

Grille d'Évaluation Agentic - Solution 1

#	Critère d'Évaluation	Poids (%)	Description Détaillée
1	Extraction, Chunking et Embedding	12	Extraction contenu des documents KB, chunking des documents (statique ou sémantique). Stockage vectoriel. Test similarité >0.8 sur échantillon KB.
2	Agent Query Analyzer	9	Résumé (<100 mots) + extraction 5-10 mots-clés. Métriques retournées : résumé, keywords
3	Agent Solution Finder (RAG)	9	Retrieve N=5 docs + reranking. Snippets triés + scores confiance (0-1).
4	Agent Evaluator et Decider	9	Confiance moyenne <0.6 → escalade. Détection émotions négatives, exceptions. Règles configurables.
5	Détection de Données Sensibles	9	Patterns pour cartes, emails, phones. Escalade 100%.
6	Agent Response Composer	9	Template : remerciements + problème + solution + action. Détection langue
7	Qualité des Réponses IA	8	les réponses au client ne doivent pas avoir des hallucination. Ton professionnel. Couvre 100% questions
8	Gestion des Erreurs du Pipeline	8	Fallbacks intelligents. Circuit breaker (retries 3x). Logging structuré avec trace_id
9	Latence du Pipeline	6	<10s end-to-end (<5s better)
10	Intégration au Back-end	7	APIs cohérentes. Webhooks ou polling. Sauvegarde async des résultats. Gestion des timeouts.
11	Stockage des Résultats IA	7	Tables : summary, keywords, rag_docs, response, confidence, escalade_reason. Index optimisés.
12	Prompt Engineering	7	rediger un bon system prompt : rôle, format, etc.
	TOTAL PRINCIPAL	100	_/100
#	Bonus	Poids (%)	Description
13	Détection Automatique de Catégorie	6	Classification (accès, technique, commercial). Confiance score. Fallback si <0.7. Précision >85% sur 50 tickets
14	Caching Embeddings/RAG	10	Cache persistant. TTL 24h. Hit rate >70%. Invalidation KB update
15	Support Multimodal KB	7	OCR pour PDFs/images. Retrieval cross-modal
16	Monitoring & Observabilité	2	Traces end-to-end. Métriques précision/coût. Alertes
	TOTAL BONUS	25	_/25
	TOTAL GÉNÉRAL	125	_/125