

# IA générative : les biais

**Professeur Stéfan Darmoni**

Responsable du Département de Santé Numérique, CHU de Rouen

Laboratoire AIMS

SaNuRN

<http://sanurn.eu>



# Les biais : (1) boîte noire

- Absence d'explicabilité de la majeure partie des outils fondés sur l'IA générative : concept de "boîte noire" (black box)
- A l'opposé des outils fondés sur l'IA symbolique des années 80-90
  - à savoir les systèmes experts, les algorithmes, la formalisation des recommandations sur les bonnes pratiques cliniques
- La solution pour minimiser ce problème : les RAG !

# Les biais : (2) hallucinations

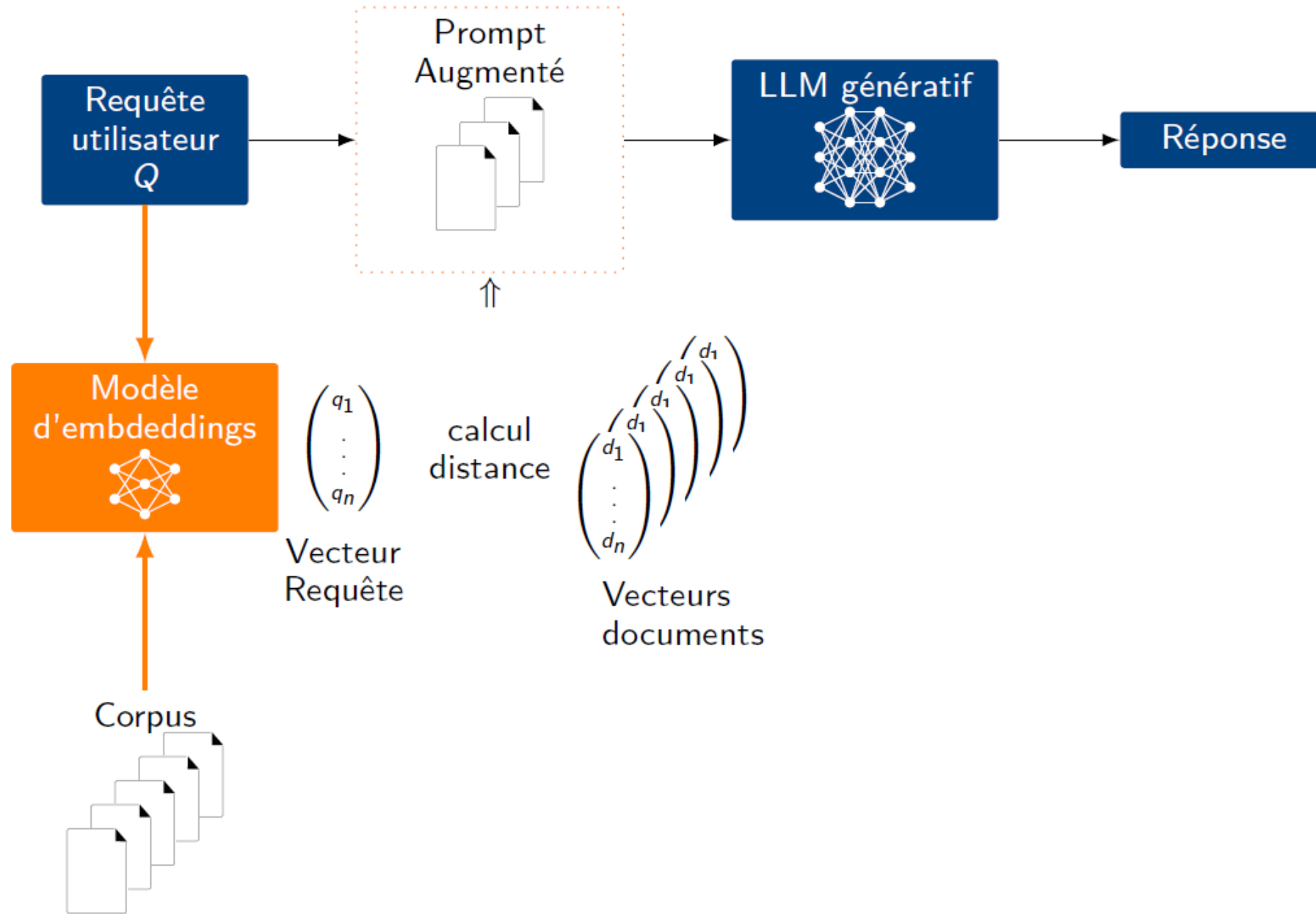
- Les outils d'IA générative sont fondées sur des données massives
- Par construction, il existe a priori des biais statistiques fondées sur ces données
- Avoir conscience de la notion de biais dans les algorithmes
- Une « hallucination » est le fait qu'un LLM n'est pas un enfant bien élevé ; quand il ne sait pas répondre à une question, au lieu de répondre honnêtement : je ne sais pas, il dit n'importe quoi : il « hallucine »
- Les hallucinations peuvent entraîner (fourni par ChatGPT) :
  - **Mauvaises décisions cliniques** si un professionnel fait confiance à une information erronée
  - **Propagation de fausses croyances médicales** chez les patients
  - **Erreur de diagnostic ou de traitement** si utilisé sans supervision
  - **Perte de confiance** dans les outils d'IA médicale
- **Exemples de conséquences :**
  - Prescrire un traitement inefficace ou dangereux
  - Ignorer une contre-indication réelle
  - Informer un patient avec de fausses données (ex. taux de survie, options de traitement)
- La solution pour minimiser ce problème : les RAG !

# Définition d'un RAG

- Un système **R**etrieval-**A**ugmented **G**eneration (RAG) est une architecture conçue pour fournir à un **L**arge **L**anguage **M**odel (LLM) génératif des informations issues d'un **corpus** externe au moment de son interrogation.
- Pour cela, il combine un **moteur de recherche d'information (retriever)** — *souvent fondé sur des représentations vectorielles sémantiques* — avec un **modèle génératif**.
- Un pipeline RAG intègre généralement les étapes suivantes :
  - 1 **Recherche** de documents pertinents dans une base de connaissances externe, en fonction de la requête ;
  - 2 **Augmentation** du prompt avec les informations ainsi récupérées ;
  - 3 **Génération** de la réponse à partir de ce prompt enrichi, via l'inférence du LLM.



# Pipeline RAG typique



# CISMeF



- Catalogue et Index des Sites Médicaux de langue Française
- Création : février 1995
- URL : [www.cismef.org](http://www.cismef.org)
- Référence la littérature grise en accès libre et gratuit
- Plusieurs milliers d'utilisateurs quotidiens
- 130 000 ressources



# LiSSa

- Lis ça !
- Référence les articles
  - de revues scientifiques,
  - dans le domaine de la Santé,
  - en français,
  - gratuits et payants.
- Le « petit PubMed francophone »
- URL : [www.lissa.fr](http://www.lissa.fr)
- Environ 1000 utilisateurs par jour ouvré
- 1,2 M d'articles recensés et indexés



# HeTOP

- Health Terminology/Ontology Portal since 2007
- DéSaN, Rouen University Hospital
- URL : [www.hetop.eu](http://www.hetop.eu)
- 109 KOS (Knowledge Organization System), 55 languages
- Partially free (46 KOS); some others after authentication (ICPC2/3)
- Over 125,000 hits from 1,700 unique machines per working day
- More than 8,900 subscribers
- In 2020, creation of the HUI (HeTOP Unique Identifier) for French concepts

# RAG CISMeF, LiSSa, HeTOP

- Dans le cadre du projet PARTAGES, financé par la BPI, premier contractant : Health Data Hub
  - Objectif : créer un ou N LLM en santé et en français, avec 7 cas d'usage différents dont désidentification et optimisation du codage PMSI...
- Développement de trois RAG sur chacun des trois outils : CISMeF, LiSSa, HeTOP



# RAG LiSSa

## Pourquoi le vaccin contre la rougeole cause-t-il l'autisme ?

Réponse : D'après les documents fournis, il n'existe aucune preuve scientifique crédible reliant le vaccin contre la rougeole à l'autisme. En fait, les documents indiquent que cette idée a été largement réfutée par la communauté scientifique.

1. **\*\*Document NLM\_PF\_24162066 (2010)\*\*** : Ce document explique que les interrogations concernant une association entre la vaccination ROR (rougeole-oreillons-rubéole) et les troubles autistiques ont commencé en 1998 suite à la publication du Dr Wakefield. Cependant, cette publication a été largement critiquée et rétractée en raison de nombreuses irrégularités. Les études épidémiologiques n'ont pas trouvé de lien entre la vaccination et l'autisme, et la plupart des auteurs de l'étude originale ont rétracté leurs conclusions.

2. **\*\*Document NLM\_21939085 (2011)\*\*** : Ce document traite de l'histoire de l'autisme et du vaccin ROR, soulignant que la controverse a été un scandale médical qui a causé une baisse de la couverture vaccinale, mais qui a été réfutée par des études scientifiques ultérieures. Invention du contenu du document  
<https://www.lissa.fr/fr/rep/articles/21939085>

3. **\*\*Document NLM\_EL\_S0992594519300765 (2019)\*\*** : Ce document mentionne que les mythes concernant le lien entre le vaccin ROR et l'autisme persistent, mais qu'ils ont été réfutés par la recherche scientifique.

En résumé, la théorie selon laquelle le vaccin contre la rougeole cause l'autisme a été largement discréditée et rejetée par la communauté scientifique. Les études épidémiologiques et les avis d'organismes de santé publique, comme le General Medical Council britannique, ont confirmé qu'il n'existe pas de lien causal entre la vaccination ROR et l'autisme.



## Pourquoi les risques de cancer de la prostate sont-ils majorés chez la personne enceinte ?

- Réponse : Sur la base des documents fournis, il n'y a pas d'informations directes ou de contexte expliquant pourquoi les risques de cancer de la prostate seraient majorés chez une personne enceinte. Les documents se concentrent principalement sur le traitement, le dépistage et les aspects cliniques du cancer de la prostate, sans mentionner de lien avec la grossesse ou les conditions spécifiques des femmes enceintes.
- Si vous avez besoin d'informations sur ce sujet, il serait préférable de consulter des sources médicales spécialisées ou des études de recherche qui abordent spécifiquement les interactions entre la grossesse et les risques de cancer de la prostate. **Absurde**

Pourquoi les risques de cancer de la prostate sont-ils majorés chez la personne enceinte ?



La question semble contenir une confusion, car les personnes enceintes ne peuvent pas développer de cancer de la prostate. La prostate est une glande spécifique aux personnes assignées homme à la naissance, et elle n'est pas présente chez les personnes assignées femme à la naissance, qui sont celles qui peuvent tomber enceintes.

Si vous avez des questions sur les risques de cancer chez les personnes enceintes ou sur un autre sujet médical, n'hésitez pas à demander des précisions ou à reformuler votre question.



# RAG LiSSa

## Pourquoi la consommation de pamplemousse est-elle déconseillée avec le paracétamol ?

Réponse : La consommation de pamplemousse est déconseillée avec le paracétamol en raison de l'interaction *potentielle* entre le jus de pamplemousse et le métabolisme des médicaments. Le jus de pamplemousse contient des composés qui inhibent le cytochrome P450 3A4, une enzyme clé dans le métabolisme présystémique de nombreux médicaments. Cette inhibition peut entraîner une augmentation des concentrations plasmatiques de médicaments qui dépendent de cette voie métabolique, augmentant ainsi le risque de surdosage et d'effets indésirables.

Bien que le document ne mentionne pas spécifiquement une interaction entre le jus de pamplemousse et le paracétamol, il est important de noter que le paracétamol est métabolisé par le foie, et toute interaction qui affecte le métabolisme hépatique pourrait potentiellement influencer sa sécurité et son efficacité. Par conséquent, il est généralement conseillé d'éviter la consommation de jus de pamplemousse avec des médicaments qui ont un métabolisme complexe ou une marge thérapeutique étroite, comme le paracétamol, pour prévenir les effets indésirables.

Références :

- Document NLM\_12611197 : "Jus de pamplemousse et médicaments: une association à surveiller?"
- Document NLM\_14758791 : "Nouvelles perspectives avec le paracétamol."
- Document NLM\_PF\_26331856 : "Pamplemousses: des interactions avec les médicaments.«

Commentaire Pr. Staccini (Nice) : Cette interaction n'est pas documentée de façon significative dans la littérature scientifique. Le paracétamol n'est pas principalement métabolisé par le CYP3A4 (mais surtout par conjugaison, via UGT et SULT). L'interaction évoquée est donc **infondée** et peut induire en erreur.



# Quelques informations complémentaires

- UE d'IA appliquée à la santé
- Ouverture en septembre 2025 dans le cadre du DIU Sciences de données (Rouen & Nice)



# Notre questionnement – Nos réponses

- Boîte noire & hallucinations IA générative – Mettre en place des RAG pour les minimiser
- Travailler avec des enseignants-chercheurs en informatique
  - Pr. Stéphane Canu, INSA Rouen & Min. Ens. Sup
  - Pr. Richard Dufour, Nantes
  - Pr. Didier Schwab, LIG, Grenoble

pour améliorer l'explicabilité des outils d'IA générative et minimiser les hallucinations

- Sur le plan de la recherche, est-il possible de construire des outils d'IA hybride, symbolique et numérique
  - La réponse : construire dans les mois qui viennent une preuve de concept : RAG Graph

# Messages Clés – *Take-Home Messages*

- Réduire les biais de l'IA générative
- Enseigner l'IA à tous les étudiants de santé et paramédicaux
- Sur le plan R&D, Tester une approche hybride (IA symbolique & IA générative) : RAG graph

# Associer IA symbolique & IA générative

- Encore du domaine de la recherche active
- Au sein du DéSaN, toujours viser une recherche appliquée
- Dans cet exemple, nous visons un démonstrateur (preuve de concept), sans nécessité absolue d'atteindre certains points de criticité (comme par exemple, un temps de réponse  $< 2s$ )
- Coopération entre le DéSaN et la startup Jarod pour développer un RAG-Graph permettant d'exploiter nativement les triplets RDF du réseau sémantique (Web sémantique) du serveur terminologique HeTOP

