

Rapport d'Analyse Pondéré (Benchmark Final)

Méthodologie du Calcul

Conformément aux critères définis pour le projet, chaque stratégie a été évaluée selon la pondération suivante :

- **Exactitude (30%)** : Score mots-clés sur le Golden Set.
- **Pertinence (20%)** : Qualité de la réponse (échelle 0-2).
- **Hallucinations (20%)** : % de réponses garanties sans invention.
- **Latence (15%)** : Temps de réponse (Bonus si < 5s).
- **Complexité (15%)** : Soutenabilité et maintenance.

Critères (Poids)	Strat A (LLM Seul)	Strat B (RAG)	Strat C (Extr.)
Exactitude (30%)	32 / 100	60 / 100	33 / 100
Pertinence (20%)	25 / 100	75 / 100	15 / 100
Hallucinations (20%)	20 / 100	85 / 100	95 / 100
Latence (15%)	10 / 100	55 / 100	100 / 100
Complexité (15%)	90 / 100	40 / 100	60 / 100
NOTE FINALE (/100)	33.6	64.3 GAGNANT	55.9

Conclusion du Calcul Pondéré

La **Stratégie B (RAG)** l'emporte avec une note finale de **64.3/100**.

Bien que la Stratégie C soit plus rapide et moins complexe, elle échoue sur les critères de **Pertinence** et d'**Exactitude** (mots-clés), ce qui la rend inutilisable pour un service public. La

Stratégie A est disqualifiée par son taux d'hallucination trop élevé.

Document technique généré pour la soutenance J6 - 13 février 2026
Projet : Assistant FAQ Intelligent pour Collectivité Territoriale