



# Grille d'Évaluation du Benchmark

Ce document sert à noter les performances de l'assistant intelligent. Pour chaque question, testez les stratégies et attribuez une note.

## 🎯 Critères de Notation

Note	Signification	Description
5	<b>Excellent</b>	Réponse exacte, complète et naturelle.
4	<b>Bon</b>	Réponse correcte mais style ou détail améliorable.
3	<b>Moyen</b>	Réponse partiellement correcte ou vague.
2	<b>Insuffisant</b>	Hors sujet, incomplet ou format cassé.
1	<b>Hallucination</b>	Informations fausses ou inventées (Critique).

## 💡 Scores Moyens Globaux

Stratégie	Score Moyen / 100	Latence Moyenne	Verdict
Stratégie A (LLM Seul)	32	16 663 ms	Insuffisant / Hallucinations
Stratégie B (RAG)	60	6 688 ms	Recommandé / Fiable
Stratégie C (Extractif)	33	130 ms	Rapide / Trop concis

## 📝 Détail des Tests (Échantillon)

ID	Question (Exemples)	Strat A (LLM Seul)	Strat B (RAG)	Strat C (Extractif)
1	Quel est le Maire ?	Note : ___ /5	Note : ___ /5	Note : ___ /5
2	Horaires de la piscine ?	Note : ___ /5	Note : ___ /5	Note : ___ /5
3	Aides au logement ?	Note : ___ /5	Note : ___ /5	Note : ___ /5
4	Prochain conseil municipal ?	Note : ___ /5	Note : ___ /5	Note : ___ /5
5	Comment trier le verre ?	Note : ___ /5	Note : ___ /5	Note : ___ /5
6				
7				
8				
9				
10				
11				
12				