

Evaluation PDF – Two Page Source Document

Page 1

This document is created specifically for testing and evaluating Document Intelligence, RAG pipelines, and grounded QA systems. The first page introduces the purpose, structure, and expected usage patterns of the document.

Key Concepts Introduced:

- PDF parsing reliability across structured and unstructured layouts.
- Chunking strategy evaluation, including paragraph, fixed-size, and semantic chunking.
- Embedding quality checks using Google GenAI Embeddings (text-embedding-004).
- Chroma in-memory and persistent index behavior during ingestion.

Page 2

The second page provides richer semantic material, suitable for evaluation of retrieval quality, semantic similarity scoring, and grounded answer generation.

Evaluation Criteria Discussed:

1. Recall@K performance across varying query types.
2. Groundedness and hallucination resistance when answering factual questions.
3. Metadata propagation across ingest → chunk → embed → store → query pipelines.
4. Behavior of query classifiers in multi-tier RAG architectures.

This document also contains intentionally overlapping semantics between pages to test whether retrieval resolves ambiguity using embeddings rather than lexical signals.