

# Notebook-style RAG System

*AI-Powered Document Intelligence for Learning & Research*

Youcheng Taing, Freeman Yiu

# Executive Summary



## Business Concept

AI-powered notebook for interacting with uploaded documents through intelligent Q&A with source citations



## Target Market

Students, researchers, educators, and knowledge workers seeking efficient document analysis



## Proposed Solution

RAG-powered system with vector search, accurate citations, and conversational interface



## Team Structure

Full-stack development with specialized backend and AI systems expertise

# Target Market



## Students

High school and university students seeking study assistance



## Researchers

Graduate students and researchers analyzing academic papers



## Self-Learners

Online course participants and independent learners



## Educators

Teachers and teaching assistants managing course materials



## Knowledge Workers

Professionals handling technical documentation

# Solution Overview



## Document Upload

Support for PDFs and various learning materials



## Vector Embeddings

Convert documents into searchable semantic representations



## RAG Technology

Retrieval-Augmented Generation for accurate, grounded answers



## Source Citations

Transparent references to original document sources



## Notebook Interface

Conversational, intuitive user experience

# Team Structure



## Team Member 1

### Full-Stack & Backend Lead

- Backend development using FastAPI
- Database design (PostgreSQL, SQLAlchemy, Alembic)
- API design and authentication
- System integration



## Team Member 2

### Frontend & AI Systems Lead

- User interface development (React)
- Retrieval-Augmented Generation implementation
- Vector database integration
- AI model optimization

# Technology Stack

## Backend

- FastAPI (Python web framework)
- PostgreSQL (Database)
- SQLAlchemy & Alembic (ORM & migrations)
- Authentication & authorization

## Frontend

- React (JavaScript/TypeScript)
- Modern UI components
- Responsive design
- Real-time updates

## AI & ML Infrastructure

- Large Language Models for natural language understanding
- Vector databases for semantic search
- RAG pipeline for context-aware responses
- Document processing and embedding generation

# Key Benefits

1

## Accuracy

Grounded answers based on actual document content

2

## Transparency

Source citations for every response

3

## Efficiency

Rapid document analysis and information retrieval

4

## Scalability

Handle multiple documents and large knowledge bases

5

## User Experience

Intuitive notebook-style conversational interface

6

## Trust

Verifiable information with clear source attribution

# Ready to Transform Document Intelligence

*Let's discuss how we can bring this vision to life*

Questions & Discussion