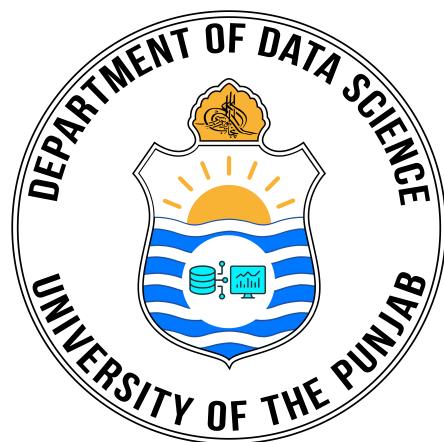


Final Year Project Proposal

# NoteScape: AI-Powered Personal Learning Workspace



*By*

Syeda Rabbiya Bukhari    BSDF22A044  
Hajra Amir                    BSDF22A027  
Lubaba Khalid                BSDF22A025  
Mahnum Zahid                BSDF22A022

*Under the supervision of*

**Dr. Syed Faisal Bukhari**

*Bachelor of Science in Data Science (2022-2026)*

**FACULTY OF COMPUTING & INFORMATION TECHNOLOGY  
(FCIT),**

**UNIVERSITY OF THE PUNJAB, LAHORE.**

**NoteScape: AI-Powered Personal Learning Workspace**

A project proposal presented to

**University of the Punjab, Lahore**

In partial fulfillment of the requirement for the degree of

*Bachelor of Science in Data Science (2022-2026)*

By

Syeda Rabbiya Bukhari BSDF22A044

Hajra Amir BSDF22A027

Lubaba Khalid BSDF22A025

Mahnum Zahid BSDF22A022

**FACULTY OF COMPUTING INFORMATION TECHNOLOGY (FCIT),**

**UNIVERSITY OF THE PUNJAB, LAHORE**

## **Executive Summary**

Students commonly rely on passive study methods—like repeatedly reading PDFs, lecture notes, or slides—and often turn to quick fixes at exam time rather than structured revision. NoteScape addresses this by providing a single AI-powered workspace where students can effortlessly upload their existing study materials, including PDFs, slide decks, code snippets, and scanned handwritten notes. The uploaded materials are automatically digitized and organized, with an optional OCR module under active research to digitize handwritten notes, subject to research accuracy. The workspace allows fast, intuitive retrieval of content, preserving the original layout and context (including diagrams and formulas) digitally through OCR-based processing.

An integrated AI chat provides instant summaries and Q&A, while the platform automatically generates flashcards and quizzes tailored directly from the uploaded materials. A built-in spaced repetition system schedules reviews effectively, aiming to reduce the manual effort involved in organizing study materials and helping students actively engage with their coursework. All files, study sessions, chat interactions, and performance data are centrally stored to ensure students retain easy, uninterrupted access.

NoteScape's primary goal is to significantly streamline the study workflow, promote active learning methods, and thereby improve student retention and efficiency. Optional features such as audio playback of summaries and full-session PDF exports may be included if time permits once core objectives have been completed. The system utilizes a modern technology stack, including React/Next.js for the front end, FastAPI microservices for the back end, PostgreSQL for persistent storage, and OCR/NLP tools, embedding proven educational methods directly into students' personal study materials.

# Contents

<b>Contents</b>	<b>4</b>
<b>List of figures</b>	<b>5</b>
<b>List of tables</b>	<b>6</b>
<b>1 INTRODUCTION</b>	<b>7</b>
1.1 Background . . . . .	7
1.2 Problem Statement . . . . .	8
1.3 Project Goals Objectives . . . . .	9
1.4 Scope . . . . .	10
1.5 High-level System Components . . . . .	10
1.5.1 Web Application . . . . .	10
1.5.2 Backend Services . . . . .	11
1.6 List of optional functional units . . . . .	11
1.7 Application Architecture . . . . .	11
1.8 System Limitations and Constraints . . . . .	12
1.8.1 Limitations . . . . .	12
1.8.2 Constraints . . . . .	13
1.9 Tools and Technologies . . . . .	14
<b>2 Related Literature Review</b>	<b>15</b>
2.1 Related Work . . . . .	15
2.2 Gap Analysis . . . . .	15
<b>References</b>	<b>16</b>

# List of figures

1.1 Application Architecture . . . . .	12
--	----

# List of tables

1.1 Tools and Technologies Used . . . . .	14
---	----

# Chapter 1

## INTRODUCTION

### 1.1 Background

Effective education requires active engagement and deep learning, but many students fall into passive habits such as rereading lecture slides, PDFs, or handwritten notes without truly interacting with the material. This often leads to poor retention, inefficient study routines, and last-minute cramming. Additionally, some study materials are dense, text-heavy, and difficult to understand, which can make learning feel overwhelming, dull, and stressful.

Several educational tools address parts of these challenges. Platforms like **Anki** and **Brainscape** offer flashcard-based learning with confidence-based spaced repetition to improve retention. However, both require students to manually create or find flashcard decks, which can be time-consuming and discouraging. **AlgorEducation** automates flashcard and quiz generation from uploaded content but produces static, non-editable materials and imposes costly subscription limits that restrict accessibility.

AI-powered tools such as **Google’s NotebookLM** provide intelligent summarization and context-aware question answering but focus more on general document interaction rather than structured study workflows. Despite their strengths, none of these solutions fully integrate the variety of study materials students use—including PDFs, slides, handwritten notes, and code snippets—into a unified, personalized learning workspace.

NoteScape is designed to fill this gap. It offers a single AI-powered workspace where students organize their study materials by class or subject, preserving the full context of each item. The platform digitizes content using OCR (with ongoing research to improve handwriting recognition), maintaining visual elements like diagrams and formulas. It delivers AI-powered chat for instant summaries and clarifications, automatically generates editable flashcards and quizzes tailored to each student’s materials, and supports confidence-based spaced repetition. Importantly, NoteScape also provides personalized revision planning to help students prepare strategically for exams and quizzes.

Beyond improving retention and reducing manual effort, NoteScape aims to make learning more engaging and less stressful by transforming dense, text-heavy content into digestible, interactive formats. By keeping all study resources, AI interactions, quizzes, and progress organized in one persistent workspace, NoteScape encourages consistent deep learning and

helps students manage their study workload effectively, ultimately supporting better academic outcomes.

## 1.2 Problem Statement

Many students struggle with inefficient and ineffective study habits that limit their academic performance. The core problem is that students often rely on passive learning techniques—such as rereading lecture notes, PDFs, and slides—withou engaging actively with the material. This approach leads to poor retention, frequent procrastination, and last-minute cramming, which increase stress and reduce learning outcomes.

Despite the availability of various digital study aids like flashcard apps (Anki, Brainscape) and AI summarization tools (NotebookLM, AlgorEducation), these solutions fall short of addressing students' comprehensive needs. They either require manual preparation, lack support for diverse file types (including handwritten notes), or generate static, non-editable study content. Moreover, most do not offer personalized revision planning or maintain learning context across sessions, which are essential for sustained engagement and mastery.

The unmet need is for a unified, intelligent platform that effortlessly transforms all kinds of study materials into active, personalized learning resources, automates revision scheduling, and preserves context over time. This solution should significantly reduce manual effort, encourage consistent deep learning, and make studying less stressful and more effective.

### **Significance of the Problem:**

Studies show that over 70% of students report using passive study methods as their primary learning approach, contributing to lower retention rates and increased exam anxiety. Research also indicates that well-implemented active learning strategies, including spaced repetition and self-testing, can improve retention by up to 50%. However, manual application of these strategies is time-consuming and often abandoned without automated support.

### **Target Customers and Market Size:**

#### **1. Undergraduate and Graduate Students:**

- Estimated 25 million students in Pakistan and neighboring countries in STEM and humanities fields who regularly use digital and handwritten materials.
- Global higher education market estimated at over 220 million post-secondary students.

#### **2. Lifelong Learners and Professionals:**

- Millions engaged in continuous professional development, online courses, and certifications requiring effective study tools.

- The global e-learning market is projected to surpass USD 400 billion by 2026, with rising demand for AI-powered personalized learning solutions.

**3. Educational Institutions and Tutors:** Schools, universities, and private tutors seeking tools to support students' independent learning and improve outcomes.

NoteScape aims to serve these customer segments by providing an accessible, all-in-one study workspace that addresses the key challenges of engagement, retention, and efficient study management.

### 1.3 Project Goals Objectives

The main goal of NoteScape is to develop an AI-powered workspace that transforms students' study materials into active, personalized learning resources, helping them study more effectively with less manual effort.

The specific objectives are:

- **Reduce Manual Effort:** Automate the digitization and organization of diverse study materials, minimizing the time students spend formatting and preparing notes.
- **Enhance Study Efficiency:** Generate editable flashcards, quizzes, and personalized revision schedules automatically to promote active learning and spaced repetition.
- **Improve Learning Engagement:** Provide an AI chat interface for instant Q&A and summaries, making complex and text-heavy materials easier to understand and less stressful.
- **Maintain Contextual Continuity:** Organize all study content, interactions, and progress by class or subject within a single persistent workspace, preserving learning context across sessions.
- **Offer Flexible, Scalable Features:** Design optional components such as audio playback and session exports that can be added based on available time and resources.

By achieving these objectives, NoteScape aims to save students' time, reduce study-related stress, and foster deeper, more effective learning habits.

## 1.4 Scope

NoteScape focuses on creating an AI-powered workspace that enables students to upload and organize diverse study materials—including PDFs, slide decks, handwritten notes, and code snippets—into a unified, interactive learning environment. The system will digitize content using OCR technology, generate editable flashcards and quizzes, provide AI-assisted chat for on-demand Q&A and summaries, and implement a confidence-based spaced repetition system with personalized revision planning. All study materials, interactions, and progress will be organized by class or subject for easy access and continuity.

### **Limitations:**

- The project targets individual users rather than multi-user collaboration or instructor dashboards.
- Handwriting OCR is an ongoing research component; its accuracy may vary depending on input quality.
- Advanced features such as audio playback of summaries and full-session PDF export are optional and will only be developed if time permits after core functionalities are complete.
- The system will not initially support integration with third-party learning management systems or social platforms.
- Performance with very large documents (100+ pages) may experience processing delays.

## 1.5 High-level System Components

The system comprises the following modules, grouped by type:

### 1.5.1 Web Application

- **Login Module:** Authenticates users, manages secure access, and supports password recovery via email.
- **Dashboard Module:** Provides an overview of classes, upcoming revisions, and quick access to study materials.
- **File Upload Module:** Allows uploading of PDFs, slide decks, handwritten notes, and code snippets with batch enhancements like auto-straightening and reordering.
- **Class Workspace Module:** Organizes all materials, AI chat logs, quizzes, and revision schedules by subject or class for seamless access.

- **AI Chat Interface:** Enables interactive, context-aware question answering and summaries tied to each class workspace.

### 1.5.2 Backend Services

- **Ingestion and OCR Service:** Processes uploaded files with OCR for printed text and ongoing research-based handwriting recognition; preserves layouts including diagrams and formulas.
- **Semantic Indexing Service:** Generates vector embeddings from digitized content and manages semantic search for fast, context-aware retrieval.
- **Content Generation Service:** Utilizes large language models (e.g., GPT-4) to produce editable flashcards, quizzes, and study summaries.
- **Revision Planning Service:** Implements confidence-based spaced repetition algorithms and generates personalized revision schedules aligned with exams and quizzes.
- **Storage Service:** Manages persistent storage of files, metadata, chat transcripts, quiz results, and mastery data using PostgreSQL and object storage.

## 1.6 List of optional functional units

- Audio playback of summaries and quizzes using text-to-speech technology.
- Export of complete study sessions as consolidated PDF packets for offline review.
- Authentication via third-party providers such as Google or Gmail.

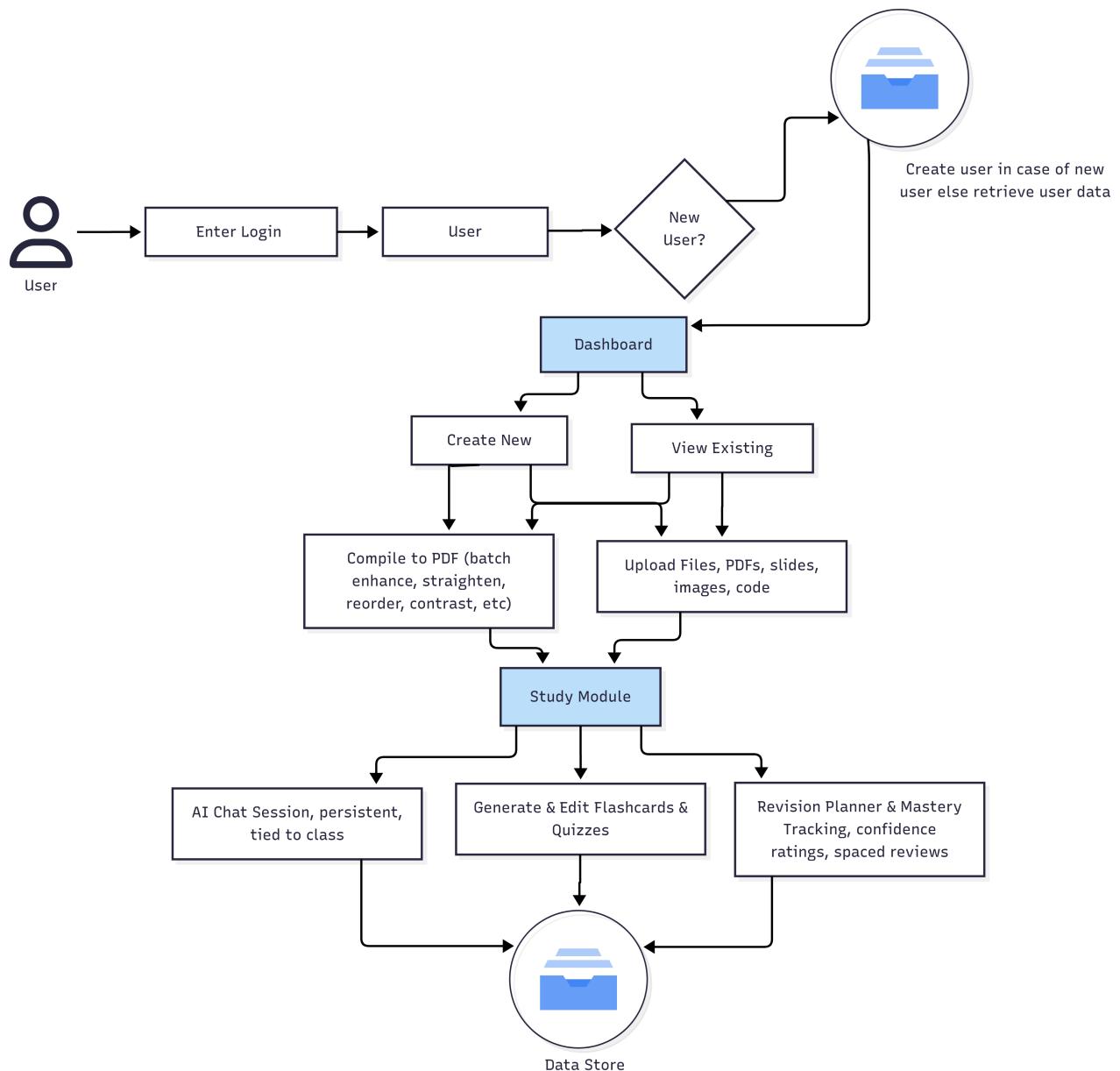
## 1.7 Application Architecture

NoteScape uses a three-tier architecture consisting of the Presentation Layer, Business Logic Layer, and Data Storage Layer.

The Presentation Layer is the user interface built with React/Next.js, enabling users to upload materials, access study modules, and interact with AI chat and quizzes.

The Business Logic Layer runs FastAPI microservices that handle file processing, OCR, AI content generation, semantic search, and revision scheduling.

The Data Storage Layer uses PostgreSQL and object storage for user data, study materials, and session information, ensuring reliable and efficient data management.



**Fig. 1.1.** Application Architecture

## 1.8 System Limitations and Constraints

### 1.8.1 Limitations

- OCR accuracy for handwritten notes may vary depending on handwriting quality and scan resolution, potentially affecting digitization quality.
- Processing large documents (e.g., 100+ pages) may introduce latency, impacting user experience during file ingestion.
- Real-time AI chat responses rely on internet connectivity; interruptions may delay interactions.

- Automated flashcard and quiz generation quality depends on the underlying language model and may require user edits for accuracy.
- Optional features like audio playback and PDF export are subject to time constraints and may not be implemented in the initial release.

### 1.8.2 Constraints

- The project scope focuses on individual student use and excludes multi-user collaboration or instructor analytics dashboards.
- Development is limited to a one-year timeline (Q3 2024 to Q3 2025) with a small team size (single developer or up to four members).
- Hardware and GPU resources for handwriting OCR and AI processing are constrained by available infrastructure.
- Integration with external learning management systems (LMS) or third-party platforms is out of scope for this project phase.

## 1.9 Tools and Technologies

**Table 1.1.** Tools and Technologies Used

Tool/Technology	Description and Version
Front-end framework for dynamic user interfaces	React framework for server-side rendering and routing (v18.2, Next.js v13)
High-performance backend API framework in Python	FastAPI (v0.115)
Relational database system for structured data storage	PostgreSQL (v16)
Storage for uploaded files and binary data	Object Storage (MinIO or AWS S3)
Vector database for semantic search and embeddings	Qdrant (v1.10)
Open-source OCR engine for printed text	Tesseract OCR (v4.1)
Deep learning-based OCR for handwriting (under research)	Custom PyTorch Model
Large Language Model integration for chat, summarization, and flashcard generation	OpenAI GPT-4 API
Library for generating semantic embeddings for search	Sentence-Transformers
Python ORM for database interaction	SQLAlchemy (v2.0)
Containerization platform for deployment and scalability	Docker
Primary programming language	Python (v3.11)

# Chapter 2

## Related Literature Review

### 2.1 Related Work

Several AI-powered educational platforms and tools have emerged to enhance learning efficiency by automating study aids and improving content accessibility. For example, Anki and Brainscape implement confidence-based spaced repetition to improve retention, but require manual flashcard creation, which can be time-consuming and limits scalability [1], [2].

Google’s NotebookLM offers AI-driven document summarization and context-aware question answering, providing a more interactive way to engage with study materials; however, it primarily focuses on general document interaction rather than structured study workflows [3].

AlgorEducation automates quiz and flashcard generation from uploaded content but produces static, non-editable resources and has restrictive subscription limits, reducing its accessibility and flexibility [4].

While these platforms contribute valuable features, none fully integrate diverse study materials—including PDFs, slide decks, handwritten notes, and code—into a unified, personalized AI workspace that supports editable study aids, persistent context-aware interaction, and strategic revision planning.

### 2.2 Gap Analysis

Current research and commercial tools often address isolated aspects of the active learning process but fall short of providing an end-to-end solution. Manual flashcard platforms lack automation and contextual integration, while AI summarization tools do not support revision scheduling or editable learning content [5], [6]. Moreover, handwriting digitization remains an ongoing challenge, with few systems effectively integrating handwriting OCR into comprehensive learning workflows [7].

NoteScape aims to bridge these gaps by combining OCR-based digitization (including handwriting under active research), AI-driven content generation, editable flashcard and quiz creation, confidence-based spaced repetition, and persistent class-specific study workspaces. This integration addresses the unmet need for an all-in-one, adaptive learning environment that streamlines student effort, preserves context, and promotes deep, active learning.

# References

- [1] D. Elmes, *Anki user manual*, <https://apps.ankiweb.net>, Accessed: 2025-07-11, 2023 (cited on p. 15).
- [2] Brainscape, *Brainscape flashcard platform*, <https://www.brainscape.com>, Accessed: 2025-07-11, 2024 (cited on p. 15).
- [3] G. Research, *Notebooklm: Ai-powered study assistant*, <https://blog.google/technology/ai/notebooklm/>, Accessed: 2025-07-11, 2024 (cited on p. 15).
- [4] AlgorEducation, *Algoreducation flashcards and quiz generator*, <https://www.algoreducation.com>, Accessed: 2025-07-11, 2024 (cited on p. 15).
- [5] J. Smith and R. Lee, “A review of active learning technologies in education,” *Journal of Educational Technology*, vol. 15, no. 3, pp. 200–215, 2022 (cited on p. 15).
- [6] X. Chen and R. Kumar, “Ai summarization tools and their role in modern education,” *International Journal of AI in Education*, vol. 9, no. 1, pp. 45–59, 2023 (cited on p. 15).
- [7] L. Wang and A. Patel, “Advances in handwriting ocr for educational applications,” in *Proceedings of the International Conference on Document Analysis*, 2021, pp. 122–130 (cited on p. 15).