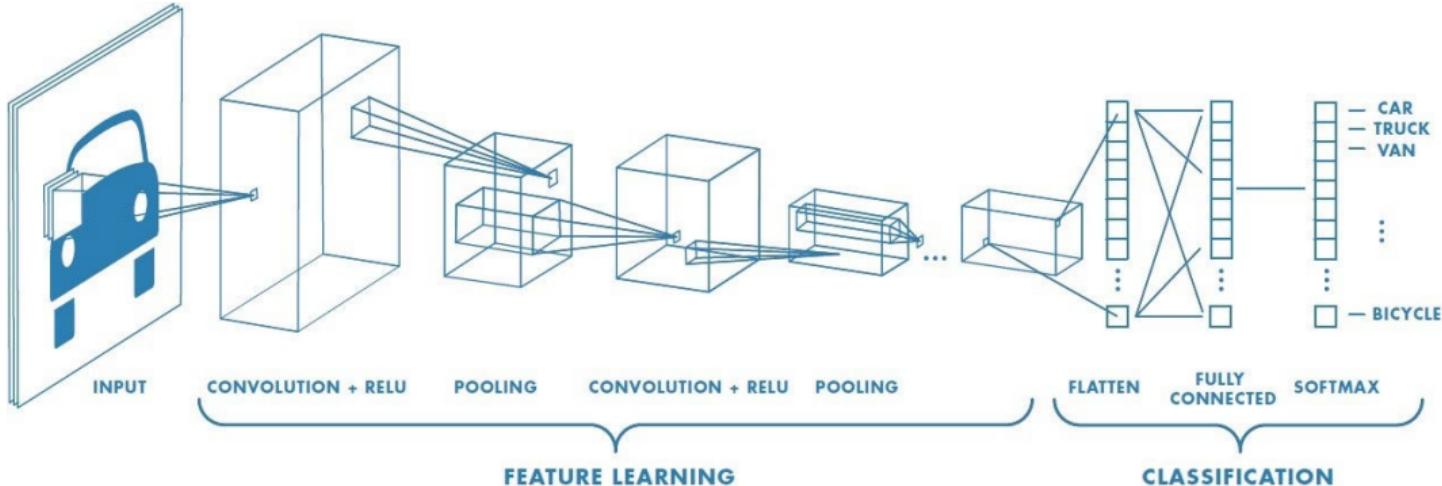
1. Convolution

Filtre d'analyse

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



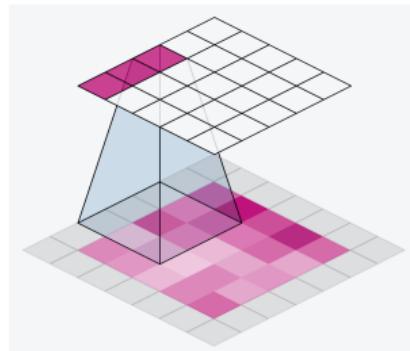
Apport de l'architecture transformer en vision



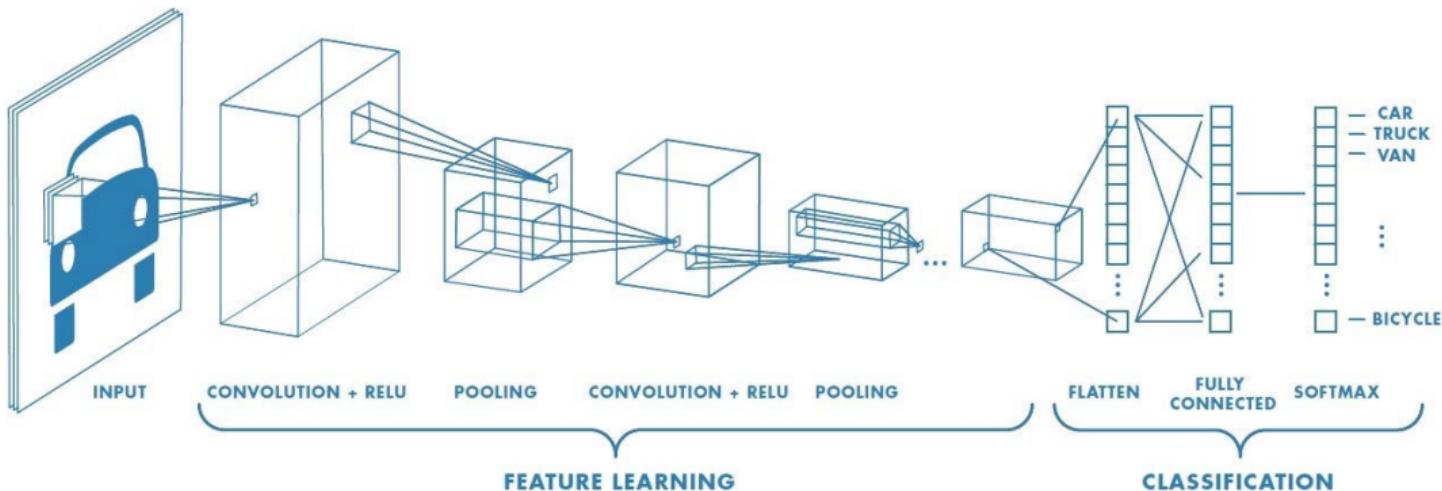
1. Convolution

Filtre d'analyse

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



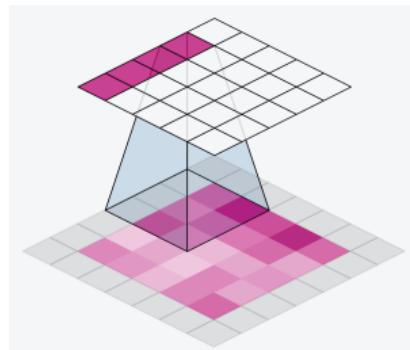
Apport de l'architecture transformer en vision



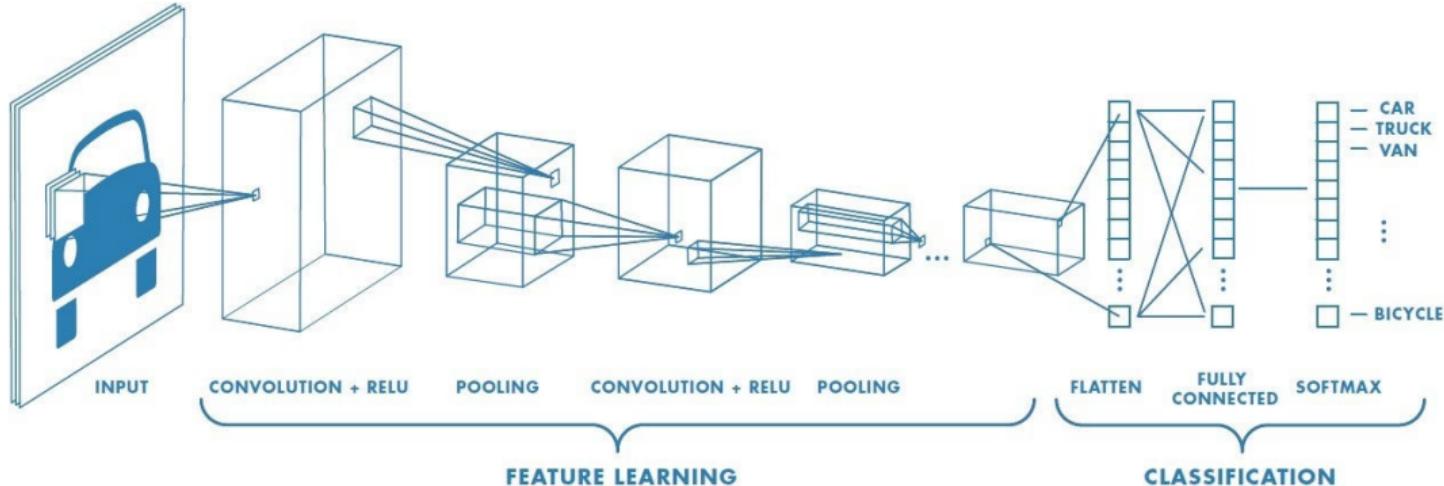
1. Convolution

Filtre d'analyse

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



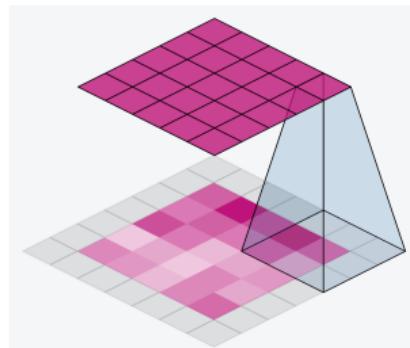
Apport de l'architecture transformer en vision



1. Convolution

Filtre d'analyse

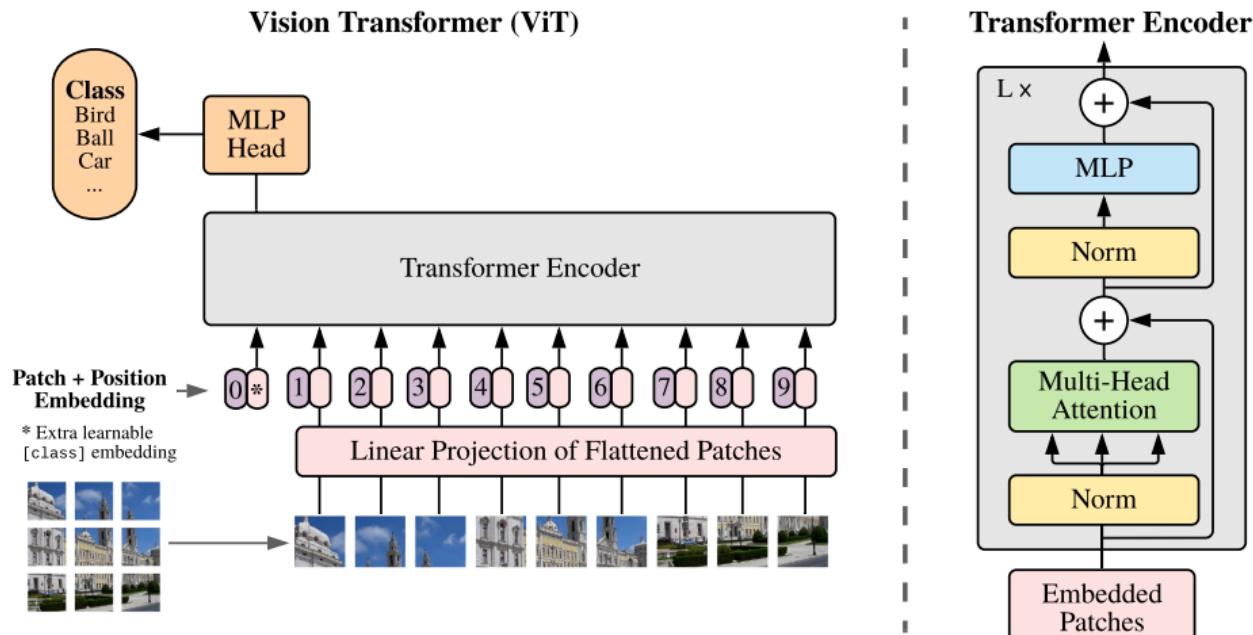
- ▶ Peu de paramètres
- ▶ Apprentissage des motifs à extraire
- ▶ Agrégation progressive des échelles



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

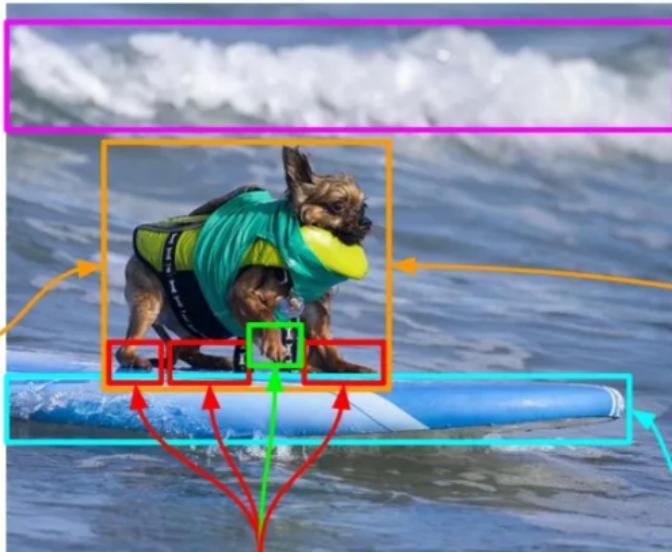
Visual Grounding

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?

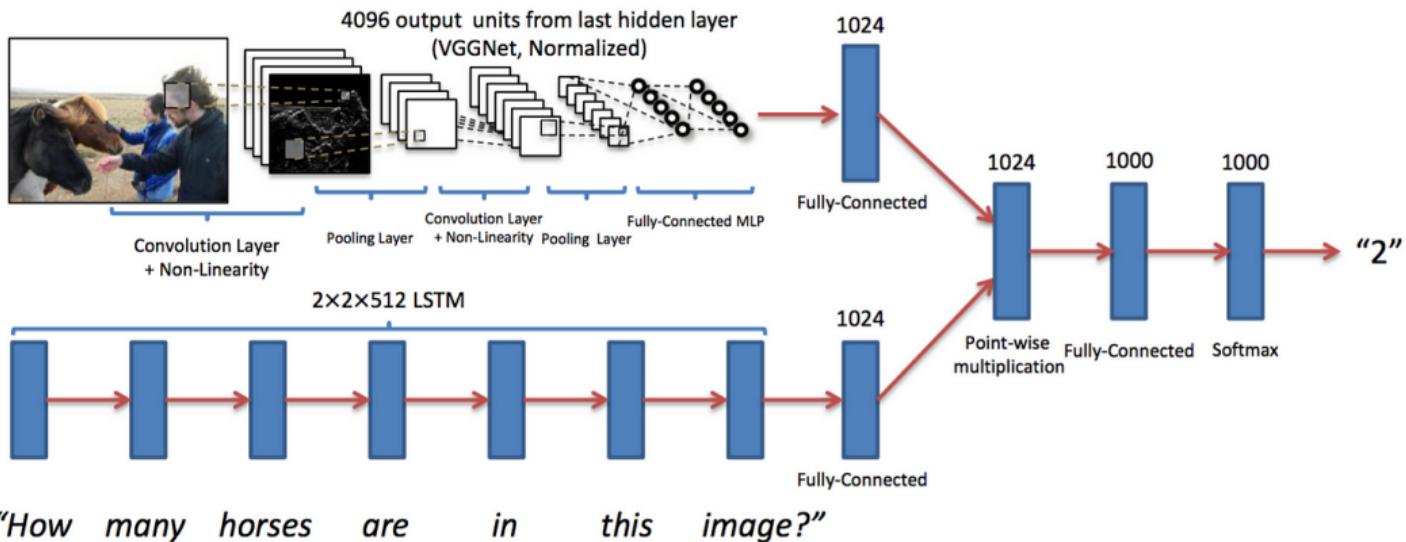
- A) On a surfboard. ✓
- B) On a table.
- C) On a garage.
- D) On a ball.

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



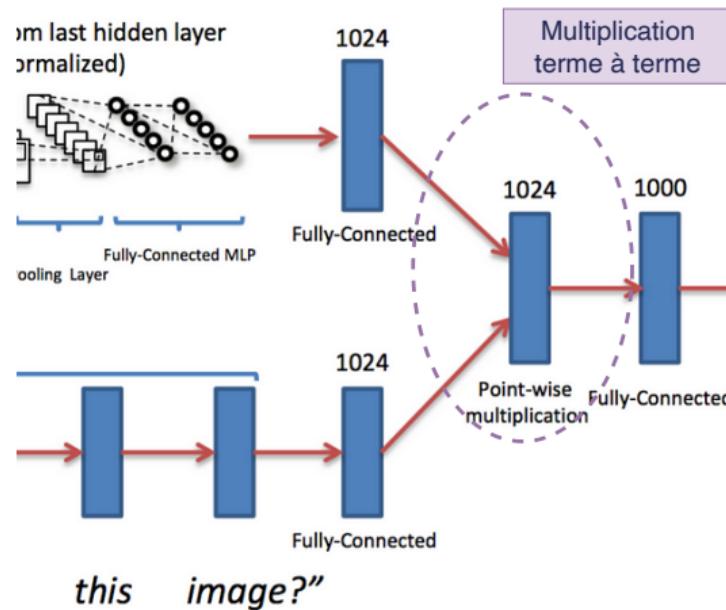
⇒ Rétro-propager l'erreur ⇒ modifier les représentations de mots + analyse de l'image

Compréhension globale / approfondie

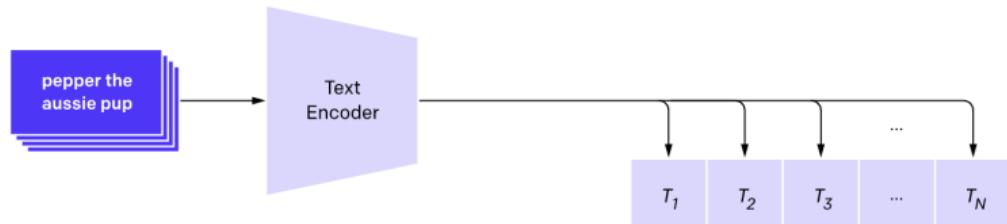
Risque de spécialisation /
Oubli de la connaissance générale

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

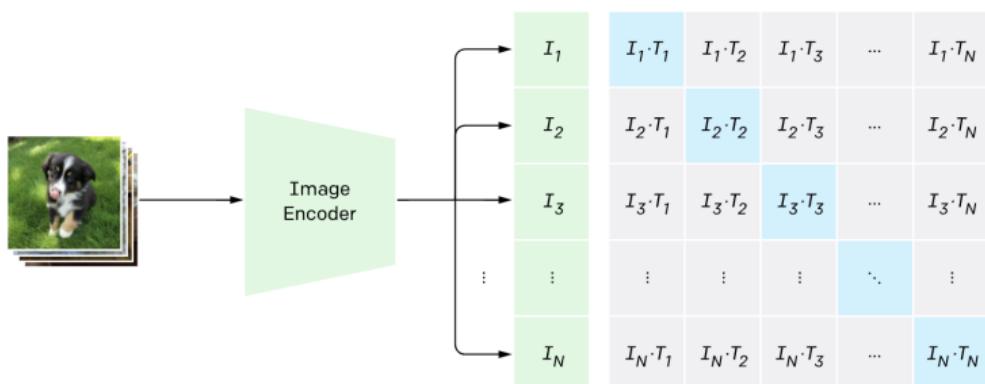


Illustration = encodeur seul (compréhension)

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

Pourquoi ne pas faire une fresque du numérique?

- 1 Mesurer/Quantifier les consommations de matières et les émissions du matériel et des algorithmes
- 2 Construire les dépendances, les causalités dans le processus
- 3 Mesurer l'impact sur l'environnement et le climat
- 4 Proposer des actions concrètes



Une démarche parallèle

- ▶ Labos 1point5
 - ▶ Une réflexion spécifique orientée IA (plus que numérique)
 - ▶ Approche technique compatible avec les enjeux environnementaux
- Section obligatoire dans les articles, reflexion sociétale