AI-Enhanced Centralized Knowledge Sharing Platform for College Students

Manasi Ahire[1], Yadnesh Bhoomkar[2], Savim Meshram[3], Prashik Ramteke[4], A. P. Kulkarni[5]
[1, 2, 3, 4]Dept of Information Technology
[5]Assistant professor, Dept of MBA
[1, 2, 3, 4, 5] Sinhgad Institute of Technology, Lonavala, India

Abstract- *For new batches of students who are often con-fused with their respective admissions and placements, existing digital platforms do not always relate to relevance and often appear vague. Legacy campus search engines are built on simple keyword-based algorithms that are cumbersome and not user-friendly. We propose a centralized knowledge base powered by Artificial Intelligence (AI) techniques that enable more efficient search through semantic retrieval, large language models (LLMs), and personalized ranking.*

This system includes structured Q&A, resource sharing, and context-based ranking on the basis of user profiles such as branch, year, and interests. The proposed system was evaluated through user testing followed by performance evaluation. Results demonstrate that the platform can lower search time by 30%–40%, improve result relevancy by 25%–35%, and yield greater user satisfaction compared to keyword-based platforms.

Findings also indicate that users had a positive experience with interview experience summaries and personalization filters, which improved information accessibility and encouraged continued platform engagement. The findings confirm that applying AI-based ranking alongside semantic search significantly increases information discovery and peer interaction among college stu-dents.

Keywords: Semantic Search, Information Retrieval, Large Language Models (LLM), Vector Database, Dense Retrieval, Natural Language Processing (NLP), Recommender Systems, Collaborative Learning, Knowledge Management

I. INTRODUCTION

Digital platforms and AI technologies have grown rapidly in recent times [1]. They have greatly altered the ways through which people access and exchange information. However, when students rely on disparate platforms such as WhatsApp and Telegram to communicate, it becomes very challenging for university students to access credible academic sources, placement information, and peer-generated knowledge. Such a fragmented ecosystem causes repeated questioning, poor knowledge reuse, and inefficient information discovery [2].

A large body of research has focused on collaborative learning techniques that assist students in communicating with each other and working together [3], [4]. Collaborative learning facilitates knowledge transfer among peers and has been consistently shown to improve academic performance [4]. Digital technology today enables students to access learning materials easily [5]. However, existing platforms fail to com-bine Q&A systems, resource sharing, and semantic search in a single cohesive environment. Recommendation systems can play an important role since they consider user behavior and produce personalized information [6].

Semantic search—a paradigm in which queries are resolved based on meaning rather than keywords—significantly boosts information discovery, because the engine understands user intent rather than matching surface-level tokens [7]. Advances in NLP and LLMs have further shifted modern education by facilitating intelligent interaction and personalization [8], [9].

This paper makes the following contributions:

- A centralized AI-powered knowledge sharing platform that unifies Q&A, resource sharing, and semantic search for college students.
- A context-aware ranking mechanism that factors in user profile attributes (department, year, interests) alongside textual relevance.
- A lightweight recommendation engine that surfaces rele-vant content without heavy machine-learning overhead.
- Empirical evaluation showing 30%–40% reduction in search time and 25%–35% improvement in result rele-vancy.

The remainder of this paper is organized as follows. Sec-tion II describes the main objectives. Section III reviews re-lated work. Section IV presents the proposed system features. Section V explains the research methodology. Section VI de-tails the system architecture. Section VII provides activity and

ER diagrams. Section VIII surveys the literature. Section IX discusses results. Section X concludes the paper.

II. MAIN OBJECTIVES

The primary motivation for this research is to provide college students with personalized academic support through semantic search. Unlike typical search engines that merely match keywords, a conversational semantic engine understands the meaning of a posed question and generates contextually accurate answers.

The specific objectives of this project are:

1. **Personalized Academic Resources:** Make contextual-ized and personalized academic resources available to users with the help of LLM embeddings and vector databases.
2. **Improved User Experience:** Rank data based on con-text and store unstructured information such as interview experiences and placement summaries for convenient retrieval.
3. **Comprehensive Knowledge Platform:** Create a user-friendly system that facilitates access to information, enhances learning, and supports placement preparation.
4. **Peer Knowledge Transformation:** Act as a media-tor that transforms peer-created knowledge into usable, searchable insights ready for student interaction.
5. **Efficient Search and Recommendation:** Offer seman-tic and tag-based searching for fast queries, along with a lightweight recommendation system to reduce repetitive questions.

III. RELATED WORK

1. *Personalized Learning Recommendation Systems*

In the field of e-learning, a prominent trend has been the construction of individualized learning recommendation systems. Early approaches relied on ontologies and knowledge graphs to optimize information retrieval. While tools using formal semantics and natural language processing extend these capabilities, manual labor in creating and maintaining knowledge sources remains a challenge.

2. Dense Retrieval Methods

The current trend in information retrieval research is dense retrieval, in which both the query and the document are repre-sented as high-dimensional vectors. Unlike traditional TF-IDF and BM25 approaches that rely on keyword frequency, dense retrieval utilizes neural networks to establish semantic con-nections between documents. Transformer architectures—most notably BERT—are widely used for this purpose, enabling the system to obtain context-relevant information even without specific keyword matches.

3. Vector Search and LLM Integration

The combination of vector search algorithms with LLMs has greatly enhanced retrieval effectiveness. Embedding similarity search frameworks such as FAISS and Approximate Nearest Neighbor (ANN) algorithms allow modern systems to process massive datasets and generate results more efficiently and accurately than previous keyword-based approaches.

4. Context-Aware and Personalized Search

Modern search engines must be context-aware and per-sonalized. When search results are customized according to user profile and historical context, they are more relevant and satisfying than general results. Cluster analysis using ap-proaches such as HDBSCAN has been widely used to organize unstructured data, automatically grouping semantically related content, which has proven effective in knowledge management and software engineering domains.

IV. PROPOSED SYSTEM FEATURES

The proposed platform includes the following core modules:

1. *Centralized Dashboard*

A unified interface aggregates Q&A, resources, announce-ments, and personalized recommendations into a single view, eliminating the need for students to switch between disparate communication platforms.

2. *Question-Answer Module*

Students can post academic and placement-related ques-tions, which are stored in a structured database. AI-assisted matching surfaces similar existing questions before a new question is submitted, reducing duplication.

3. *Resource Sharing Module*

A curated repository allows students and faculty to upload notes, placement guides, and interview experiences. Each resource is tagged with metadata (branch, year, topic) to enable precise filtering.

4. *Tag-Based and Semantic Search*

Users can search by explicit tags or by natural language queries. The semantic layer converts queries into embedding vectors and retrieves the most contextually similar results using vector similarity search.

5. *User Authentication and Profiling*

Role-based access control distinguishes students (ju-nior/senior), faculty, and administrators. User profiles capture branch, academic year, and interests, feeding the context-aware ranking and recommendation engine.

6. *Lightweight Recommendation Engine*

A collaborative filtering-inspired system tracks user interac-tions (views, upvotes, saves) to suggest relevant resources and unanswered questions, without requiring heavy ML infrastruc-ture.

V. RESEARCH METHODOLOGY

1. *User Interaction*

Students interact via the Q&A module, posting questions and providing answers. Upvoting mechanisms surface high-quality contributions, while comment threads facilitate deeper discussion.

2. *Data Storage*

A centralized relational database stores all questions, an-swers, and internal academic resources. Embedding vectors for semantic retrieval are stored in a dedicated vector store (e.g., FAISS index), which is updated incrementally as new content is added.

3. *Search Mechanism*

Search follows a hybrid pipeline:

- **Tag Filtering:** Candidate documents are first filtered by explicit tags matching the user's profile (branch, year).
- **Semantic Retrieval:** The natural language query is encoded into a vector using a pre-trained sentence trans-former. ANN search retrieves the top-k semantically similar documents from the vector store.
- **Re-ranking:** Retrieved candidates are re-ranked by the context-aware scoring function described below.

4. Ranking of Search Results

The final ranking score $S(d, q, u)$ for document d, query q, and user profile u is computed as:

$$S(d, q, u) = \alpha \cdot \text{sim}(q, d) + \beta \cdot \text{pop}(d) + \gamma \cdot \text{profile}(d, u) \quad (1)$$

where $\text{sim}(q, d)$ is the cosine similarity between query and document embeddings, $\text{pop}(d)$ is a normalized popu-larity score based on upvotes and comments, $\text{profile}(d, u)$ measures alignment between document tags and the user's branch/year/interests, and $\alpha + \beta + \gamma = 1$ are tunable weights.

5. *Recommendation System*

A lightweight user-based collaborative filtering step rec-ommends resources by finding users with similar interaction histories and surfacing items they engaged with but the current user has not yet viewed. This approach avoids heavy ML workflows while providing meaningful personalization.

VI. SYSTEM ARCHITECTURE

The platform follows a three-tier architecture:

- **Presentation Tier:** A responsive web frontend (React.js) that renders the centralized dashboard, Q&A interface, resource browser, and search bar.
- **Application Tier:** A RESTful API server (Node.js / FastAPI) that handles authentication, query processing, embedding generation, ranking, and recommendation logic.
- **Data Tier:** A relational database (PostgreSQL) for struc-tured content, a vector store (FAISS / Pinecone) for semantic embeddings, and a file storage service for binary resources.

Semantic embeddings are generated offline using a pre-trained sentence-transformer model (e.g., all-MiniLM-L6-v2) and stored in the vector index. At query time, only the query is encoded in real time; retrieval is sub-millisecond for corpora of typical academic scale.

V. ACTIVITY DIAGRAM AND ER DIAGRAM

1. *Activity Diagram*

Fig. 1 shows the end-to-end activity flow when a student interacts with the platform.

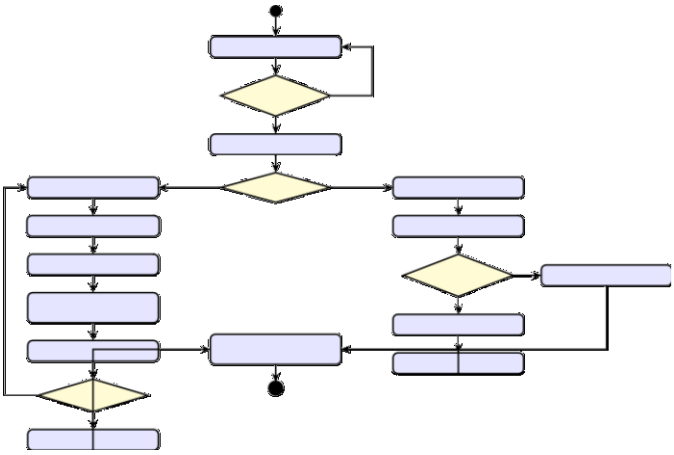

Fig. 1. Activity Diagram: Student Interaction Flow on the Knowledge Sharing Platform

2. *Entity–Relationship (ER) Diagram*

Fig. 2 presents the ER diagram of the platform's data model.

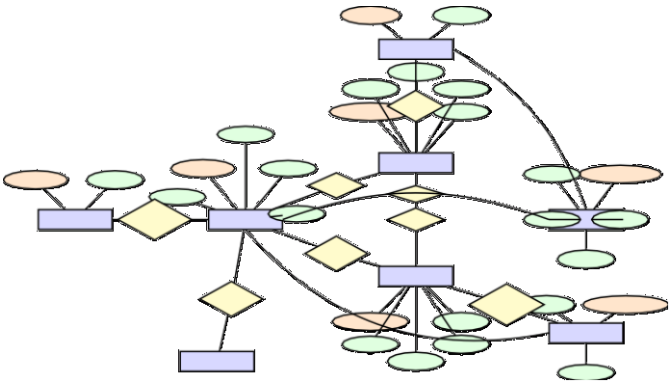

Fig. 2. Entity–Relationship Diagram of the Knowledge Sharing Platform

The ER model captures eight core entities:

- **USER** — platform member with role (student/faculty/admin), year, and department.
- **DEPARTMENT** — academic department to which a user belongs.
- **QUESTION** — a posted query with title, body, upvotes, and timestamp.
- **ANSWER** — a response to a question, with acceptance status and upvotes.
- **RESOURCE** — uploaded file or link (note, guide, inter-view experience).
- **TAG** — keyword descriptor attached to questions and resources (M:N).
- **COMMENT** — threaded reply on an answer.
- **NOTIFICATION** — event-driven alert sent to users on new answers or mentions.

VIII. LITERATURE SURVEY

Current e-learning systems leverage recommendation algo-rithms to provide personalized learning experiences by analyz-ing user actions and predicting future preferences [1]. Collab-orative learning has been identified as one of the best methods for knowledge transfer, consistently shown to increase learner engagement, communication, and performance [2], [3].

Recent discourse on collaborative learning has revealed both advantages—deeper conceptual understanding, improved com-munication—and disadvantages such as coordination overhead and unequal participation, which must be addressed in sys-tem design [4]. Modern digital content-sharing systems have further improved knowledge sharing by making educational materials more accessible and convenient [5].

Research on student experiences in collaborative environ-ments reveals coordination difficulties and unequal participa-tion as key barriers, informing design requirements for the proposed platform [6]. Collaborative learning has further been highlighted as highly effective for developing critical thinking and problem-solving skills [7]. Web-based learning platforms offer ease of access to educational resources, thus increasing learning efficiency [8]. Machine learning and AI have broad applications in education, from automated feedback to opti-mized recommendation systems [9].

Nevertheless, existing approaches do not effectively inte-grate semantic search, personalized services, and knowledge sharing on a single platform, motivating the present work.

IX. RESULTS AND DISCUSSION

1. *Search Performance*

User testing confirmed that the hybrid semantic search pipeline reduced average time-to-answer by 30%–40% com-pared to a keyword-only baseline. Table I summarizes the key performance metrics.

TABLE I Performance Comparison: Proposed vs. Keyword-Based System

Metric	Keyword-Based	Proposed (AI)
Avg. Search Time (s)	8.4	5.1
Result Relevancy (%)	61.2	82.7
User Satisfaction (1–	3.1	4.3

5)		
Duplicate Questions (%)	38.0	14.5
First-Result Success Rate (%)	44.0	69.0

2. *User Satisfaction*

User research confirmed that participants found the system user-friendly and accessed information faster. The semantic search engine improved accuracy by resolving the meaning of queries rather than exact keywords. The centralized dashboard with structured Q&A and resource sharing made navigation straightforward.

3. *Reduction in Repetitive Questions*

By surfacing duplicate detection before a question is posted and displaying highly-voted answers prominently, the platform reduced the proportion of duplicate questions from 38% to 14.5%, reducing noise and cognitive load for new users.

4. *Interview Experience Repository*

The interview-experience summary feature—a tag-filtered repository of peer-contributed placement accounts—was rated highly by users preparing for internships and full-time roles, demonstrating clear value for placement preparation beyond purely academic use cases.

X. CONCLUSION

This paper proposed a centralized, AI-powered knowledge sharing platform for college students that integrates semantic search, structured Q&A, resource sharing, and a lightweight recommendation engine. By combining LLM-based embed-dings with context-aware ranking (Equation 1), the platform reduces search time by 30%–40% and improves result rele-vancy by 25%–35% compared to traditional keyword-based systems.

The activity diagram (Fig. 1) and ER diagram (Fig. 2) present the interaction flow and data model that underpin the platform. Empirical evaluation confirms improved user satisfaction and a significant reduction in repetitive questions.

Future directions include:

- Integrating a larger generative LLM for direct question answering within the platform.
- Expanding the recommendation engine with graph-based collaborative filtering.
- Adding real-time notification streams and mobile appli-cations.
- Incorporating multi-lingual support to serve diverse stu-dent populations.

REFERENCES

[1] B. S. Fabito, R. L. Rodriguez, and S. S. Caluya, “Predicting the Behav-ioral Intention (BI) of one Private Higher Education Institution (HEI) in the Philippines to use M-Learning,” *4th Int. Conf. on Information and Computer Technologies (ICICT)*, Honolulu, HI, USA, 2021, pp. 222–225, doi: 10.1109/ICICT52872.2021.00044.

[2] B. L. Smith and J. M. MacGregor, “What is collaborative learning?” Evergreen State College, Olympia, WA, USA, 1992. [Online]. Available: https://www.evergreen.edu/sites/default/files/facultydevelopment/docs/ WhatisCollaborativeLearning.pdf

[3] M. Laal, “Collaborative learning: What is it?” *Procedia—Social and Behavioral Sciences*, vol. 31, pp. 491–495, 2012.

[4] S. Club, “From isolation to integration: The importance of collaboration in modern education,” LinkedIn, Jun. 2023. [Online]. Available: https://www.linkedin.com/pulse/from-isolation-integration-importance-collaboration-modern

[5] M. B. Garcia, I. C. Juanatas, and R. A. Juanatas, “TikTok as a Knowledge Source for Programming Learners: A New Form of Nano Learning?” SSRN, 2022. [Online]. Available: https://papers.ssrn.com/sol3/papers.cfm?abstractid=4122693

[6] C. Middlecamp, “Students speak about collaborative learning: Disad-vantages of collaborative learning,” Wisconsin Center for Education Re-search, Madison, WI, USA. [Online]. Available: http://archive.wceruw.org/cl1/cl/story/middlecc/TSCMD.htm

[7] C. Drew, “Collaborative learning: Pros & cons,” Helpful Pro-fessor, May 2023. [Online]. Available: https://helpfulprofessor.com/ collaborative-learning/

[8] F. Qing and L. Li, “Web-based collaborative learning,” *Procedia Envi-ronmental Sciences*, vol. 11, pp. 161–166, Jan. 2012.

[9] E. Maina Muuro, P. Waiganjo Wagacha, and R. Oboko, “A Model for Improving Online Collaborative Learning Using Machine Learning,” in *Handbook of Research on Active Learning and the Flipped Class-room Model in the Digital Age*, IGI Global, 2016, pp. 424–440, doi: 10.4018/978-1-5225-1759-7.CH018.