

Exploration du RAG pour la génération de réponses à des questions en contexte éducatif : étude sur les données SCIQ

Sarah Nouali

Doctorante en Informatique
R2I LIS UMR 7020 CNRS

Ismail Badache

Maître de conférences
R2I LIS UMR 7020 CNRS

Patrice Bellot

Professeur des universités
R2I LIS UMR 7020 CNRS



IA-ÉDU

Enjeux, Défis et Perspectives

30 juin 2025



Plan

L'Intelligence Artificielle et l'Education ————— 01

Qu'est-ce que le Retrieval Augmented Generation (RAG) ? ————— 02

Notre stratégie RAG pour la génération de réponses à des questions ————— 03

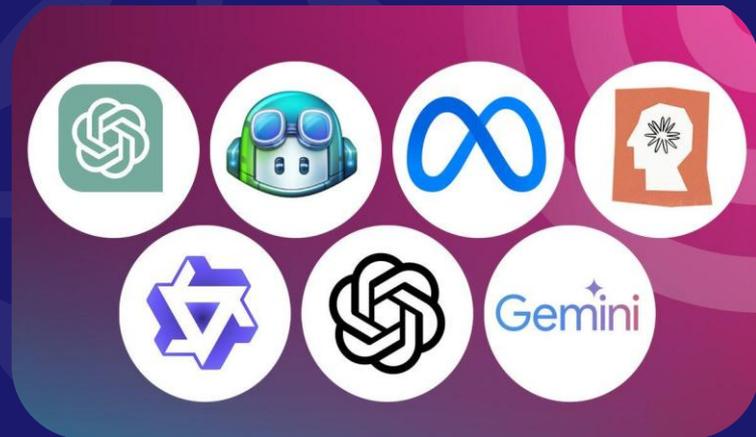
Conclusion ————— 04



1. L'Intelligence Artificielle et l'Education

a. IA Générative :

- Les grands modèles de langues (LLM) : capables de comprendre et générer du texte en langage naturel
- AI Act¹ a classé les systèmes d'IA dans l'éducation à haut risque:
 - x Risques d'information incorrecte, absurde ou obsolète (Hallucination)
 - x Risques de discrimination ou de biais

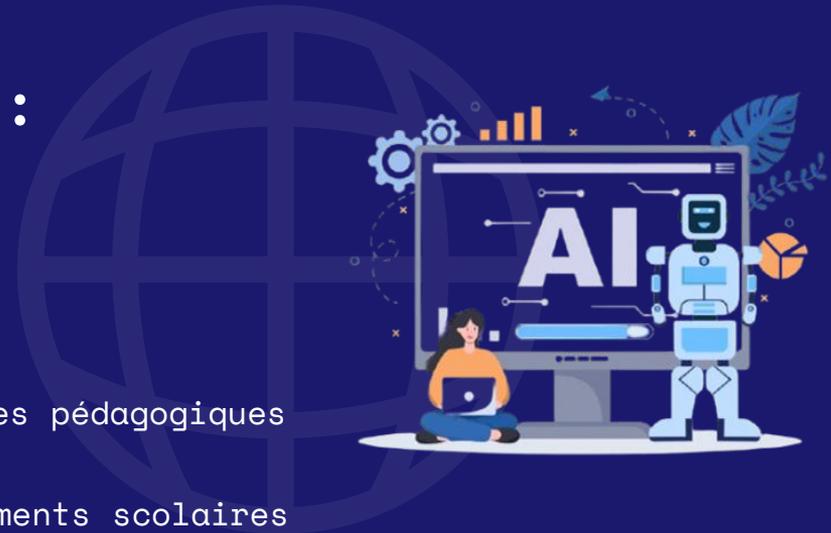


1: <https://www.entreprises.gouv.fr/decryptages-de-nos-experts/le-reglement-europeen-sur-lintelligence-artificielle-publics-concernes>

1. L'Intelligence Artificielle et l'Education

b. Quelques usages éducatifs :

- Aider les élèves
 - ◆ Systèmes de tutorat intelligent
 - ◆ Apprentissage ludique
- Apporter un soutien aux enseignants
 - ◆ Curation intelligente de ressources pédagogiques
 - ◆ Détection de plagiat
- Assister les institutions et établissements scolaires
 - ◆ Gestion des admissions
 - ◆ Planification des cours et des emplois du temps



2. Qu'est-ce que le Retrieval-Augmented Generation (RAG) ?

a. Définition :

Le RAG représente une avancée majeure dans le domaine de l'IA générative en combinant :

Les capacités génératives des grands modèles de langage (LLMs)

Un moteur de recherche documentaire

Générer une réponse basée sur des connaissances contextuelles fiables

Contrairement aux LLMs traditionnels, le RAG s'appuie sur des sources vérifiables et actualisées, garantissant pertinence et fiabilité des réponses.



2. Qu'est-ce que le Retrieval-Augmented Generation (RAG) ?

b. Fonctionnement
du RAG dans notre
contexte





c. Bénéfices du RAG pour les élèves

Accès personnalisé à l'information:

- reformuler ou simplifier des notions complexes
- Idéal pour l'apprentissage autonome et la remédiation.

Réponses précises et sourcées:

- des réponses justifiées avec des extraits de documents fiables
- L'élève apprend à vérifier les sources et à comprendre les contenus.

Apprentissage actif:

- Encouragement de la curiosité en permettant aux élèves de poser des questions en langage naturel.
- L'élève apprend à vérifier les sources et à comprendre les contenus.

Accessibilité et disponibilité:

- Disponible 24/7, même en dehors des heures de cours.
- Aide précieuse pour les élèves absents ou en difficulté.



c. Bénéfices du RAG pour les enseignants

Gain de temps :

- Le système peut répondre automatiquement aux questions fréquentes des élèves.

Valorisation des ressources pédagogiques :

- Le système peut être alimenté par les documents internes de l'enseignant (cours, photocopiés, supports PDF...).

Suivi individualisé :

- Analyse des types de questions posées pour détecter les zones de difficulté.
- L'enseignant peut adapter sa pédagogie en fonction des besoins réels.

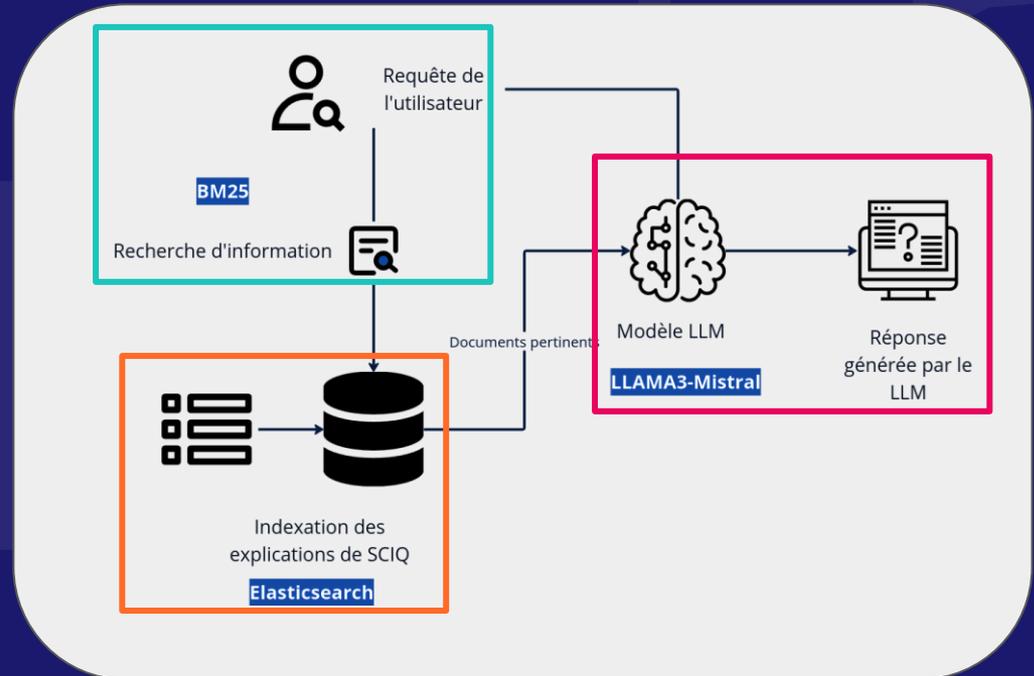
Support à l'évaluation :

- Évaluation formative possible avec feedback immédiat.

3. Notre stratégie RAG pour la génération de réponses à des questions dans un contexte d'apprentissage

a. Architecture de notre système RAG

- Intégration des ressources
- Recherche contextuelle
- Génération personnalisée



3. Notre stratégie RAG pour la génération de réponses à des questions dans un contexte d'apprentissage

b. Les données *Science* *Questions (SCIQ)*

- En anglais
- 13000 questions à choix multiples
- Libre accès

Question	What type of organism is commonly used in preparation of foods such as cheese and yogurt ?
Réponses	A : mesophilic organisms B : protozoa C : gymnosperms D : viruses
Support	Mesophiles grow best in moderate temperature, typically between 25 C and 40 C (77 F and 104 F). Mesophiles are often found living in or on the bodies of humans or other animals. The optimal growth temperature of many pathogenic mesophiles is 37 C (98 F), the normal human body temperature. Mesophilic organisms have important uses in food preparation, including cheese, yogurt, beer and wine.

3. Notre stratégie RAG pour la génération de réponses à des questions dans un contexte d'apprentissage



Transform faults are the site of massive what ?

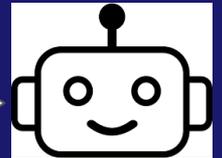
c. Cas d'usage de notre proposition

Réponse: earthquakes

Support: ID_support :94

Transform faults are the site of massive earthquakes.

Explication: In the provided context, it is stated that "Transform faults are the site of massive earthquakes" (Paragraph 94)





3. Notre stratégie RAG pour la génération de réponses à des questions dans un contexte d'apprentissage

d. Résultats de l'expérimentation

MRR: Mean Average Precision

Modèle	MRR	Rappel	R@5
Solution RAG	0,7247	0,8425	0,8175

Table 1: Evaluation de la recherche de documents pertinents

3. Notre stratégie RAG pour la génération de réponses à des questions dans un contexte d'apprentissage

d. Résultats de l'expérimentation

EM: Exact match | DL: Distance de Levenshtein | SS:similitude sémantique | LLM: Evaluation par LLM

Stratégie	EM1	EM2	DL	DL_Part	SS1 ¹	SS2 ²	P	R	F1	LLM
<i>bl-mist</i>	0,27	0,51	0,29	0,61	0,41	0,36	0,22	0,54	0,31	0,82
<i>bl-llama</i>	0,39	0,46	0,42	0,59	0,51	0,51	0,45	0,47	0,46	0,75
<i>rag-llama</i>	0,46	0,56	0,50	0,68	0,59	0,55	0,49	0,56	0,52	0,79
<i>rag-mist</i>	0,38	0,74	0,41	0,82	0,54	0,47	0,15	0,78	0,25	0,94

Table 2: Evaluation des réponses générées, avec RAG (rag) ou sans RAG (bl), avec Llama 3 (llama) ou Mistral 7b (mist).

1: <https://huggingface.co/sentence-transformers/paraphrase-MiniLM-L6-v2>

2: <https://huggingface.co/sentence-transformers/all-mpnet-base-v2>

Conclusion

Limites et défis rencontrés

Le jeu de données utilisé est bien connu et largement exploité dans la littérature, ce qui signifie que certains LLMs ont probablement été entraînés sur ce contenu. Cela peut expliquer les performances élevées observées pour les modèles de référence, LLM seuls.

Evaluation de l'explication du LLM

Qualité documentaire

La performance du système dépend directement de l'exhaustivité et de la qualité des bases documentaires intégrées.

Prompt engineering

Chaque modèle LLM peut réagir différemment au même prompt

Affiner la recherche de documents pertinents

Merci pour votre attention

Questions-Réponses
Echanges

