



2

## Plan de la présentation

- IA et science des textes
- Récentes avancées en TAL et LLMs
- Application à la science des textes
- Conclusion

# IA et science des textes

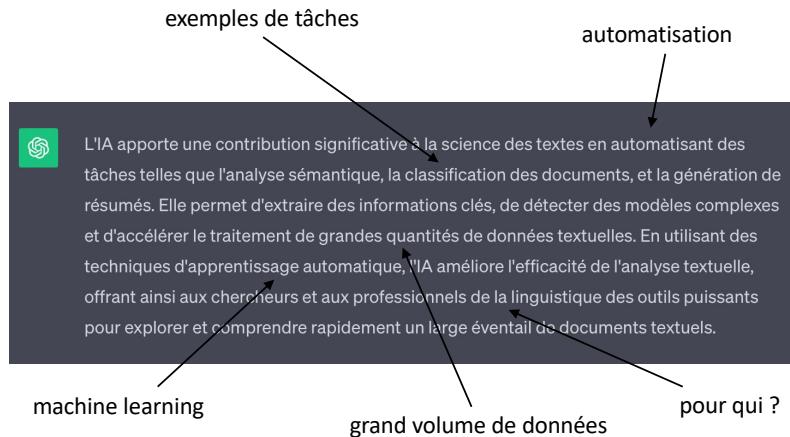
Julien Velcin, laboratoire ERIC, Université Lumière Lyon 2

Dans la suite, je vais écrire un début de phrase et tu vas le compléter comme un modèle de langue simple.

Bien sûr, je suis prêt. Commencez quand vous voulez !

Regenerate 4

4



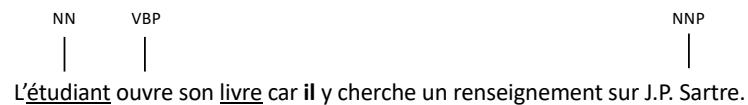
5

## D'Enigma à ChatGPT

- Le Traitement Automatique des Langues (TAL) est un défi dès les premiers travaux en IA. Il permet de :
    - décomposer un texte en ses **constituants**
    - identifier (découvrir) le **sens** des mots et des expressions
    - découvrir des **motifs** (*patterns*) pour classer les textes ou générer du texte
  - Quelques applications :
    - **chercher** de l'information dans les BD et le Web (moteurs de recherche)
    - traduire automatiquement des textes, les **comparer**
    - classer des textes en fonction de leur thématique, des opinions véhiculées...
    - résumer un document, **dialoguer** pour répondre à des questions...

6

## Annotation automatique

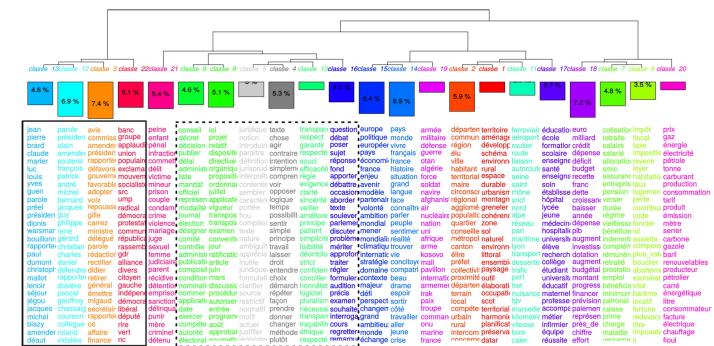


Pour les catégories grammaticales, des logiciels comme TXM utilisent **TreeTagger** qui est basé sur des modèles d'apprentissage automatique, comme les Hidden Markov Models (HMM), mais d'autres modèles ont été proposés : SVM, CRF...

Des problèmes plus complexes, comme l'identification d'entités nommées (NER), nécessitent de recourir à des modèles plus récents, souvent basés sur le *deep learning*.

7

## Segmentation par thématique



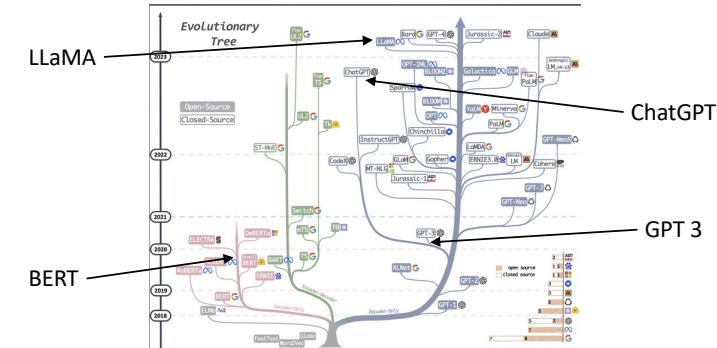
(mondes lexicaux à l'Assemblée Nationale, tirée d'une présentation d'Iramuteq par P. Ratinaux)

8

# Récentes avancées en TAL et LLMs

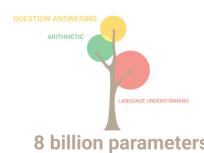
Julien Velcin, laboratoire ERIC, Université Lumière Lyon 2

Succès des grands modèles de langue (LLMs)



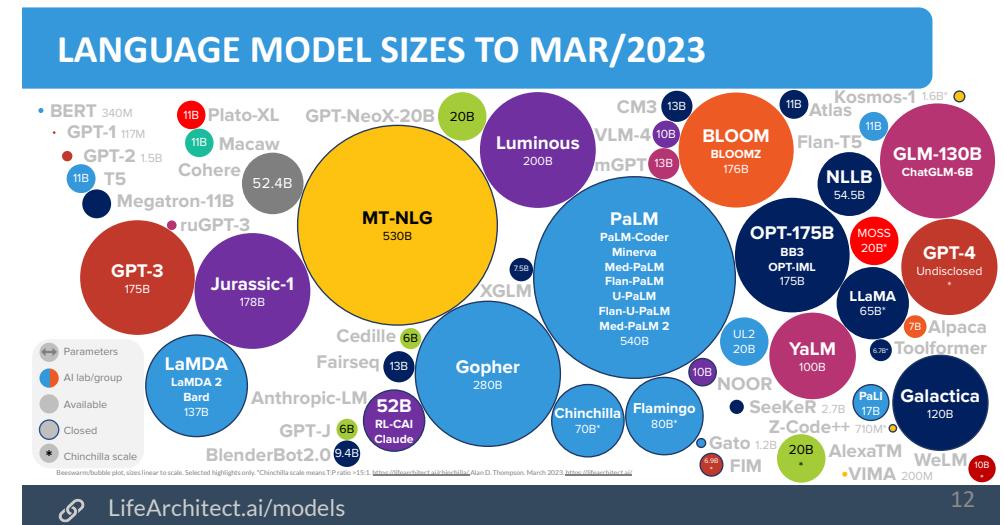
10

Des modèles de plus en plus larges



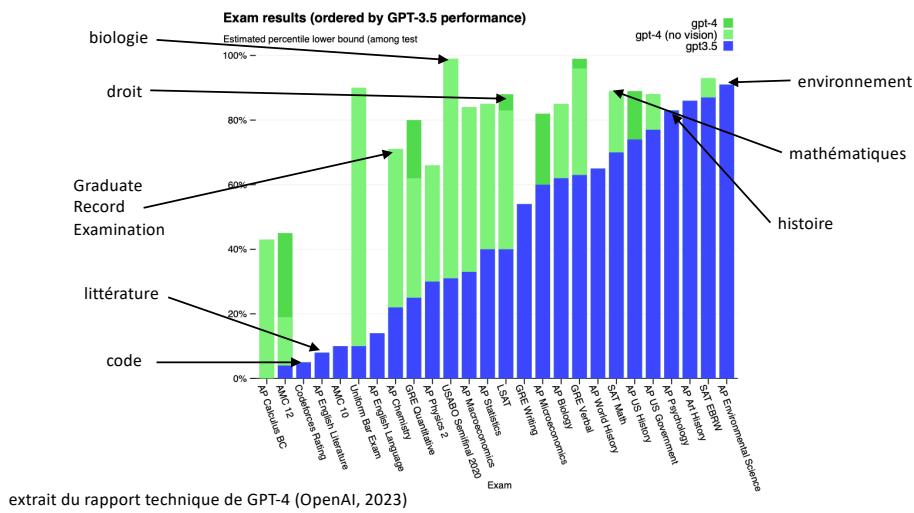
<https://blog.research.google/2022/04/pathways-language-model-palm-scaling-to.html?m=1>

11



12

[LifeArchitect.ai/models](http://LifeArchitect.ai/models)



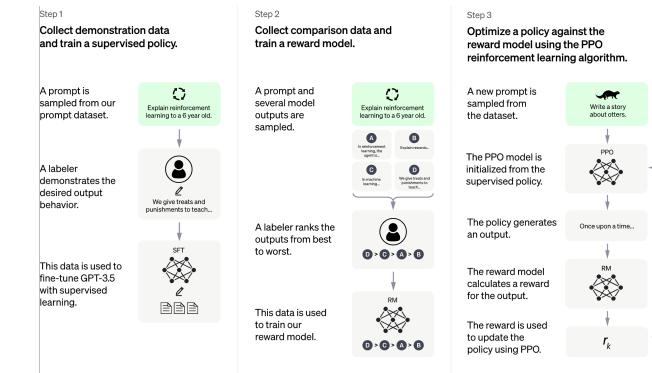
extrait du rapport technique de GPT-4 (OpenAI, 2023)

13

## L'essor des modèles libres

- LLaMA est proposé par Meta en février 2023 (Touvron et al., 2023), disponible pour la communauté, entraînés sur données publiques, nouvelle version avec LLaMA 2 en juillet 2023
  - Mistral AI, fondé par des anciens de Google DeepMind et Meta, propose un modèle libre à 7B de paramètres, Mistral-7B, qui dépasse LLaMA 2 en septembre 2023
  - Ces modèles peuvent être affinés (*fine-tuned*) avec des instructions pour les utiliser comme *chatbots*
  - Disponibles « facilement » sur le hub d’HuggingFace

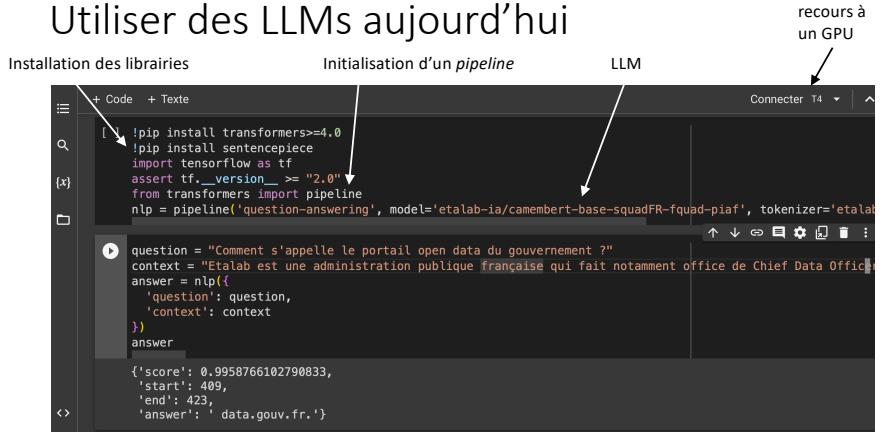
15



<https://openai.com/blog/chatgpt>

14

## Utiliser des LLMs aujourd’hui



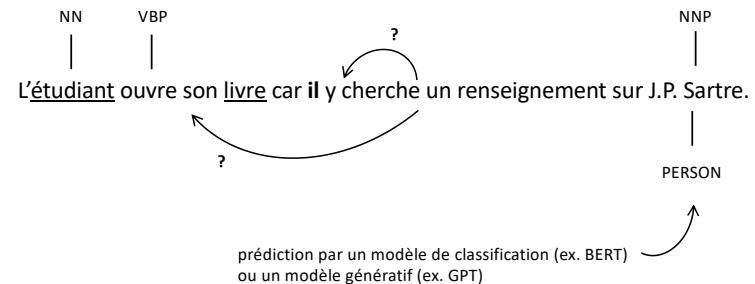
<https://colab.research.google.com>

16

## Enrichissement du texte

# Application à la science des textes

Julien Velcin, laboratoire ERIC, Université Lumière Lyon 2



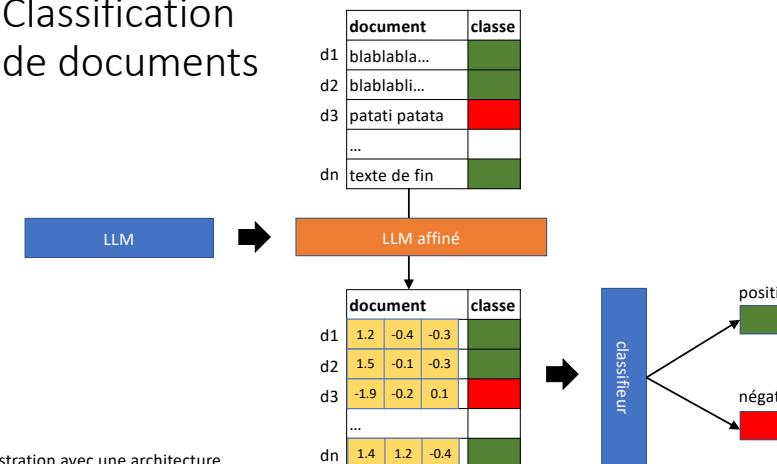
18

## Recherche d'information



19

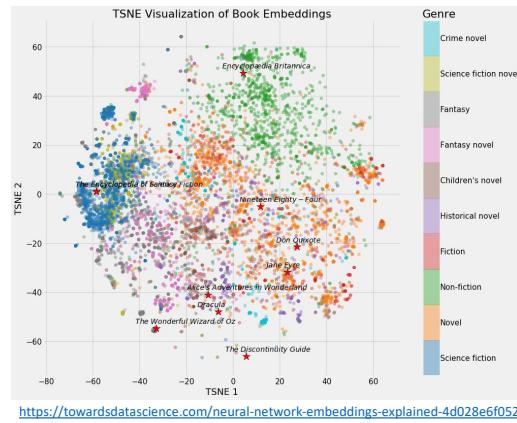
## Classification de documents



20

## Plongement de textes et représentations

- La plupart des applications nécessitent de « plonger » (embed) les textes dans des espaces vectoriels. Ces vecteurs sont une **représentation** qui vise à capturer la sémantique des textes.

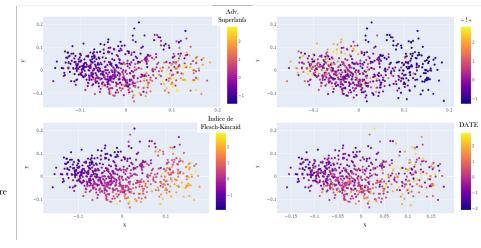
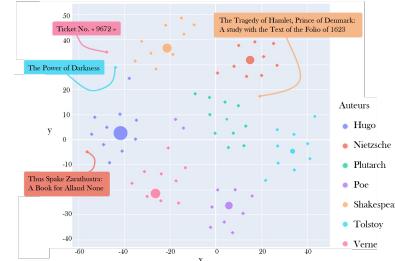


21

<https://towardsdatascience.com/neural-network-embeddings-explained-4d028e6f0526>

## Application à l'analyse du style littéraire

- Projet ANR LIFRANUM , thèse d'E. Terreau  
<https://marge.univ-lyon3.fr/projet-lifranum>



22

## Identification et réduction des biais

- Projet ANR DIKé, thèse d'I. Proskurina sur les biais dans les modèles réduits (ou compressés).
- Pour donner un exemple simple :

```
fill_masker("The editor stopped the driver and asked [MASK] for a ride")
```

[MASK] Bert Predictions with confidence:  
➤ 90% confidence: [MASK] = him  
➤ 10% confidence: [MASK] = her

Language Models can predict HIM instead of HER given the context and [MASK]

## Conclusion

Julien Velcin, laboratoire ERIC, Université Lumière Lyon 2

23

## Quelques conclusions personnelles

- Beaucoup de ressources « prêtées à l'emploi »
- La plupart sont issues des travaux en TAL et en machine learning
- Nécessité de travailler avec des experts pour l'évaluation et l'affinage des modèles (mais attention aux coûts, aux biais...)
- Difficulté de trouver des problématiques de recherche en informatique (ou mathématiques)
- Toujours un chainon manquant entre les développeurs de modèles et d'algorithmes et les chercheurs/utilisateurs en SHS

25

## Références

- Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2019. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. Prépublié en 2018 mais finalement accepté à NAACL 2019 (~80k citations)
- Ouyang, Long, et al. "Training language models to follow instructions with human feedback." NeurIPS 2022 (~2300 citations)
- Irina Proskurina, Guillaume Metzler, Julien Velcin: The Other Side of Compression: Measuring Bias in Pruned Transformers. IDA 2023
- Irina Proskurina, Guillaume Metzler, Julien Velcin: nouveau papier accepté au BabyLM Challenge, ACL 2023
- Enzo Terreau, Antoine Gourru, and Julien Velcin. Writing Style Author Embedding Evaluation. Workshop Evaluation and Comparison of NLP Systems, co-situé avec EMNLP 2021
- Touvron, Hugo, et al. "Llama: Open and efficient foundation language models." arXiv preprint arXiv:2302.13971, 2023, ~1300 citations

26