

# RAG System Evaluation

A high-performing Retrieval-Augmented Generation (RAG) system is more than a technical achievement; it's a source of truth that your organisation can depend on. To move beyond subjective impressions, a structured evaluation is essential.

This worksheet provides a simple, powerful framework to measure the practical value of your RAG implementation. Use it to gain an objective understanding of your system's current performance and to identify clear pathways for improvement. By focusing on these core pillars of quality, you ensure your AI delivers intelligence you can trust and act on.

## The FACT Framework

We assess RAG systems against four crucial dimensions. This FACT framework provides a balanced view of performance, ensuring the system is not only accurate but also genuinely useful for the end user.

- **Fidelity:** How accurately does the response reflect the source documents, without contradiction or distortion?
- **Answer Relevance:** How directly does the answer address the specific intent behind the user's query?
- **Completeness:** Does the answer contain all key information from the source material needed to satisfy the query?
- **Trustworthiness:** Is the response properly grounded in the retrieved context, with clear source attribution?

## Your System's FACT Profile

Rate each dimension from 1 (Poor) to 5 (Excellent) to map your system's current capability.

- **Fidelity:** \_\_\_\_\_ / 5
- **Answer Relevance:** \_\_\_\_\_ / 5
- **Completeness:** \_\_\_\_\_ / 5
- **Trustworthiness:** \_\_\_\_\_ / 5

## Performance Scorecard

A robust evaluation requires testing against a diverse set of real-world queries. Use this scorecard to document system performance on a case-by-case basis. This granular analysis is key to pinpointing specific weaknesses and tracking the impact of system tuning.

A consistent, high average FACT score across many queries is the hallmark of an enterprise-ready RAG solution. Analyse low-scoring responses to identify patterns—are sources inadequate, is retrieval failing, or is the generation layer struggling with synthesis?

Test Query	System Response (Summary)	FACT Score (1-5)	Notes / Improvements

**Ready to translate AI possibility into practical reality?** Book a 30-minute discovery call at [adsum.ai](https://adsum.ai)