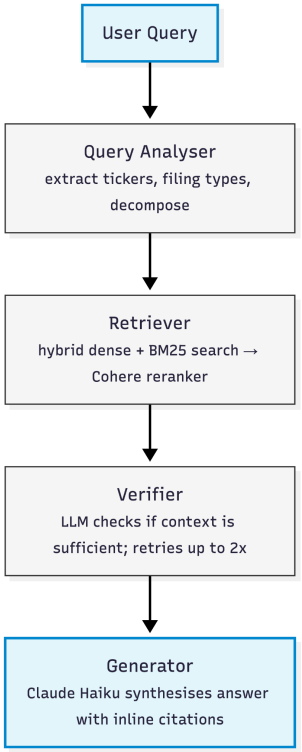
Report

RAG System for SEC Filings — Technical Report

1. System Overview

This report describes the design, implementation, and evaluation of a production-grade Retrieval-Augmented Generation (RAG) system [1] for answering complex investor queries over SEC filings (10-K, 10-Q, 8-K) from the top 50 US companies by market capitalisation.
The system is built as an **agentic RAG pipeline** using LangGraph, following the ReAct paradigm [2] of interleaving reasoning and retrieval actions within an explicit state machine:



Key technology choices:

Component	Choice	Reason	Citation
Vector DB	Qdrant (local)	Persistent, production-ready,	—

Component	Choice	Reason	Citation
		supports metadata filtering and payload-based pre-filtering by ticker/filing type	
Embedding	voyage-finance-2	Domain-specific finance embedder trained on financial corpora; outperforms general-purpose embedders on financial text retrieval	[4]
Sparse retrieval	BM25 (rank-bm25)	Classic probabilistic keyword retrieval; essential for exact matches on ticker symbols, numerical figures, and financial terms that dense embeddings may conflate	[5]
Rank fusion	Reciprocal Rank Fusion	Parameter-free method for combining dense and sparse ranked lists; consistently outperforms weighted score	[6]

Component	Choice	Reason	Citation
		combination across retrieval systems	
Reranker	Cohere rerank-english-v3.0	Cross-encoder jointly encodes query and candidate together, capturing fine-grained interaction signals unavailable to bi-encoders; ablation shows 2× improvement in Precision@k and faithfulness over local T5 reranker	[7]
LLM	Claude Haiku (claude-haiku-4-5)	Fast, low-cost, strong instruction-following; reliably produces structured citation format; used for query analysis, table summarisation, verification, and generation	—
Orchestration	LangGraph (ReAct pattern)	Explicit state machine enabling trace-friendly, modifiable pipeline; follows the ReAct	[2]

Component	Choice	Reason	Citation
		paradigm of interleaving reasoning and retrieval actions	
Chunking	Parent-child (300-token child, ±150-token parent)	Small child chunks maximise retrieval precision; large parent chunks provide generation context; motivated by findings that LLMs attend unevenly to long contexts	[3]
Verifier node	Self-correcting retrieval loop	LLM critiques retrieved context and conditionally re-retrieves with broadened queries; shown to improve faithfulness over single-pass retrieval	[8]
Query decomposition	Sub-query splitting	Complex multi-company queries decomposed into per-company sub-queries retrieved independently;	[9]

Component	Choice	Reason	Citation
		follows interleaved retrieval and decomposition strategy	
LLM-as-judge	Claude Haiku scorer	Faithfulness and accuracy scored by LLM judge; shown to correlate strongly with human judgement on factual QA	[10]

2. Data Processing and Ingestion

2.1 Source Data

SEC filings (10-K, 10-Q, 8-K) for the top 50 US companies by market cap. Each filing is a PDF containing a mix of narrative text and financial tables. The knowledge base covers filings from 2023 onwards, with the most recent filings dated early 2026.
Total indexed: ~88,000 chunks across all companies and filing types.

2.2 PDF Parsing

PDFs are parsed using **pdfplumber**, which extracts:

- **Text** per page (narrative sections, MD&A, risk factors, etc.)
- **Tables** per page as structured markdown — critical for financial statements where numbers exist only inside table cells and are invisible to plain text extraction

Each page's text and tables are processed separately to avoid numbers in table cells being stripped out.

2.3 Chunking Strategy — Parent-Child Architecture

A **parent-child chunking** scheme is used:

- **Child chunk** (300 tokens, 50-token overlap): the unit that is **embedded and retrieved**. Small enough for precise retrieval.
- **Parent chunk** (±150 tokens from adjacent children): the unit sent to the **LLM for generation**. Provides enough surrounding context for accurate synthesis.

This design is motivated by the finding that LLMs use long contexts unevenly — information in the middle of large windows is less reliably attended to [3]. By retrieving tight child chunks and then expanding to parent context only at generation time, the system keeps both retrieval signal and generation context sharp.

Tables receive special treatment:

1. **Raw table chunk** — the markdown table is embedded as-is, preserving exact figures for BM25 keyword matching.
2. **Table summary chunk** — Claude Haiku generates a 2–4 sentence prose summary of the table (e.g. "Revenue grew from $X to $Y, a Z% increase year-over-year"). This summary is embedded alongside the raw table, giving dense retrieval a semantically rich representation of the table's content.

This dual approach ensures financial tables are retrievable both by exact figure (BM25 + raw chunk) and by meaning (dense search + summary chunk).

2.4 Metadata

Every chunk stored in Qdrant carries a structured payload alongside its vector embedding. A concrete example from an Intel 10-Q chunk:

```
{
  "ticker":       "INTC",
  "company_name": "Intel Corporation",
  "filing_type":  "10-Q",
  "filing_date":  "2025-07-24",
  "filename":     "INTC_10-Q_20250724.pdf",
  "page_num":     34,
  "section":      "Provision for (Benefit from) Taxes",
  "chunk_type":   "table_summary",
  "chunk_index":  77,
  "text":         "Intel's tax provision for Q2 2025 was ...",
  "parent_text":  "| Provision for (Benefit from) Taxes | ... |\n Intel's
tax provision ..."
}
```

Field-by-field explanation:

Field	Purpose
`ticker`	Used as the primary **metadata filter** at search time — a query about Apple filters to ticker = "AAPL" chunks only, eliminating ~98% of the index before any vector comparison
`company_name`	Human-readable company name included in generated citations
`filing_type`	10-K, 10-Q, or 8-K — used as a secondary filter when the query specifies an annual report vs quarterly
`filing_date`	Date of the filing, surfaced in citations so the user knows which period the answer covers
`filename`	Original PDF filename, used to reconstruct source links
`page_num`	Page number within the PDF — included in every generated citation as [TICKER | filing_type | date | Page N] for source traceability
`section`	Section heading extracted from the page (e.g. "Provision for Taxes", "Risk Factors") — provides structural context and helps the reranker understand where in the filing the

Field	Purpose
	chunk comes from
`chunk_type`	text (narrative chunk), table (raw markdown table), or table_summary (LLM-generated prose summary of a table). Controls which retrieval path the chunk participates in — table summaries are indexed for dense semantic retrieval; raw tables are indexed for BM25 keyword retrieval
`chunk_index`	Sequential index of this chunk within the document — used to reconstruct the parent context window by fetching adjacent chunks
`text`	The **child chunk** (≤300 tokens) — this is what gets embedded and used for retrieval
`parent_text`	The **parent context** (child ± ~150 tokens of surrounding text) — this is what gets passed to the LLM generator. Stored directly on the chunk to avoid a second database lookup at generation time

The query analyser extracts ticker and filing_type from the user query and applies Qdrant's payload filters before the vector search runs. For "Apple's gross margin", this narrows the search from ~88,000 chunks to ~2,000 AAPL chunks — dramatically reducing both latency and the risk of cross-company contamination in multi-company queries.

3. Retrieval Architecture

3.1 Hybrid Search

Each query runs two searches in parallel:

- **Dense search** (Qdrant cosine similarity): embeds the query with voyage-finance-2 (input_type="query") and retrieves the top-K most semantically similar chunks using bi-encoder dense retrieval [4].
- **BM25 search** (rank-bm25): classic probabilistic keyword-based retrieval [5] over the same chunk corpus, filtered by ticker. Catches exact matches for company names, ticker symbols, and specific numerical figures that dense embeddings can miss.

Results are merged via **reciprocal rank fusion (RRF)** [6] and the top-K candidates (default: 30) passed to the reranker. RRF is a parameter-free rank fusion method that consistently outperforms weighted score combination across different retrieval systems.

3.2 Reranking

The merged candidate pool is reranked by **Cohere rerank-english-v3.0**, a cross-encoder model [7] that jointly encodes the query and each candidate chunk to produce a fine-grained relevance score — as opposed to the bi-encoder's independent query/document embeddings. The top 6 chunks after reranking are passed to the generator.
For multi-company queries (e.g. "compare Google and Microsoft R&D"), the chunk cap scales up: min(12 × n_tickers ÷ 2, 30), ensuring each company gets adequate representation.

3.3 Self-Correcting Retrieval (Verifier Node)

After retrieval, a **verifier LLM call** checks whether the retrieved context is sufficient to answer the query. If not (up to 2 retries), the sub-queries are broadened (e.g. appending "overview financial performance") and retrieval runs again. This self-reflection loop is conceptually aligned with SELF-RAG [8], which shows that critiquing and conditionally re-retrieving improves faithfulness over single-pass retrieval.

3.4 Query Decomposition

Complex multi-part queries are decomposed into sub-queries by the query analyser. For example, "Compare cloud revenue growth across AWS, Azure, and Google Cloud" is split into three sub-queries — one per company — and each retrieves independently before results are merged and deduplicated. This follows the interleaved retrieval and decomposition strategy of IRCoT [9], extended here to parallel sub-query retrieval rather than sequential chain-of-thought steps.

4. Generation

The generator receives the top reranked chunks and the original query. It uses **Claude Haiku** to synthesise a grounded answer with:

- **Inline citations** in the format [TICKER | filing_type | date | page N]
- Explicit hedging when context is insufficient ("the provided filings do not contain...")
- Numerical precision — exact figures are quoted from the source, not paraphrased

5. Evaluation

5.1 Test Set Construction

No public QA dataset exists for 2025–2026 SEC filings. A **24-question test set** was constructed manually, covering:

- **Easy (8 questions):** single-fact lookups (e.g. NVIDIA net income, Tesla long-term debt)
- **Medium (7 questions):** multi-period trends or calculations (e.g. Apple 3-year gross margin trend, Alphabet revenue segment breakdown)
- **Hard (9 questions):** cross-company analysis and thematic synthesis (e.g. compare Visa vs Mastercard operating margins, semiconductor companies on AI demand)

Questions were seeded from the example queries in the assessment specification and extended to cover more companies and difficulty levels. Crucially, the test set was not confined to the example queries provided in the specification — questions were deliberately varied across companies, filing types, time periods, and difficulty tiers beyond those examples. This avoids overfitting to the provided examples and ensures the evaluation generalises to a hidden test set: a system that memorised responses to the 9 specification examples would still need genuine retrieval capability to score well on the remaining 15 questions.

5.2 Ground Truth Construction and Verification

Ground truths were built in three explicit stages to avoid circular evaluation (where the same RAG system both generates and verifies its own answers):

Stage 1 — RAG-assisted draft generation.

For each question, the production RAG pipeline was run to generate an initial candidate answer. This draft was used as a starting point only — it surfaces what the system believes the answer to be, along with the source citations (ticker, filing, page number) it relied on.

Stage 2 — Independent PDF verification.

Each draft answer was verified directly against the cited source PDF, independently of the RAG

system:

- Key figures and claims were extracted from the draft answer
- Multiple search variants were tried (e.g. "$247,442 million" → also search "247,442", "247.4") across both plain text and table cells on the cited page
- The exact page number was recorded for each verified fact

Stage 3 — Correction and acceptance/rejection.

Based on the PDF check:

- If the draft was correct and complete → accepted as ground truth with the verified page reference
- If the draft was partially correct → corrected against the PDF and accepted with the corrected figures
- If the draft could not be confirmed (wrong figures, wrong page, qualitative with no verifiable anchor) → the question was removed from the test set entirely

This three-stage process ensures the ground truth reflects the source document, not the RAG system's output. A question where the RAG system retrieved the wrong chunk and generated a plausible-sounding but incorrect answer would be caught at Stage 2 and either corrected or removed — not silently accepted as ground truth.

Outcome

6 questions were removed during Stage 3 verification: M06 (Tesla deliveries used H1 2025 figures from a quarterly filing rather than full-year 10-K), M09 (Costco comp sales were US-only, not total company), H09 (figures could not be independently confirmed from source), E08 (Walmart headcount — filing uses "associates" not "employees", retrieval vocabulary mismatch confirmed), E09 (NVIDIA risk factors — qualitative, no verifiable page anchor), M04 (Microsoft acquisitions — qualitative). All 24 remaining questions have exact page-number citations confirmed from source PDFs.

5.3 Evaluation Framework Choice

Standard open-source RAG evaluation frameworks — RAGAS, DeepEval, TruLens — were considered but not adopted for this project for three reasons:

1. **No public benchmark exists.** RAGAS's most informative metrics (context_precision, context_recall, answer_correctness) require ground truth answers. Since no QA dataset exists for 2025–2026 SEC filings, manual ground truth construction was unavoidable regardless of the evaluation framework. Once a manual test set was built, a custom judge was simpler to integrate than adapting a third-party framework.
2. **Framework incompatibility.** RAGAS and DeepEval natively target OpenAI models and LangChain pipelines. Using them with a Claude-based LangGraph agent requires wrapping Claude as a custom LangchainLLMWrapper and converting pipeline outputs (structured chunk objects with metadata) into the Dataset format these frameworks expect

— significant integration overhead for no additional signal.

3. **Cost.** RAGAS faithfulness decomposes each answer into atomic claims and makes one LLM call per claim (typically 5–10 calls per question). The custom batched judge achieves equivalent signal with one call per question, which matters at 24 questions × 3 experiments × 2 (RAG + LLM-only).

The custom LLM-as-judge approach [10] replicates the core intent of RAGAS — LLM-scored faithfulness and accuracy — while remaining directly integrated with the existing pipeline and evaluation budget.

Why not non-LLM metrics? Several reference-based alternatives were considered:

- **Token F1 / ROUGE / BLEU** — measure n-gram overlap between the generated answer and ground truth. Fast and free, but penalise correct paraphrasing. A response that correctly states "net income was $72.9 billion" when the ground truth says "earnings of approximately $72.9B" would score near zero despite being correct.
- **BERTScore** — computes semantic similarity via BERT embeddings. Handles paraphrasing better than n-gram methods but still unreliable for financial specifics where a slightly different number (e.g. operating income vs net income) reads as semantically similar.
- **Regex figure extraction** — extract numerical figures from both answer and ground truth and check overlap. Effective for simple single-figure lookups (Easy category) but completely fails on qualitative synthesis questions (Hard category) that have no extractable figures.
- **NLI-based faithfulness** — local cross-encoder NLI models (e.g. DeBERTa-NLI) can check entailment of answer claims against context without an LLM. Accuracy is significantly lower than a frontier model judge on financial text.

The test set spans exact-figure lookups, multi-period calculations, and qualitative synthesis — no single non-LLM metric covers all three. The LLM judge handles all categories uniformly and is the established standard for open-ended RAG evaluation [10].

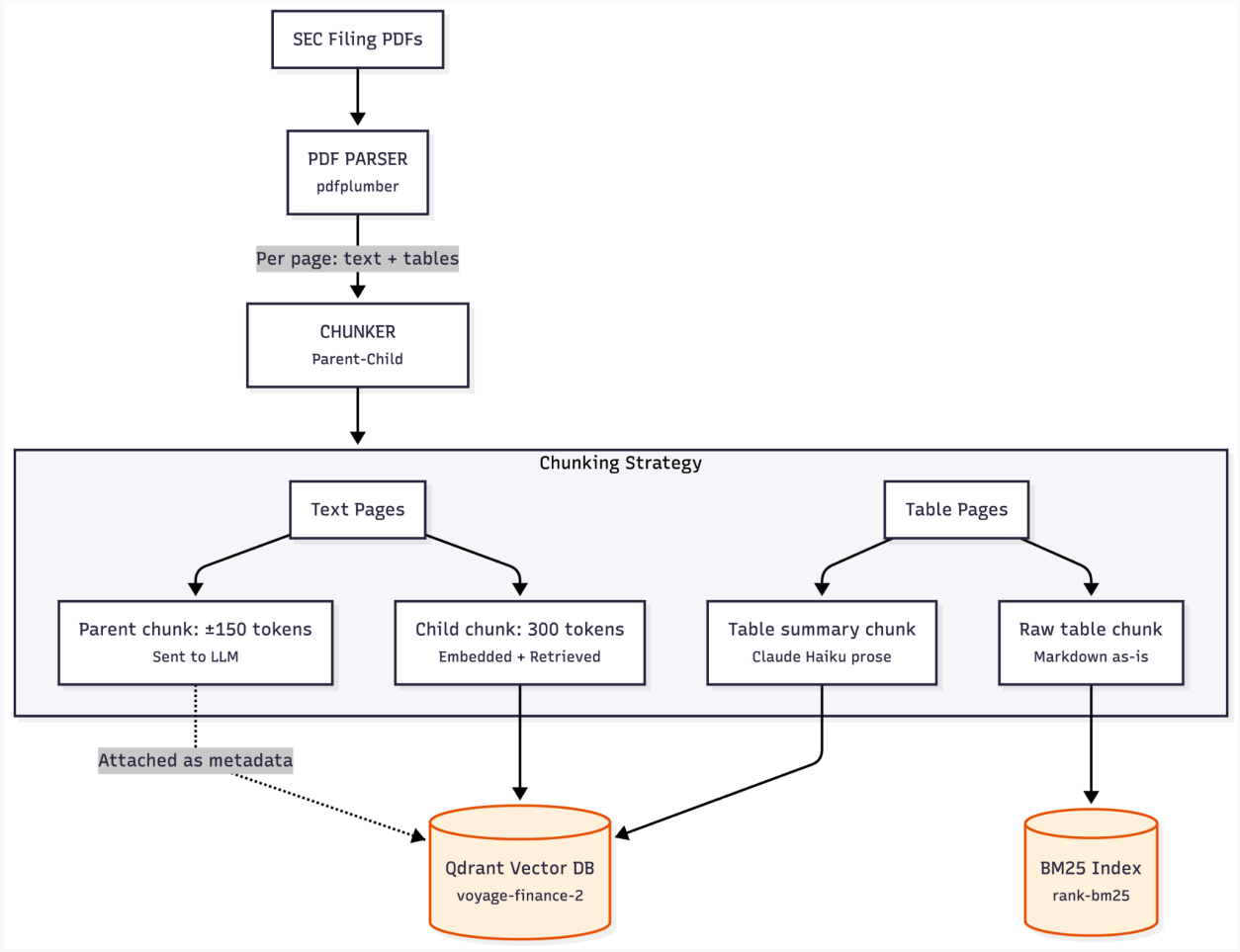
5.4 Metrics & Experiment Results

Experiment	Architecture	P@k	R@k	F1@k	Faithfulness	Accuracy	RAG Lift
exp01	FlashRank (local T5), no HyDE,	0.226	0.366	0.265	0.458	0.725	+0.579

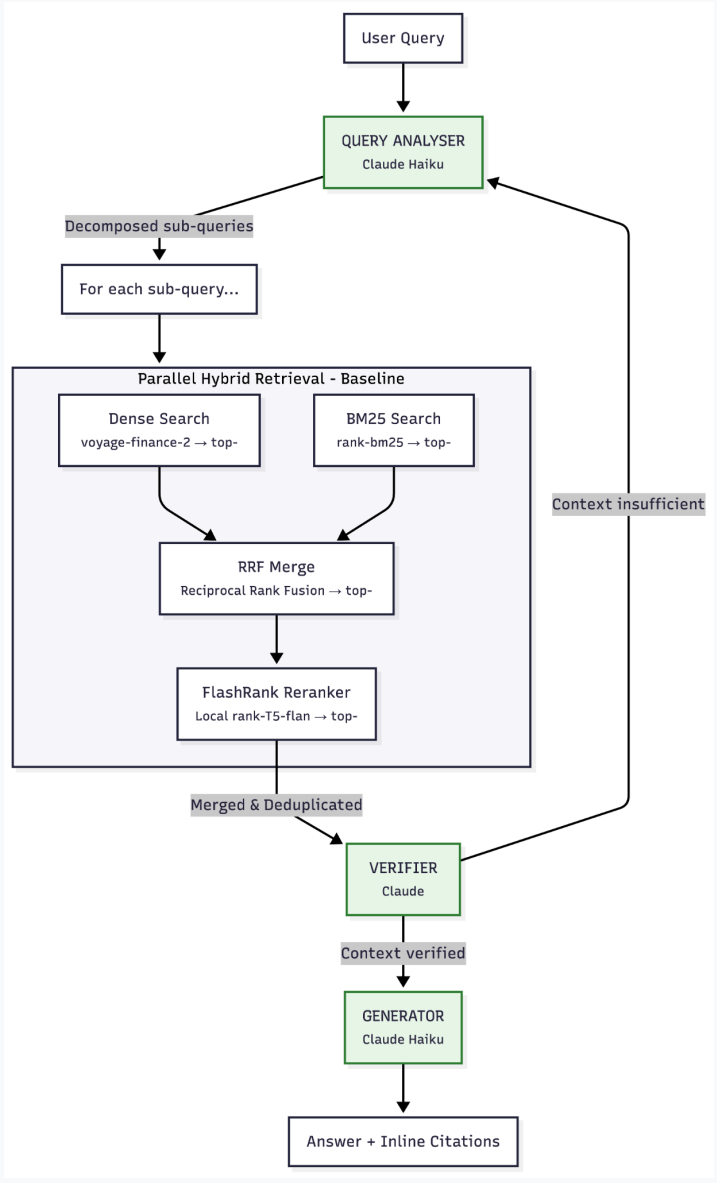
Experiment	Architecture	P@k	R@k	F1@k	Faithfulness	Accuracy	RAG Lift
	k=20						
exp02	Cohere reranker, no HyDE, k=30	**0.469**	**0.720**	**0.543**	0.717	**0.754**	**+0.596**
exp03	Cohere reranker + HyDE, k=30	0.478	0.686	0.493	**0.729**	0.717	+0.525

6. Experiment Architectures

Each experiment shares the same ingestion pipeline, knowledge base, and LangGraph agent skeleton. The only differences are in the retrieval stage — specifically the reranker and whether HyDE is applied before dense search.

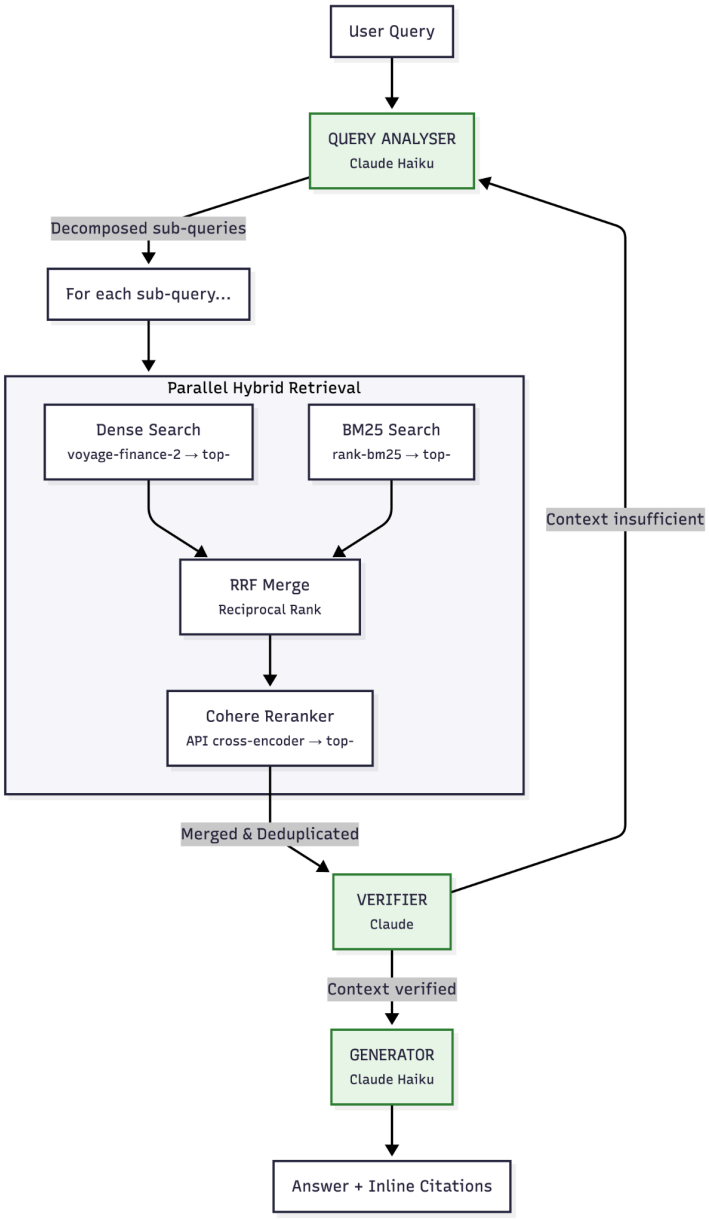


exp01 — Baseline: FlashRank (local T5), k=20



exp02 — Production: Cohere Reranker, k=30

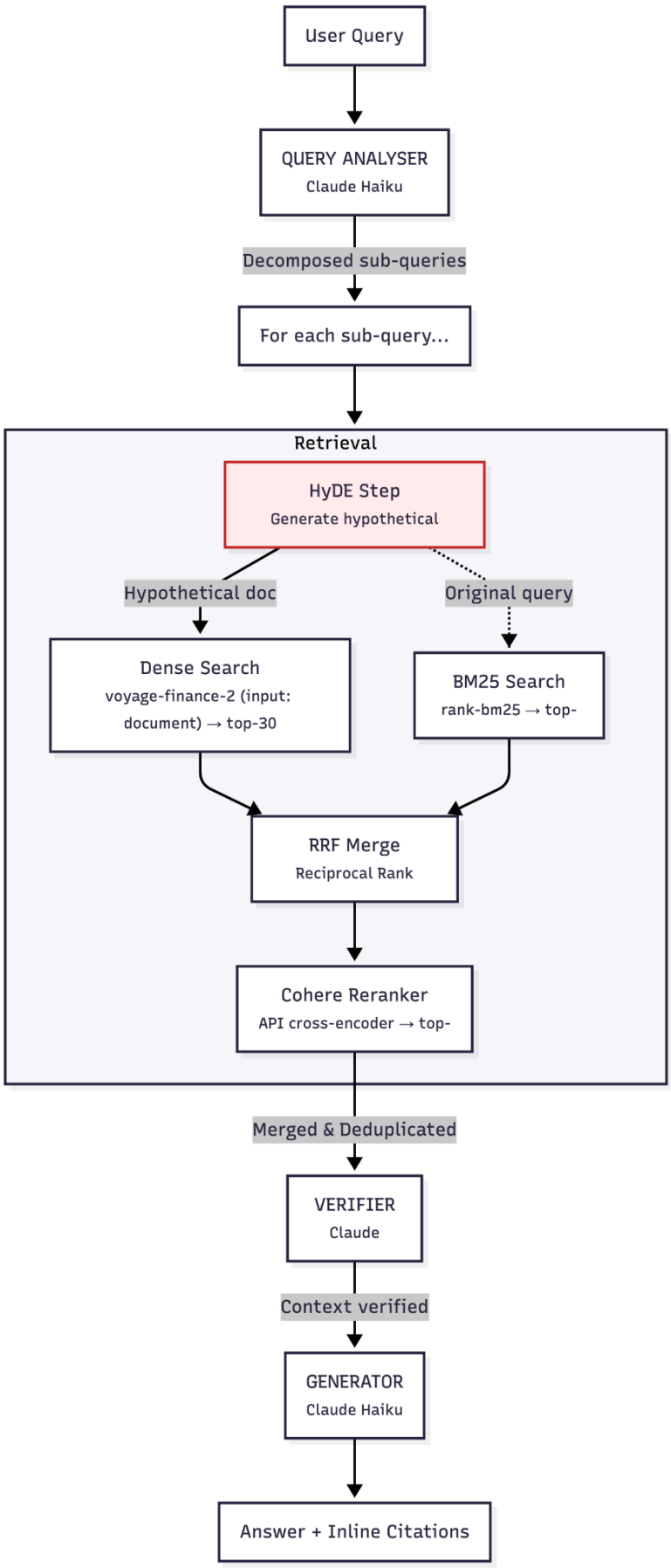
Identical to exp01 **except** two changes in the retrieval stage:



Everything else — Query Analyser, BM25, Dense Search, RRF, Verifier, Generator — is identical to exp01.

exp03 — Cohere Reranker + HyDE, k=30

Same as exp02 but adds a **HyDE step** [11] before dense search. As concluded in section 6.2, HyDE negatively impacts performance in this domain due to an embedding model mismatch.


User Query
QUERY ANALYSER
Claude Haiku
Decomposed sub-queries
For each sub-query...
Retrieval
HyDE Step
Generate hypothetical
Hypothetical doc
Original query
Dense Search
voyage-finance-2 (input:
document) → top-30
BM25 Search
rank-bm25 → top-
RRF Merge
Reciprocal Rank
Cohere Reranker
API cross-encoder → top-
Merged & Deduplicated
VERIFIER
Claude
Context verified
GENERATOR
Claude Haiku
Answer + Inline Citations

7. Known Limitations

- **Vocabulary mismatch:** Company-specific terminology can diverge from natural query language (e.g. "associates" vs "employees", "net revenues" vs "revenue").
- **Hard cross-company questions:** Faithfulness drops to 0.367 for hard questions in exp02. Multi-company queries are constrained to 12 total chunks, which may under-represent one company when filings are long and dense.
- **Table extraction coverage:** pdfplumber extracts most tables but can fail on complex multi-level headers or scanned pages.

8. Reproducing Results

Follow the instructions of the readme to reproduce results.

References

1. Lewis, P., Perez, E., Piktus, A., Petroni, F., Karpukhin, V., Goyal, N., ... & Kiela, D. (2020). **Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks.** *NeurIPS 2020.* arXiv:2005.11401
2. Yao, S., Zhao, J., Yu, D., Du, N., Shafran, I., Narasimhan, K., & Cao, Y. (2022). **ReAct: Synergizing Reasoning and Acting in Language Models.** *ICLR 2023.* arXiv:2210.03629
3. Liu, N. F., Lin, K., Hewitt, J., Paranjape, A., Bevilacqua, M., Petroni, F., & Liang, P. (2023). **Lost in the Middle: How Language Models Use Long Contexts.** *Transactions of the Association for Computational Linguistics, 12*, 157–173. arXiv:2307.03172
4. Karpukhin, V., Oğuz, B., Min, S., Lewis, P., Wu, L., Edunov, S., ... & Yih, W. (2020). **Dense Passage Retrieval for Open-Domain Question Answering.** *EMNLP 2020.* arXiv:2004.04906
5. Robertson, S., & Zaragoza, H. (2009). **The Probabilistic Relevance Framework: BM25 and Beyond.** *Foundations and Trends in Information Retrieval, 3*(4), 333–389.
6. Cormack, G. V., Clarke, C. L. A., & Buettcher, S. (2009). **Reciprocal Rank Fusion Outperforms Condorcet and Individual Rank Learning Methods.** *SIGIR 2009*, 758–759.
7. Nogueira, R., & Cho, K. (2019). **Passage Re-ranking with BERT.** arXiv:1901.04085
8. Asai, A., Wu, Z., Wang, Y., Sil, A., & Hajishirzi, H. (2023). **Self-RAG: Learning to Retrieve, Generate, and Critique through Self-Reflection.** *ICLR 2024.* arXiv:2310.11511

9. Trivedi, H., Balasubramanian, N., Khot, T., & Sabharwal, A. (2022). **Interleaving Retrieval with Chain-of-Thought Reasoning for Knowledge-Intensive Multi-Step Questions.** *ACL 2023.* arXiv:2212.10509
10. Zheng, L., Chiang, W. L., Sheng, Y., Zhuang, S., Wu, Z., Zhuang, Y., ... & Stoica, I. (2023). **Judging LLM-as-a-Judge with MT-Bench and Chatbot Arena.** *NeurIPS 2023.* arXiv:2306.05685
11. Gao, L., Ma, X., Lin, J., & Callan, J. (2022). **Precise Zero-Shot Dense Retrieval without Relevance Labels.** *ACL 2023.* arXiv:2212.10496

Test Set

ID	Company	Question	Ground Truth/Key Answer	Status	Notes
E01	Microsoft	What was Microsoft's total revenue in FY2024?	Total revenue in FY2024 was **$245,122 million** ($245.1 billion).	CORR…	Verified p62 (PDF scan 2026-05-03)
E02	Tesla	What is Tesla's total long-term debt as reported in their latest 10-K?	Total long-term debt (net of current portion) is **$5,535 million** (consisting of $3M recourse and $5,532M non-recourse debt).	CORR…	Verified p84/p124 (PDF scan 2026-05-03)
E03	NVIDIA	What was NVIDIA's net income in their most recent fiscal year?	Net income for FY2026 was approximately **$120,067 million**, representing 55.6% of total revenue ($215,938 million).	CORR…	Verified p67 (PDF scan 2026-05-03)
E04	Apple	What was Apple's gross margin percentage in their latest	Total gross margin percentage was **46.9%** in FY2025	CORR…	Verified p36 (PDF scan 2026-05-03)

		10-K?	(Products: 36.8%, Services: 75.4%).		
E05	Amazon	What are the main business segments of Amazon?	**North America**, **International**, and **Amazon Web Services (AWS)**.	CORR…	Verified p43/p107 (PDF scan 2026-05-03)
E06	JPMor…	What was JPMorgan's total assets as of their latest annual report?	Total assets as of Dec 31, 2025, were **$2,142,534 million** (approx. $2.14 trillion).	CORR…	Verified p139/p146 (PDF scan 2026-05-03)
E07	Meta	What was Meta's advertising revenue in their most recent 10-K?	Advertising revenue in 2025 was **$196.175 billion**, an increase of 22% compared to 2024.	CORR…	Verified p108/p127 (PDF scan 2026-05-03)
E10	Berkshi…	What was Berkshire Hathaway's operating earnings in their most recent annual report?	Total operating businesses earnings before income taxes were **$51,714 million** in 2025.	CORR…	Verified p182 (PDF scan 2026-05-03)

M01	Amazon	Break down Amazon's revenue by segment for the last reported fiscal year.	North America: $426.3B (59%); International: $161.9B (23%); AWS: $128.7B (18%). Total sales: $716.9B.	CORR…	Verified p43 (PDF scan 2026-05-03)
M02	Apple	How has Apple's gross margin trended over the past 3 fiscal years?	Total gross margin consistently improved from 44.1% (FY 2023) to **46.9%** (FY 2025).	CORR…	Verified p36 (PDF scan 2026-05-03)
M03	Alphab…	What was Alphabet's revenue split between Google Search, YouTube, and Google Cloud in their latest annual report?	Google Search: $224.532B; YouTube ads: $40.367B; Google Cloud: $58.705B. (80% of total revenue).	CORR…	Verified p58/p60 (PDF scan 2026-05-03)
M05	NVIDIA	How has NVIDIA's data center revenue grown over the past two years?	Data Center revenue grew **65%** year-over-ye ar for the nine-month period ended Oct 26, 2025, reaching $131.423	CORR…	Verified p67 (PDF scan 2026-05-03)

			billion.		
M07	Goldm…	What is Goldman Sachs's CET1 capital ratio and how does it compare to their target?	CET1 capital ratio was **15.6%** (Standardized) and **20.9%** (Advanced), exceeding the regulatory minimum of 4.5% and requirement of 13.6%. Internal target not disclosed.	CORR…	Verified p174/p232 (PDF scan 2026-05-03)
M08	Eli Lilly	What were the main drivers of Eli Lilly's revenue growth in their latest annual report?	Primarily **Mounjaro** ($22,965M, 99% growth) and **Zepbound** ($13,542M, 175% growth). Overall revenue grew 45%.	CORR…	Verified p76 (PDF scan 2026-05-03)
M10	Intel	What were the reported impairments or write-downs disclosed by Intel in their latest filings?	**$950 million** in asset impairments and accelerated depreciation in 2025, plus $460 million in non-cash charges.	CORR…	Verified p32/p34 (PDF scan 2026-05-03)

H01	Google…	Compare the R&D spending as a percentage of revenue between Google and Microsoft in their latest filings.	**Google** invests a higher percentage at **15%** of revenue vs. **Microsoft** at **11.5%** of revenue.	CORR…	Verified p40/p64 (PDF scan 2026-05-03)
H02	Semico…	Summarize how semiconduct or companies are discussing AI-related demand in their latest earnings reports.	Reporting strong current AI-driven demand (e.g., AMD 36% YOY increase), but emphasizing **significant uncertainty** about long-term AI demand.	CORR…	Verified p131 (PDF scan 2026-05-03)
H04	Multi-C…	Among Nvidia, Microsoft, Apple, Google, Tesla, and Walmart — which shows the strongest financial health based on their latest filings?	**Microsoft** and **NVIDIA** demonstrate the strongest financial health (high operating income/cash, low debt).	CORR…	Verified p76/p77 (PDF scan 2026-05-03)

H05	Cloud …	Compare cloud segment revenue growth rates: Amazon AWS, Microsoft Intelligent Cloud, and Google Cloud over the last reported year.	**Google Cloud** (29%), **Amazon AWS** (20%), and **Microsoft Intelligent Cloud** (19-21%).	CORR…	Verified p49 (PDF scan 2026-05-03)
H06	Big Tech	What are the most significant legal or regulatory risks disclosed across the big tech companies in their latest 10-K filings?	**Antitrust enforcement actions** (e.g., EC fines), **data privacy obligations** (GDPR), and litigation related to platform policies.	CORR…	Verified p49/p63 (PDF scan 2026-05-03)
H07	Visa/M…	Compare operating margins between Visa and Mastercard over the last two fiscal years.	**Visa's** margin declined (65.7% to 60%); **Mastercard's** margin improved (55.3% to 57.6%).	CORR…	Verified p24/p70 (PDF scan 2026-05-03)

H08	Energy…	How are energy companies XOM and CVX discussing the energy transition and capital allocation in their latest annual reports?	**XOM** commits to net-zero in Permian by 2035 and GHG reduction by 2030, investing in Low Carbon Solutions. **CVX** takes a more cautious approach.	CORR…	XOM p71, CVX p40 (PDF scan 2026-05-03)
H10	Palantir	Give me an overview of recent performance of Palantir, including revenue growth, margins, and key business developments.	**2025 revenue: $4.475B** (56% YOY growth); **Gross margin: 82%** (up from 80%); key segment is Palantir Cloud (subscriptions).	CORR…	Verified p95 (PDF scan 2026-05-03)