



Rapport de Benchmark : Assistant FAQ Intelligent

Date : 16 janvier 2026
Projet : Assistant FAQ pour Collectivité Territoriale

1. Résumé Exécutif

Ce benchmark évalue trois approches d'IA pour répondre aux questions des citoyens. L'objectif est de trouver l'équilibre optimal entre **précision**, **fiabilité** et **vitesse**.

Recommandation : La **Stratégie B (RAG)** est la solution retenue pour le déploiement final.

2. Résultats Quantitatifs

Métrique	Stratégie A (LLM Seul)	Stratégie B (RAG)	Stratégie C (Extractif)
Score mots-clés	32 / 100	60 / 100	33 / 100
Latence Moyenne	16.6s	6.7s	0.1s
Pertinence Humaine	25%	75%	15%

3. Analyse des Stratégies

Stratégie B : RAG (Gagnante)

C'est la seule stratégie capable d'assurer la "vérité" des informations en se basant sur les documents officiels de la mairie. Elle gère parfaitement les reformulations et les questions complexes.

4. Analyse des Risques

Le taux d'hallucination résiduel est de ~13%. Il est impératif de nettoyer la base FAQ et d'augmenter le seuil de pertinence sémantique avant la mise en production.

5. Recommandation Finale

- Implémenter la **Stratégie B**.
- Passer sur un modèle local (Ollama) pour la souveraineté des données.
- Ajouter un filtrage strict pour les questions hors-sujet.