



Retrieval-Augmented Generation Web Service

Enhancing code documentation with company-specific knowledge

https://github.com/vassily60/Web_Service_RAG - Vassily Lombard

Current Issue

Writing high-quality, consistent technical documents is time-consuming

Normal LLM outputs → too generic, not company-specific

Need: A system that generates context-aware, structured documentation

Solution



RAG Pipeline

- LLM
- Company Knowledge Base



Integration

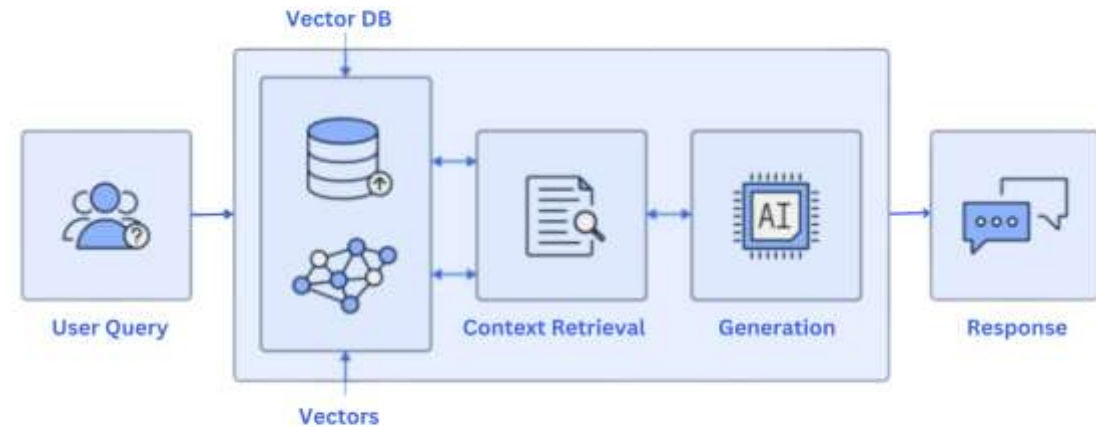
- Adapt answer (style + structure)
- Context Specific Output



Generation

- Higher Relevance
- Faster Generation

RAG Architecture



Key Features



Ask questions or request technical docs (e.g., “*API architecture doc*”)



Search by keywords or synonyms

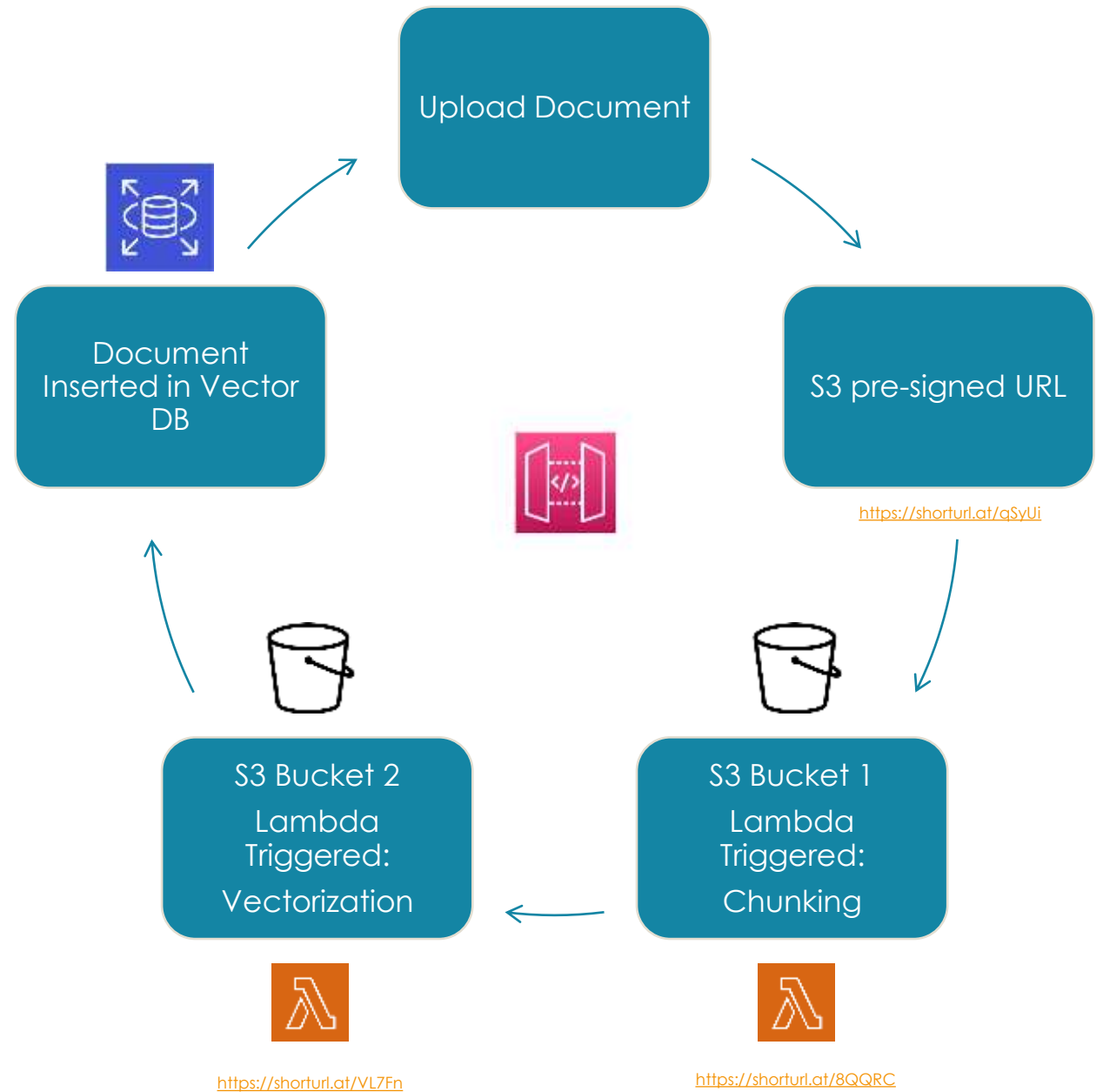


Tracks request cost (tokens + latency)



Output: **faster, consistent, context-specific docs**

Technical Workflow Doc Upload



Impact & Value

Saves time writing technical docs

Provides structured, consistent outputs

Improves knowledge sharing across teams

Scalable for different document types and industries

https://github.com/vassily60/Web_Service_RAG