


Development plan

From litun nayak <litunnayak09@hotmail.com>

Date Mon 4/7/2025 8:37 AM

To Litun Nayak <nayaklitun9@gmail.com>; DIBYAKANTA NAYAK <dibyakantanayak2@gmail.com>

PHASE 1 – CORE SYSTEM (MVP)

 Goal: Let users upload any doc (PDF, Word, Excel, etc.) and **chat with it** using GPT-4-Turbo.

1. FEATURE SCOPE (MVP)

- Upload & store any doc (PDF, DOCX, PPTX, XLSX, TXT, Scanned PDFs, Images)
 - Extract + chunk content
 - Embed and store in Vector DB
 - Chat with GPT-4 using RAG (retrieval augmented generation)
 - Simple API / UI for file upload and asking questions
-

2. USER WORKFLOW (PHASE 1)

Step 1: Upload

- User uploads file (drag & drop)
- Supported: PDF, DOCX, XLSX, PPTX, TXT, JPG/PNG, Scanned PDF

Step 2: Auto-processing

- File → Text Extractor → Clean text
- Image/Scanned file → OCR → Clean text

Step 3: Chunk + Embed

- Text is broken into chunks
- Each chunk is embedded and stored in a **vector DB** with file metadata

Step 4: Ask Questions

- User types a question
 - Your system:
 - Embeds the question
 - Finds most relevant text chunks from Vector DB
 - Passes them + the question to GPT-4-Turbo
 - Response is returned, optionally with citations
-

⚙️ 3. TECH STACK (PHASE 1)

Component	Stack
Backend	Python (FastAPI)
Text Extraction	PyMuPDF, python-docx, openpyxl, python-pptx
OCR	Tesseract, EasyOCR
Vector DB	FAISS or ChromaDB (local), Weaviate/Qdrant later
Embeddings	text-embedding-3-small (OpenAI)
LLM	gpt-4-turbo with vision (for OCR results or raw images if needed)
Frontend (later)	Next.js + Tailwind (Phase 2)
File Storage	Local for now, Wasabi/Bunny/Cloudflare R2 later

📦 PHASE 2 – FULL APP + UI

💡 Goal: Beautiful web app + better workflow integration

🔥 Features

- Full Frontend UI (dashboard, chat, upload manager)
 - Multiple file support in one session
 - Save chat history
 - Tag/search uploaded files
 - Highlight referenced sections
 - Export summary as PDF/Word
-

👤 USER WORKFLOW (PHASE 2)

Step 1: Create a workspace (optional)

- Think Notion-style: “My Docs”, “Client A Docs”

Step 2: Upload multiple files

- Multiple formats at once

Step 3: Ask questions like:

“What are the main points across all 3 docs?”
“Summarize section 3 of this doc.”
“What’s the invoice total in this Excel sheet?”

Step 4: Share or export answers


- Save chats
 - Export as PDF/Markdown
 - Share link or email it
-

💻 TECH ADDITIONS

Feature	Stack
Frontend UI	Next.js + TailwindCSS + Redux or Zustand

User Auth	Clerk/Auth0 or custom JWT
Storage	Wasabi + Bunny CDN for delivery
Background jobs	Celery + Redis
Chat UI	shadcn/ui, ChatGPT-style interface
Websockets	For live progress updates on file parsing
OCR Viewer	Display images + text overlay for source clarity

PHASE 3 – ADVANCED AI FEATURES

 Goal: Killer features that make your product viral-worthy

AI POWER-UPS

- “Summarize this document” button (in 1 click)
- Entity extraction (names, amounts, deadlines)
- Smart search across documents
- Extract tables from PDFs/images
- Visual chart summarizer (with chart-to-text)

Integrations

- Google Docs + Notion: Add-on for “Ask AI” on any doc
 - Gmail Add-on: “Chat with attached doc”
 - Slack Bot: Drop in a file → chat with it
 - Zapier Integration
-