Replicating Technical Specification Documents using Retrieval–Augmented Generation (RAG)

Name: Yoshita Banerjee
Guide: Prof. Senthilmurugan Subbiah
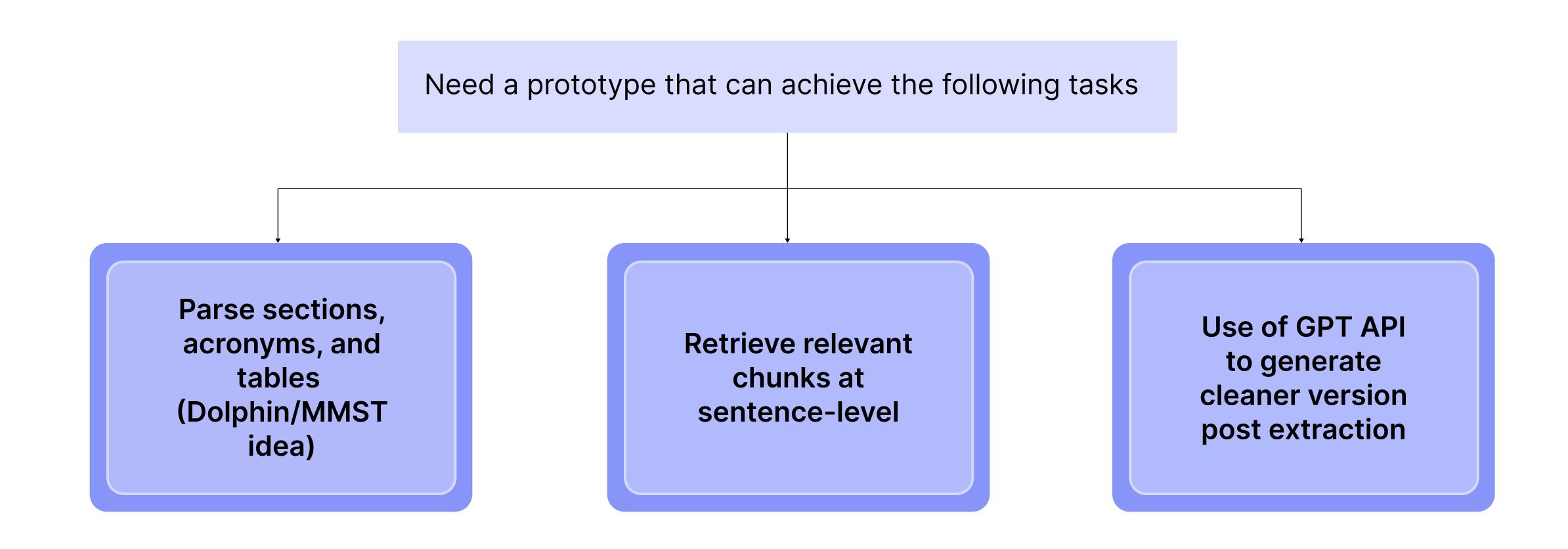*Dept: Chemical Engineering, BTP 1*

Problem Statement

→ Technical specification documents are dense, multimodal (texts, tables, figures), and domain specific

→ Large language models (LLMs) often hallucinate or misinterpret acronyms/definitions

→ Requirement: A pipeline that can reliably retrieve, structure and regenerate technical documentation

Literature Review

- **Vanilla RAG** *(Lewis et al. 2020)* : Retrieval + Generation, but weak for domain specific docs

- **RAG for technical documents** *(Soman and Roychowdhury, 2024)* : Found that sentence-level retrieval and splitting term-definition pairs work best

- **MA-RAG** *(Nguyen et al., 2025)*: Multi-agent approach (Planner, Step-Definer, Extractor, QA Writer) for handling ambiguities

- **SAMAC-R3-MED** *(Li et al., 2025)*: Multi-modal semantic trees (MMST) + knowledge graphs for engineering documents

Gaps to be solved

Existing systems do not focus on replicating structured specification documents

Need a prototype that can achieve the following tasks

Parse sections, acronyms, and tables (Dolphin/MMST idea)

Retrieve relevant chunks at sentence-level

Use of GPT API to generate cleaner version post extraction

Proposed Solutions

Workflow for BTP-1

1. Parser- Extract headers, paragraphs, acronyms, tables

2. Retriever: Sentence-level retrieval (FAISS, SBERT embeddings)

3. Generator: GPT API returns regenerated section

4. Evaluation: Use simple metrics → cosine similarity, BLEU/ROUGE scores

Deliverable - *A demo where upon giving an input query or section → model regenerates a spec-style output with inline citation*

RETRIEVAL QUALITY	Evaluation Metrics	GENERATION QUALITY
• Cosine similarity • Recall@k		• BLEU • ROUGE

Scope of BTP-2

→ Incorporate multi-agent MA-RAG workflow (Planner, Step-Definer, Extractor, QA Writer)

→ Handle tables and figures using layout-aware parsing (Dolphin, MMST)

→ Improve evaluation with expert feedback loops and better semantic metrics

Conclusion

- BTP-1: Build a demonstrable RAG prototype for specification document replication

- BTP-2: Extend to multi-agent + multimodal retrieval for robust work

References

RAG systems for Technical Documents

Multi-Agent RAG

Multi-Agent collaboration of R3 models for multi-modal documents

Dolphin (Document Image Parsing via Heterogenous Anchor Prompting