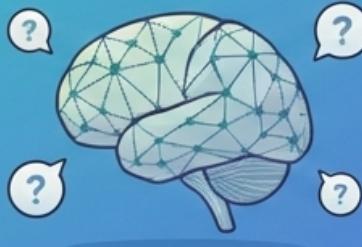


# Protocole de benchmark : Comparaison des 3 Stratégies IA pour un Assistant FAQ

Comparatif visuel pour sélectionner l'approche la plus performante d'assistant intelligent.

## COMPARATIF DES 3 STRATÉGIES

Stratégie A : LLM Seul



### Simple & Rapide

- Risque élevé d'hallucinations (réponses incorrectes).



Mistral-7B



Fiabilité

Stratégie B : RAG (Retrieval-Augmented Generation)



### Fiable & Contextuel

- S'appuie sur des documents pour la précision.
- Plus complexe et lent.



Fiabilité



Vitesse

Stratégie C : Q&A Extractif



### Extrêmement Précis

- Zéro hallucination.
- Génère des réponses "robotiques".



Fiabilité



Style de Réponse

## PRINCIPAUX RISQUES ET SOLUTIONS

RISQUE:  
Hallucinations

SOLUTION:  
Utiliser un prompt strict et des stratégies comme le RAG.

RISQUE:  
Latence > 5s

SOLUTION:  
Choisir des modèles d'embeddings légers et optimisés.

RISQUE:  
Indisponibilité de l'API

SOLUTION:  
Privilégier un modèle local ou mettre en place un système de cache.