**RAGBot Project Analysis & Recommended Structure**

**Current Implementation Status**

|  |  |  |
| --- | --- | --- |
| Component | Status | Files Present |
| RAG Engine - Document Ingestion | **COMPLETE** | document\_loader.py, vectorstore\_builder.py, ingest.py |
| RAG Engine - Query Processing | **COMPLETE** | query\_processor.py, rag\_chain\_builder.py, vectorstore\_loader.py, query.py |
| LLM Integration | **COMPLETE** | llm\_loader.py |
| Vector Database | **COMPLETE** | ChromaDB setup with chroma.sqlite3 |
| Web Search Integration | **COMPLETE** | tavily\_tool.py |
| Query Expansion | **COMPLETE** | query\_expander.py |
| Logging & Utils | **COMPLETE** | logger.py, eval\_logs.json |
| Sample Data | **COMPLETE** | Multiple test documents (txt, pdf, docx) |
| Basic Testing | **COMPLETE** | Test retrieval functionality |

**Current Staus  
What's Done (60% Complete):**

* Core RAG engine with document ingestion and query processing
* ChromaDB vector database integration
* LLM integration with query expansion
* Web search via Tavily API
* Logging and utilities
* Sample data and basic testing

**ToDo Status**

**Current Implementation Status**

|  |  |  |  |
| --- | --- | --- | --- |
| **Component** | **Status** | **Files Present** | **Missing/TODO** |
| **FastAPI Backend** | **MISSING** | **None** | **Complete API server implementation** |
| **PyQt Desktop UI** | **MISSING** | **Empty ui folder** | **Complete desktop application** |
| **Authentication System** | **MISSING** | **None** | **Google OAuth integration** |
| **Cloud Services Integration** | **MISSING** | **None** | **AWS RDS, S3, ELB setup** |
| **Cython Core Module** | **MISSING** | **None** | **Performance-critical business logic** |
| **App Packaging** | **MISSING** | **None** | **PyInstaller + MSIX/DMG setup** |
| **Configuration Management** | **PARTIAL** | **.env file** | **Complete config system** |
| **Database Models** | **MISSING** | **None** | **SQLAlchemy models for user data** |
| **API Documentation** | **MISSING** | **None** | **OpenAPI/Swagger docs** |
| **Docker Setup** | **MISSING** | **None** | **Containerization for deployment** |
| **CI/CD Pipeline** | **MISSING** | **None** | **GitHub Actions or similar** |

**What's Missing (35%):**

* FastAPI backend API server
* PyQt desktop application
* Google OAuth authentication
* AWS cloud services integration
* Cython performance modules
* App packaging and distribution