

# L'INTELLIGENCE ARTIFICIELLE EN 2023

18 Octobre 2023

Table ronde ATOS-DELL

Vincent Guigue

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

<https://vguigue.github.io>



A. Turing



Dartmouth conference  
Artificial Intelligence

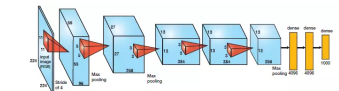


Y LeCun  
G. Hinton

Backpropagation -  
neural networks



Deep learning



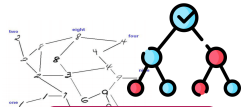
AlphaGo



Winters of AI



Nouveaux  
Langages



Nouveaux  
Paradigmes



Nouveaux  
Objets



AlphaFold

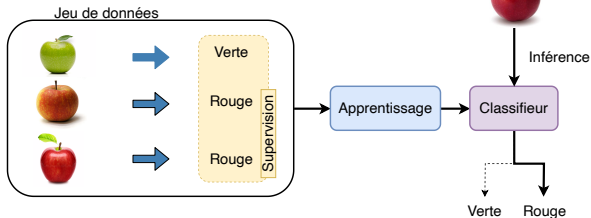


Nouvelles  
applications

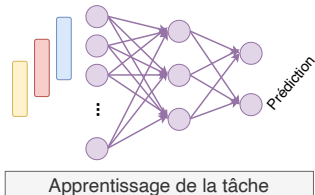
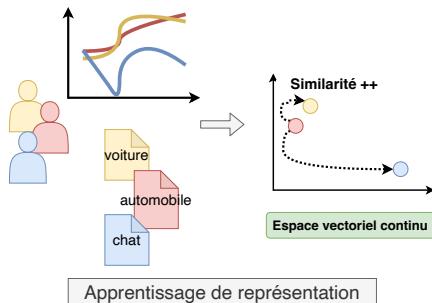


## Du machine learning...

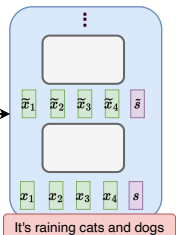
Input (X)	Output (Y)	Application
email	spam? (0/1)	spam filtering
audio	text transcript	speech recognition
English	Chinese	machine translation
ad, user info	click? (0/1)	online advertising
image, radar info	position of other cars	self-driving car
image of phone	defect? (0/1)	visual inspection



## ... Au deep-learning



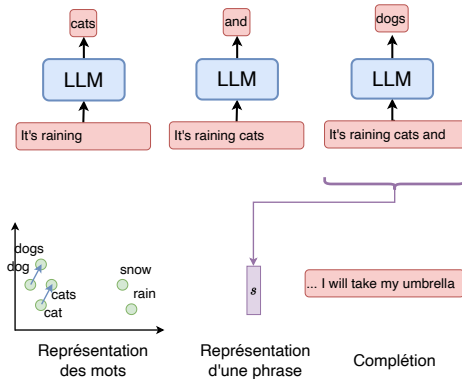
## Architecture Transformer



## Corpus large

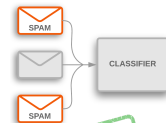


## Pré-entraînement



## Language Model

## Affinage



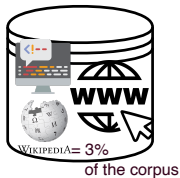
**POSITIVE**  
Great product

**NEGATIVE**  
Worst experience in my life

Sebastian Thrun **PERSON** started at **Google** **ORG** in **2007** **DATE**, few  
nim seriously. "I can tell you very senior CEOs of major **American** **NORP**  
and turn away because I wasn't worth talking to," said **Thrun** **PERSON**, fir  
higher education startup Udacity, in an interview with **Recode** **ORG** **E**

María no abofeteó a la bruja verde  
| | | | X  
Maria did not slap the green witch  
arg1 relation arg2

# chatGPT



Données massives  
+ bien structurées

## Modèle *géant* - LLM

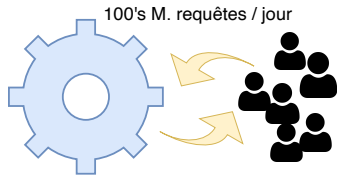
LLM (2017)



80M param.



175G param.

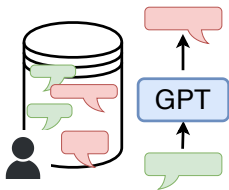


100's M. requêtes / jour

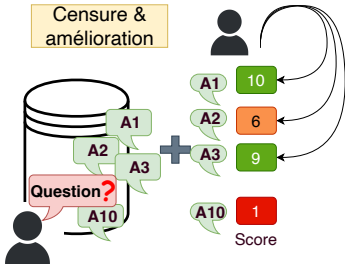
Modèle optimisé /  
quantifié

## Données humaines & interactions

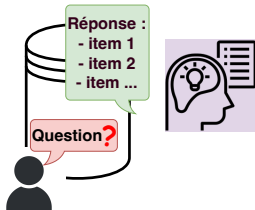
Suivi de dialogue



Censure &  
amélioration



Structuration  
des réponses



Dialogue sur code  
informatique



## Savoir le construire

### Construction du modèle de base

### Raffinement du modèle

### Exploitation du modèle

#### Maitrise des données

- Collecte/équilibrage
- Nettoyage

#### Entraînement

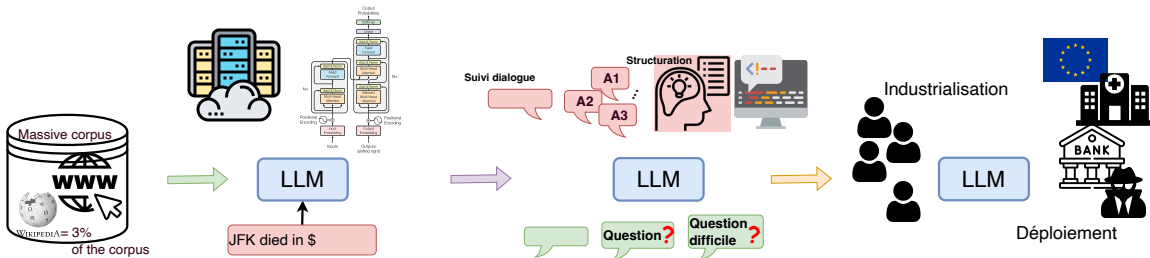
- Puissance machine (milliers de GPU)
- Architecture/recherche ML

#### Maitrise & construction des données

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

#### Optimisation / Limitation du coût

- compétences MLOps
- déploiement local



## Savoir s'en servir

Fiabilité / hallucinations

Prompting

Biais

## Savoir le vendre

Microsoft Tay, Meta Galactica, ...

Modèle imparfait, imprévisible