



UTILISER LES GRANDS MODÈLES DE LANGAGES À BON ESCIENT : ÉTUDE DE CAS



STRUCTURE PRESENTATION

- ▶ Définition de l'IA
- ▶ Différents modèles et principes de base
- ▶ Utilisations
- ▶ Limitations
- ▶ Etude de cas par le Prof. Delphine Courvoisier

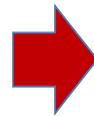
DEFINITION

L'intelligence artificielle est une branche de l'informatique qui vise à créer des machines capables d'effectuer des tâches qui requièrent généralement l'intelligence humaine.

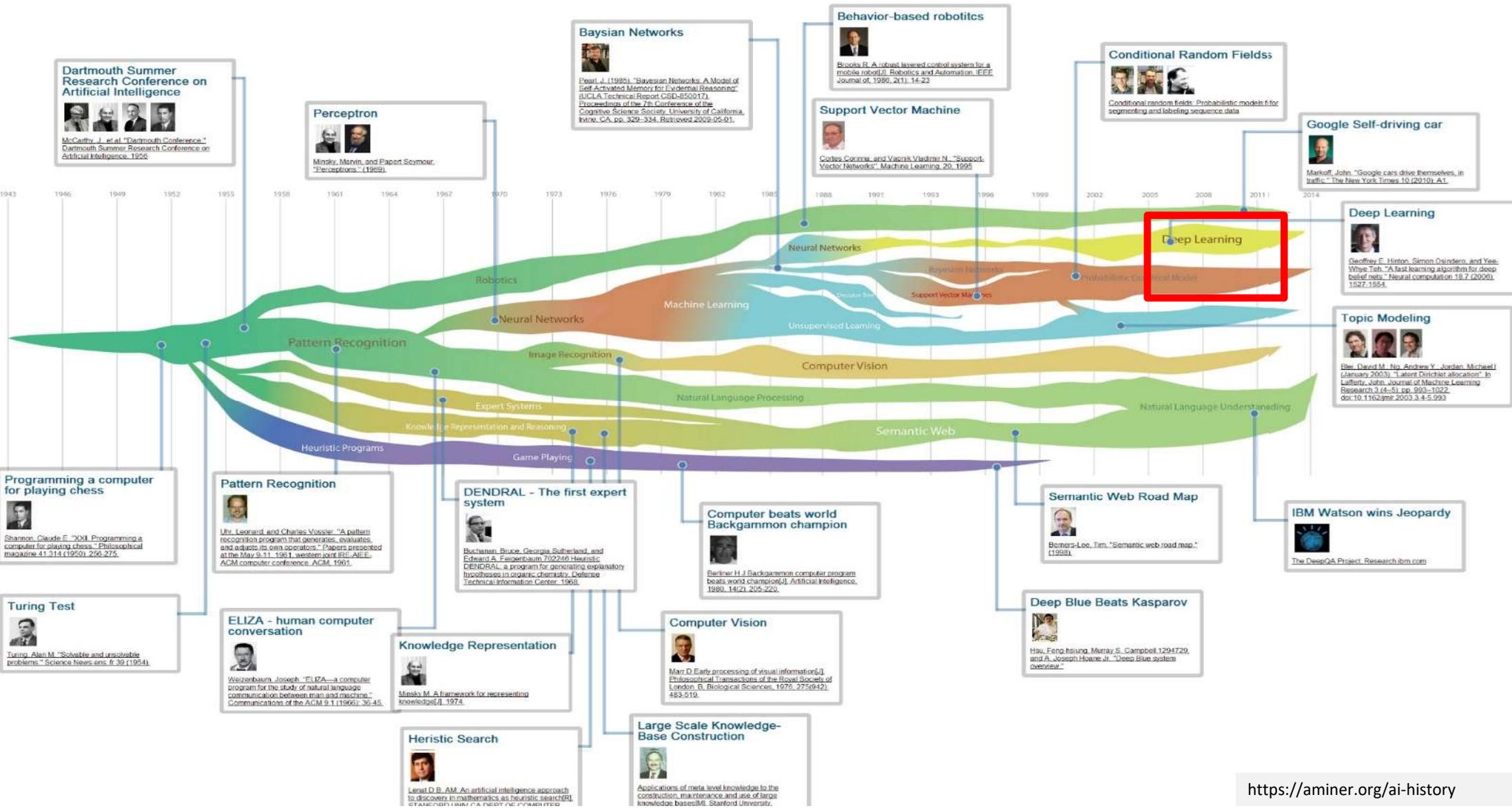


DEFINITION

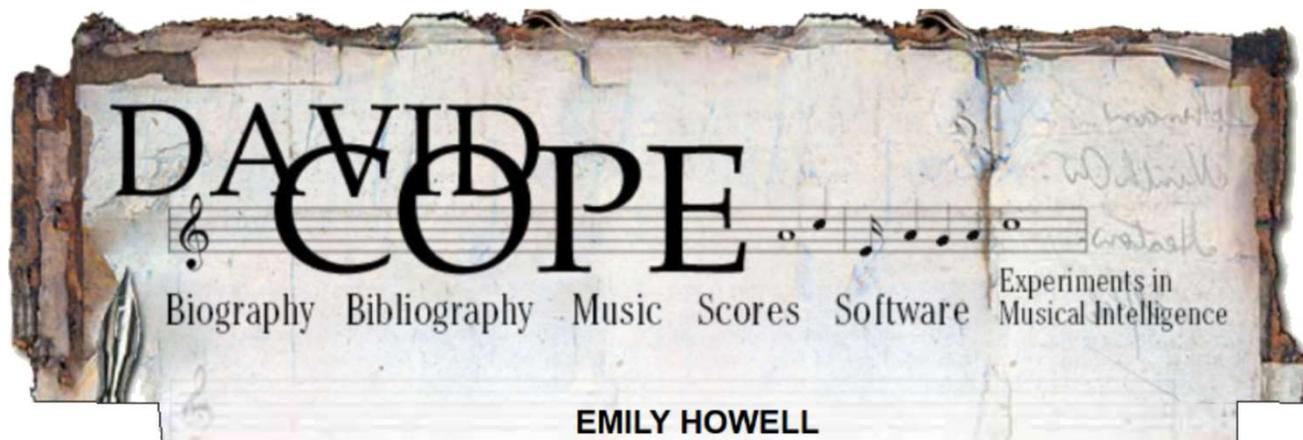
Ces tâches comprennent l'acquisition de l'information, l'apprentissage par l'expérience (apprentissage automatique), le raisonnement, l'auto-correction et la création/décision finale.



HISTORIQUE



RESULTATS ETONNANTS



Bach style chorale



Emmy Vivaldi



Emmy Beethoven beg 2

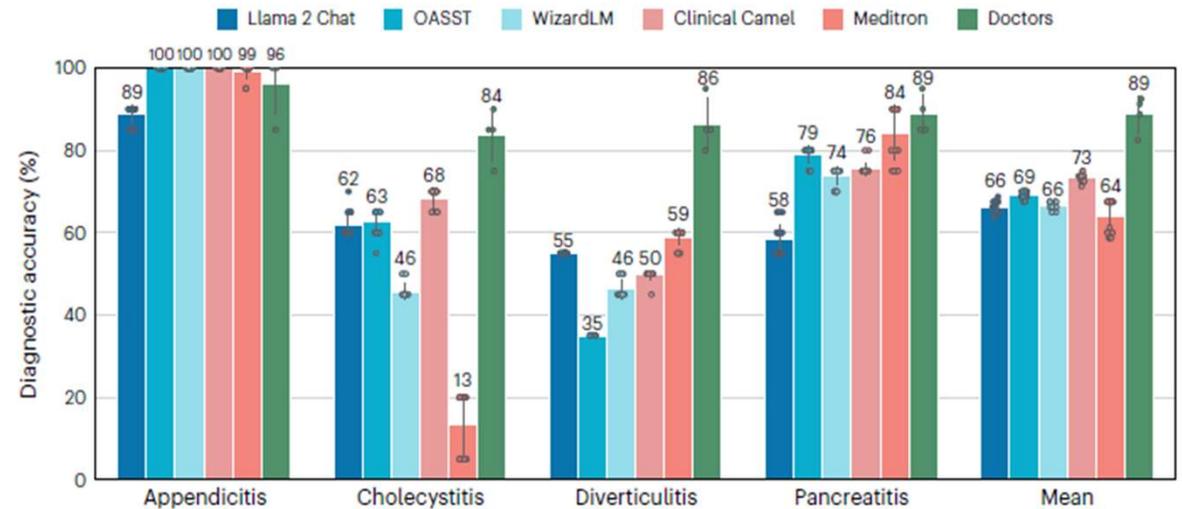


Intelligence artificielle en musique

Expériences

David Cope, The Emily Howell project

ET POURTANT.....



P Hager et al. Nature medicine 2024

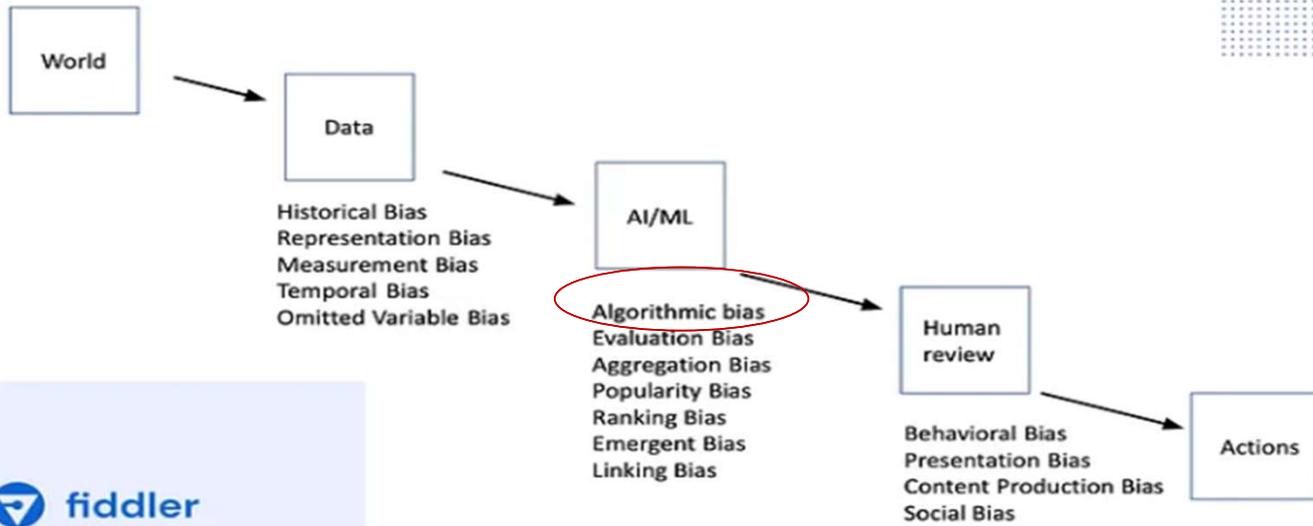
«qu'es que la nyctalopie ?»

ChatGPT

«La nyctalopie, ou vision nocturne, est la capacité à voir dans des conditions de faible luminosité. Les personnes atteintes de nyctalopie ont des difficultés à voir dans l'obscurité ou dans des environnements mal éclairés. Cela peut être causé par plusieurs facteurs, notamment des troubles rétinien, des carences nutritionnelles (comme une carence en vitamine A) ou des maladies oculaires. Si tu as des questions spécifiques à ce sujet, n'hésite pas «.....»

BIAIS DES LLM

Bias



EVALUATION DE LA VALIDITE DES ALGORITHMES:EXEMPLE

n > 20'000



University Hospitals of Geneva
Echocardiography Imaging Unit

Transthoracic Echocardiography: 28.10.2018

Name	John Doe	Sex	Male
D.O.B	06.01.1975	Weight	80 kg
Size	178 cm	Exam quality	Average

Clinical indications: Evaluation of left ventricular function
Clinical information: Atrial fibrillation since 2007

Measure mode 2D/TM

IVS thickness: 0.77cm	LV diam: 6.7 cm	LVM: 240.4 g
	LV indx diam: 3.4	LVMi: 121.4 g/m ²
MAPSE: 0.56 cm	TAPSE: 1.8 cm	

Summary

The left ventricle is moderately dilated with a severely decreased ejection fraction, visually estimated to be 15-20% subject to tachycardic AF. Severe global hypokinesia..



University Hospitals of Geneva
Imaging Department

Cardiac MRI: 09.11.2018

Name	John Doe	Sex	Male
D.O.B	06.01.1975	Weight	80 kg
Size	178 cm	Exam quality	Average

Clinical indications: Reduced LV ejection fraction
Technic: Exam performed on SIEMENS AERA 1.5 Tesla

Description

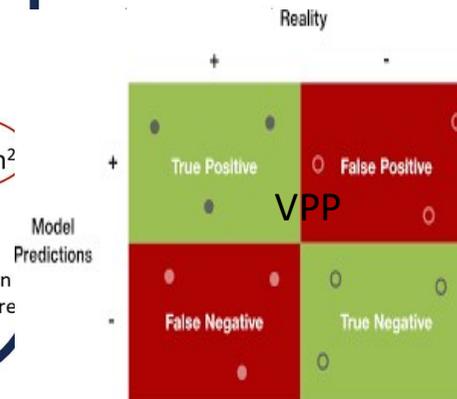
The atria are severely dilated, calculated at 39 cm² for the left atrium and 29 cm² for the right atrium. LVH with a myocardial mass of 144 g/m².

Left Ventricle

EDV	201 ml/m ²
Cardiac output	5.20 l/min
LVMi	144 g/m ²



3'943 ECG annotés
avec les deux
modalités
74.1 % accord



ÉTUDE DE CAS

Delphine Courvoisier
Denis Mongin
Service qualité des soins

Jean-Luc Falcone
Christophe Charpiloz
Faculté des sciences

DIMENSIONS DE LA QUALITÉ

- ▶ Soutien moral
 - ▶ Respect de l'individu
 - ▶ Information au patient
 - ▶ Information à la famille, à l'entourage
 - ▶ Information en chirurgie
 - ▶ Bien-être physique
 - ▶ Sortie
 - ▶ Coordination
 - ▶ Impression générale
- hôtellerie
Autre

ENQUÊTE AUX HUG

- ▶ Quand: 10-14 jours après la sortie
- ▶ Qui administre: auto-rapporté, adulte, enfant
- ▶ Comment: ~20 questions, réparties en 9 dimensions
 - ▶ Envoi électronique via e-mail ou sms
 - ▶ Choix de la langue par le patient: FR, EN, ES, DE
- ▶ Résultats
 - ▶ % d'insatisfaction par dimension
 - ▶ Commentaires libres (~40% de patients laissent un commentaire)

VOLUME

- ▶ ~100'000 patients contactés
 - ▶ ~35'000 répondants
 - ▶ ~11'000 commentaires libres
- ▶ Classification dans la semaine en:
 - ▶ valence (positif, négatif ou neutre)
 - ▶ dimensions de la satisfaction
- ▶ Environ 5 minutes par commentaire
- ▶ → >104 jours-personne

TRAITEMENT MANUEL DES COMMENTAIRES

1. Ouvrir le commentaire dans REDCap
2. Dé-identifier le commentaire (si nécessaire)
3. Séparer le commentaire en sous-commentaires présentant une dimension de la satisfaction et une valence identique
4. Coder chaque sous-commentaire en dimension de la satisfaction et valence
5. Mettre un point d'attention si le commentaire nécessite un traitement supplémentaire
6. S'occuper des problèmes de continuité des soins (p.ex. vérification de l'équipe en charge du patient, prise de contact avec le patient, ...)

TRAITEMENT AUTOMATIQUE DES COMMENTAIRES

1. Télécharger le commentaire dans REDCap
 - ~~2. Dé-identifier le commentaire (si nécessaire)~~
 3. Séparer le commentaire en sous-commentaires présentant une dimension de la satisfaction et une valence identique
 4. Classifier chaque sous-commentaire en dimension de la satisfaction et valence
 5. Annoter le commentaire dans REDCap
-
6. Dé-identification par les membres de l'équipe
 7. Vérification humaine d'un échantillon → contrôle qualité

1. TÉLÉCHARGER LE COMMENTAIRE

- ▶ Accès via API
- ▶ Vu que les données sont dé-identifiées, l'entraînement de l'IA a lieu sur les serveurs UNIGE.
 - ▶ Solution pour le développement mais pas pour la production car beaucoup de commentaires contiennent des données identifiantes

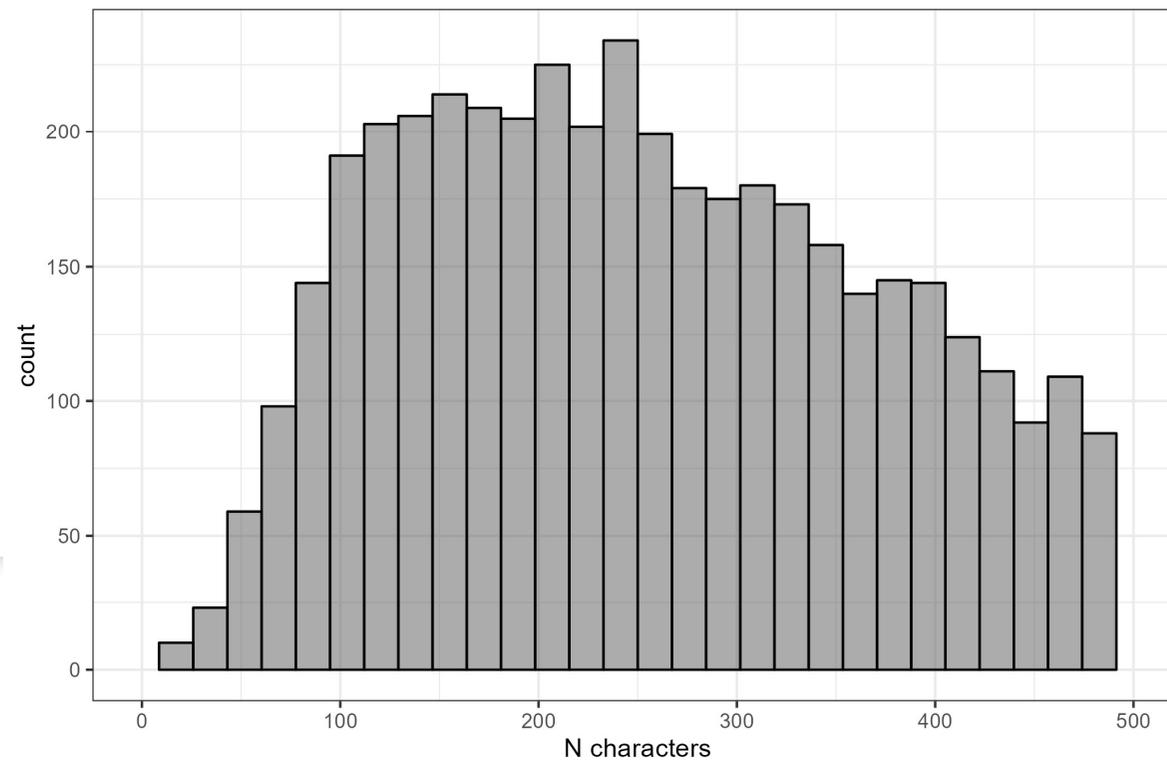
3. SÉPARER LE COMMENTAIRE

- ▶ Choisir une longueur minimale (ex. 100 caractères)
- ▶ Faire la première fenêtre de texte (0 à 100)
- ▶ Utiliser l'IA pour classifier en dimension et en valence
- ▶ Décaler la fenêtre de 1 mot
- ▶ Utiliser l'IA pour classifier en dimension et en valence
- ▶ ...

3. SÉPARER LE COMMENTAIRE

- ▶ Sur la base de l'histogramme des nombres de caractères
- ▶ 13% ont moins de 200
- ▶ <5% ont moins de 100
- ▶ Choix de la fenêtre: 100

comments with more than one dimension
13% of comments with 2 dims have less than 200 char

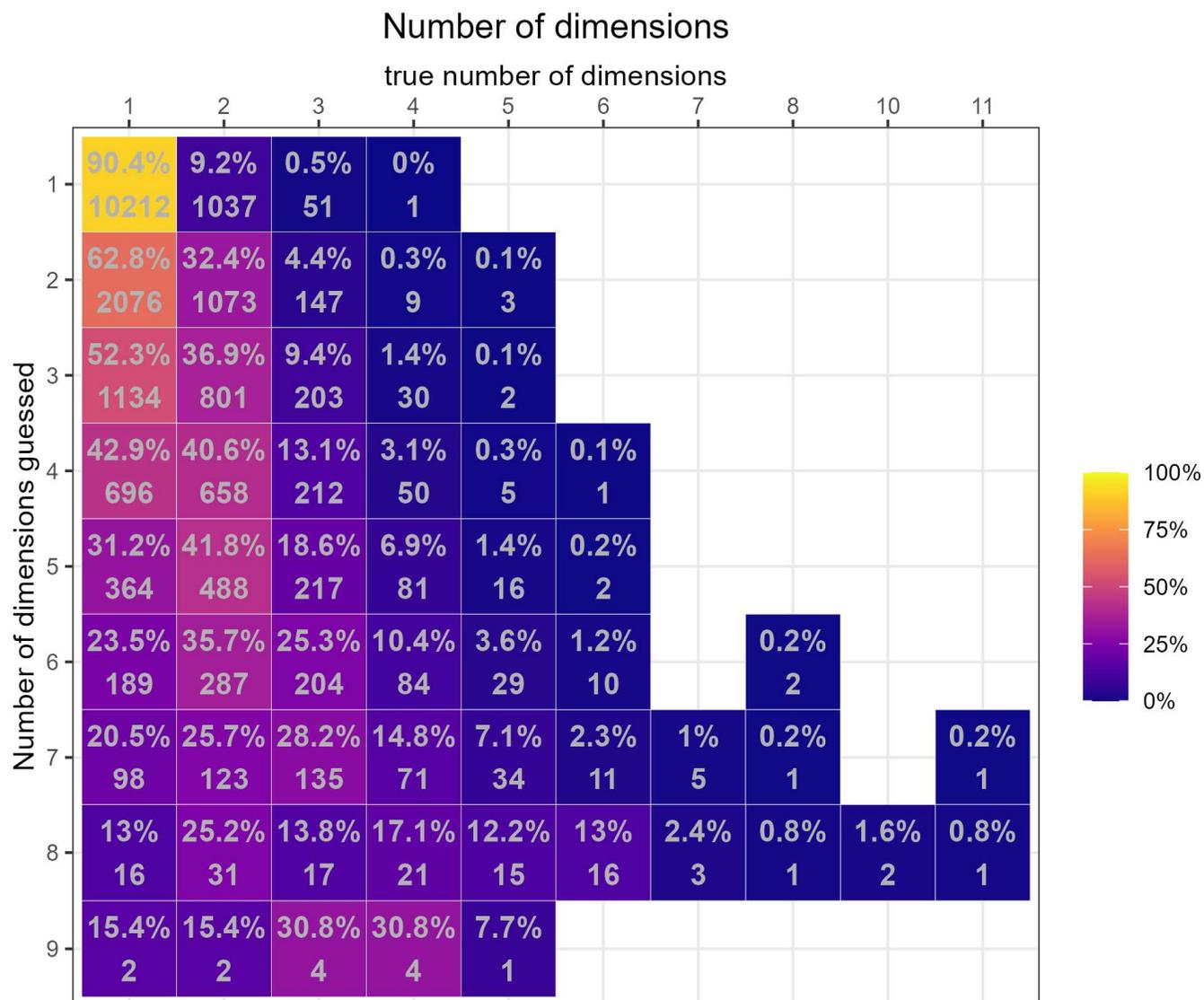


4. CLASSIFICATION DES SOUS-COMMENTAIRES

- ▶ L'IA code le sous-commentaire en
 - ▶ une seule dimension: bien-être, impression générale, hotellerie
 - ▶ une seule valence: positif, négatif, neutre
- ▶ Nous évaluons le nombre de sous-commentaires estimés et la valeur prédictive positive: % de sous-commentaires classés par l'IA
 - ▶ Dans la bonne dimension
 - ▶ Dans la bonne valence
 - ▶ Avec les deux justes

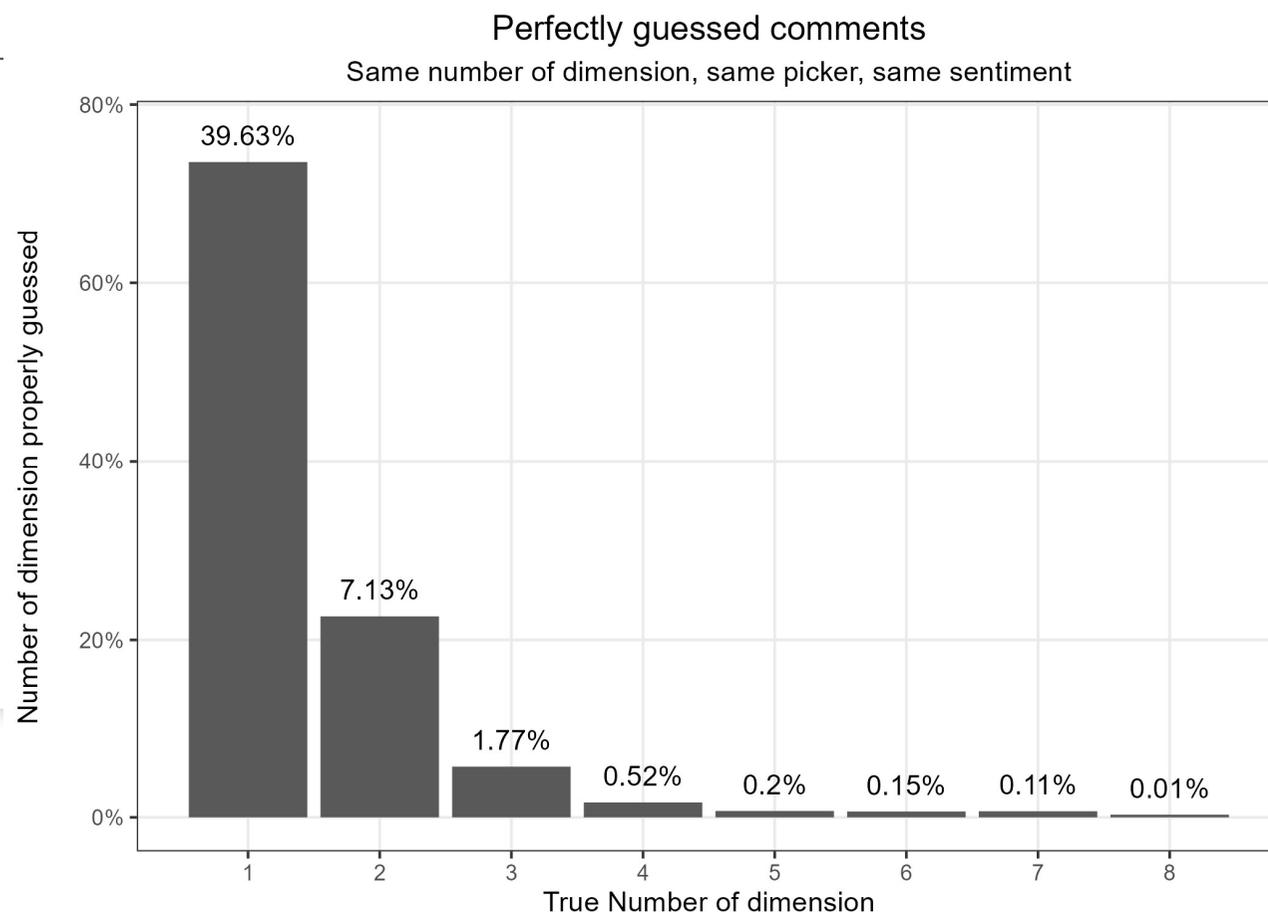
NOMBRE DE SOUS-COMMENTAIRES

- ▶ La plupart des commentaires (10'212) sont courts
- ▶ 90.4% sont classés comme ayant un seul sous-commentaire
- ▶ L'IA a tendance à sur-estimer le nombre de sous-commentaires



COMMENTAIRES AVEC LE BON NOMBRE DE SOUS-COMMENTAIRES ET CHACUN CORRECTEMENT CODÉS

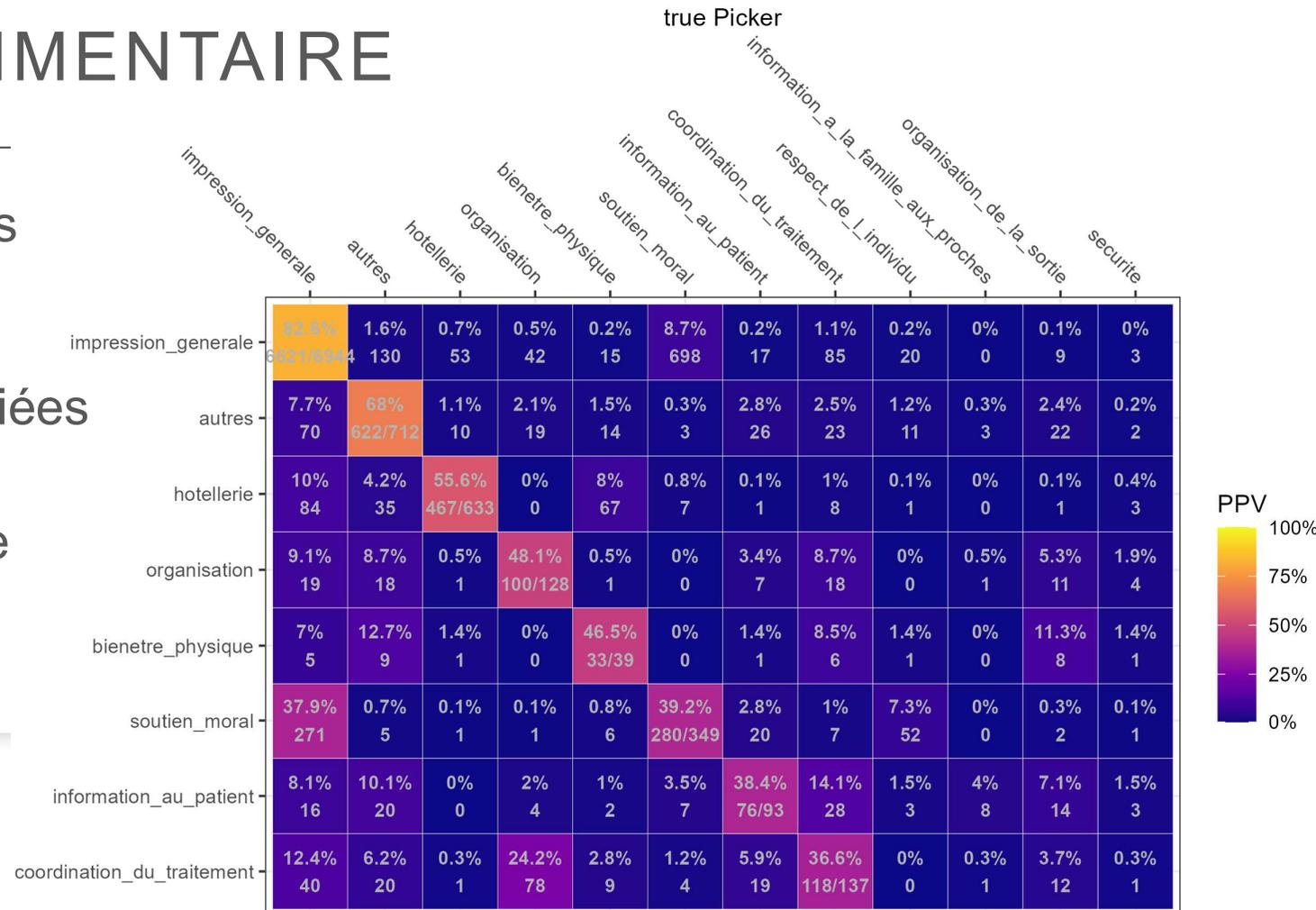
- ▶ Quasi 40% des commentaires ont 1 sous-commentaire
- ▶ Parmi ceux-ci, ~75% sont classifiés justes en dimension et en valence
- ▶ 30% des commentaires totaux (75% de 40%)



UN SOUS-COMMENTAIRE

Parmi les commentaires avec un seul sous-commentaire, quelles dimensions sont classifiées justes (VPP)?

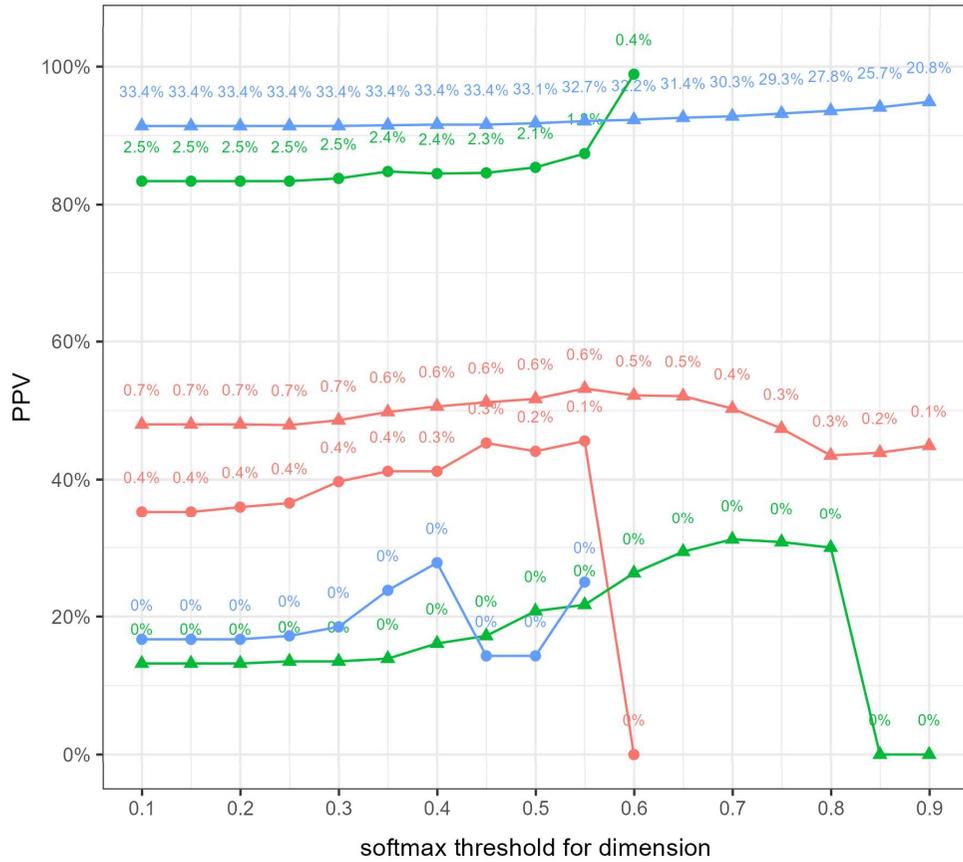
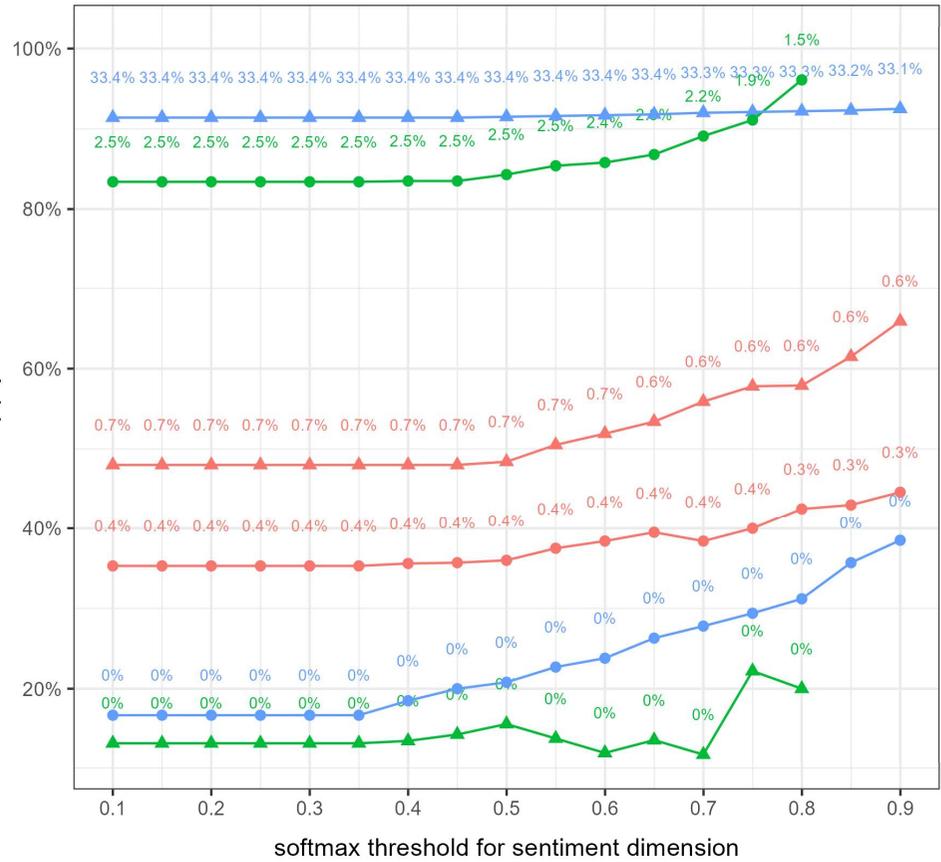
- ▶ Impression générale
- ▶ Autre
- ▶ Hôtellerie?



VARIER LES SEUILS DE DÉCISION DE L'IA

- ▶ L'IA obtient pour chaque sous-commentaire et dimension une valeur de prédiction.
- ▶ Elle choisit ensuite la dimension en prenant la valeur de prédiction (softmax threshold) la plus haute
- ▶ On peut augmenter la VPP en choisissant le seuil nous-mêmes.
- ▶ Si le seuil est trop bas, on ne prend pas la prédiction

Positive predictive value (PPV) and % of total comment properly guessed
Only one dimension, for Impression générale and autre



Picker dimension
● autres
▲ impression_generale

Sentiment
— negative
— neutre
— positive

PREMIÈRE IMPLÉMENTATION

- ▶ Seuil 0.75 pour valence
- ▶ Pas de seuil pour dimension
- ▶ En considérant impression général positive, et autres neutre:
 - ▶ ~90% VPP
 - ▶ ~35% du total des commentaires
- ▶ Nécessite un serveur HUG dédié avec capacité GPU

ÉVALUATION QUALITATIVE

- ▶ Gold standard du codage humain pas excellent
 - ▶ Dans les erreurs sur le codage de l'impression générale, le codage humain est souvent discutable
 - ▶ Définitions des dimensions complexes et se recouvrent en partie
 - ▶ Un seul codeur pour chaque commentaire
 - ▶ Plusieurs codeurs différents, dont des civilistes

ÉTAPES SUIVANTES

- ▶ Mise en place au sein des HUG (serveur dédié)
- ▶ Ajout des résultats aux questions fermées pour améliorer la classification
→ Deuxième implémentation
- ▶ Dé-identification ?

An aerial photograph of Rio de Janeiro, Brazil, showing the bay, mountains, and dense urban development. The word "Questions" is overlaid in white text on a dark blue semi-transparent rectangle in the center of the image.

Questions