



Database-Driven Architecture for Academic Knowledge Modeling and Automated Practice Paper Generation

Haji Saad¹, Shaikh Mohammed Abbas², Tiwari Aditya³, Naik Faiz⁴, Shimpi Manasi⁵

Computer Engineering, Anjuman – I – Islam’s Abdul Razzak Kalsekar Polytechnic, Navi Mumbai, India¹⁻⁴

Lecturer, Computer Engineering, Anjuman – I – Islam’s Abdul Razzak Kalsekar Polytechnic, Navi Mumbai, India⁵

Abstract: Students in diploma-level technical education often find it difficult to access academic content that is properly organized according to exam requirements. Most existing tools like study planners and question banks work separately and do not combine syllabus structure, previous year questions, and exam patterns in one place. This paper proposes a database architecture designed for managing academic content and supporting exam preparation. The system organizes syllabus elements such as subjects, units, topics, and intended learning outcomes using a structured relational model. It also includes previous year questions that are mapped to syllabus units and classified using Bloom’s taxonomy. In addition, the system represents exam patterns using slot-based structures and blueprint constraints. This helps in retrieving structured data and allows automatic generation of practice papers. Overall, the design shows how a well-organized database can support better academic understanding, question mapping, and exam focused preparation.

Keywords: Database-driven architecture, Academic knowledge modeling, Practice paper generation, Bloom’s taxonomy, Question mapping, Exam pattern modeling

I. INTRODUCTION

Students in diploma-level technical education often struggle to access and use academic resources in a properly organized way. Important data like syllabus content, previous year questions, and exam patterns are usually scattered, making it hard to understand their connections and use them effectively for exam preparation.

A major limitation of current approaches is the lack of a unified data model that represents academic knowledge in a structured and connected form. While some systems provide study materials or question banks, they do not bring together syllabus hierarchy, learning outcomes, and exam patterns in one place. Because of this, it becomes difficult to analyze important topics, understand question trends, or generate exam focused practice content.

To solve this problem, this paper proposes a relational database architecture for managing academic knowledge. The system models syllabus components such as subjects, units, topics, and intended learning outcomes, and connects them with previous year questions and exam patterns. Questions are further improved by mapping them to syllabus elements and classifying them using Bloom’s taxonomy.

The main goal of this work is to show how a well-designed database schema can support efficient storage, structured retrieval, and meaningful use of academic data. The proposed architecture can be used for applications like intelligent question retrieval and practice paper generation, while keeping the focus on database design and data modelling.

II. LITERATURE REVIEW

A. AI-Based Learning Systems

The AI-based learning system offers personalized suggestions to students according to their performance and behavioural responses. Although this system improves flexibility to a greater extent, it does not consider personalization and lacks an integrated structure with a syllabus hierarchy and examination-oriented data.

B. Question Paper Generation Systems

The current question paper generation system uses rule-based and AI-based techniques to develop examination question papers. Although these techniques consider examination-related constraints like marks and difficulty levels, they lack an integrated structure with a syllabus hierarchy.



C. Bloom's Taxonomy-Based Systems

Bloom's taxonomy-based systems classify questions into cognitive levels such as remembering, understanding, and applying. Many approaches use keyword-based methods, which may not fully capture context.

The proposed system also adopts a keyword-based mapping using predefined verbs, as it is simple, efficient, and suitable for structured diploma-level question patterns.

D. NLP-Based Question Processing

Recent advancements in question paper generation and classification are based on natural language processing techniques. Although these techniques are highly effective in processing questions and generating examination question papers, they lack an integrated structure with previous year questions and academic frameworks.

The proposed system integrates syllabus hierarchy, question intelligence, and examination patterns within a single database structure to provide a structured and examination-oriented academic analysis system.

III. SYSTEM OVERVIEW

The proposed system is an integrated academic support system with the aim of assisting diploma students with their academic preparation. The system is designed to be structured and exam-oriented.

A. System Overview

The system is built around the concept of providing a study planner, question bank, question mapping engine, practice paper generator, and student interface as a single system with a centralized database.

B. System Components

The system is comprised of the following modules:

- Personalized Study Planner
- Question Bank Module
- Question Mapping Engine
- Practice Paper Generator
- Student Interface

C. System Workflow

The system starts with student input parameters, followed by the generation of the study roadmap and interaction with the academic database to retrieve the questions and enrich the data through the question mapping engine to generate structured practice papers as per the examination pattern.

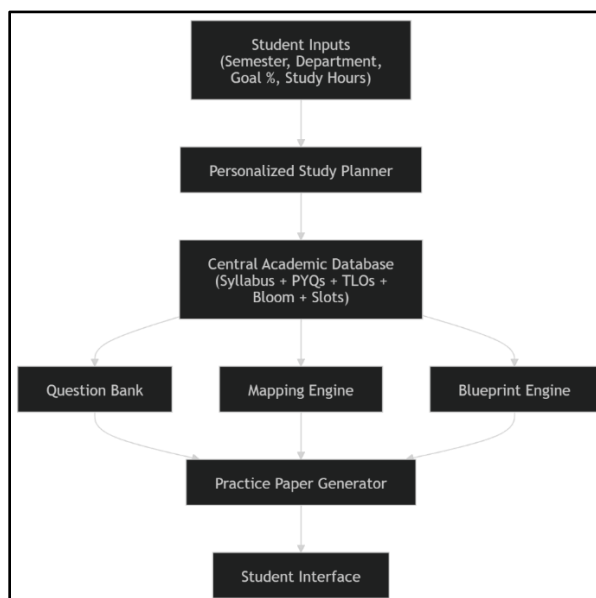


Fig. 1 Overall System Architecture of the Proposed Academic Support Platform



IV. DATABASE ARCHITECTURE

A. Database Design Goals

The database is created in a manner that it helps in the management of structured data in academics, thereby allowing intelligent operations while being constraint-driven. The specific goals of the database are as follows:

- Representation of Hierarchical Syllabus Structure
- Incorporating Question Intelligence in Academic Units
- Representation of Examination Patterns/Slot Structures
- Blueprint-Driven Paper Generation Support
- Scalability of Multi-Subject/Multi-Scheme Support

B. Core Entity Groups

The database is logically divided into separate groups of entities to promote modularity and clarity:

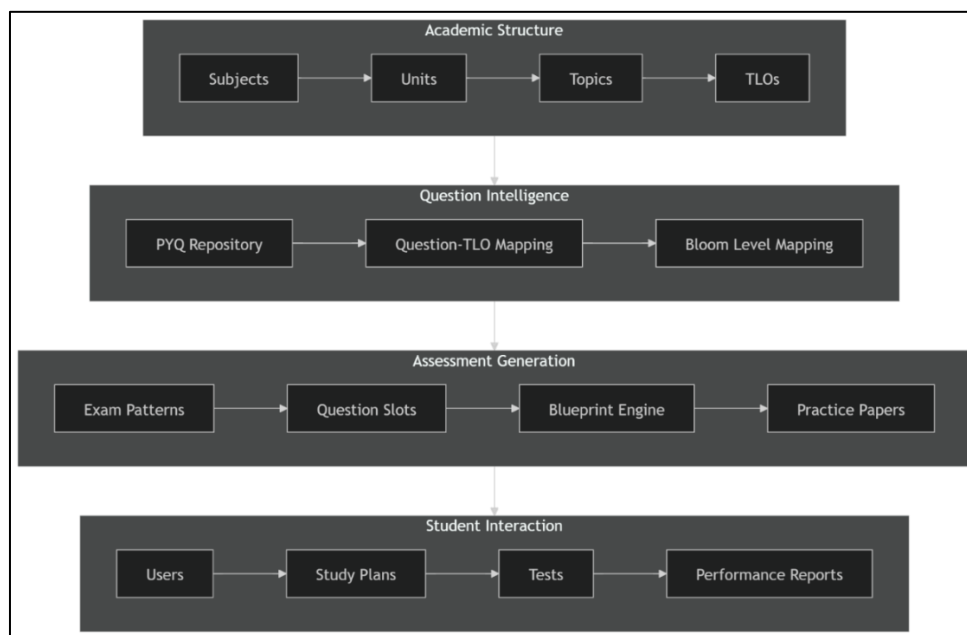


Fig. 2 Layered Database Architecture of the Proposed Academic Knowledge System

1) Academic Structure Layer:

This layer represents the syllabus structure, i.e., the hierarchy of subjects, units, topics, and learning outcomes.

2) Question Intelligence Layer:

This layer consists of the database for the previous year's questions, along with their mapping to the syllabus and Bloom's level classification.

3) Assessment and Exam Modeling Layer:

This layer consists of the structure for the examination, i.e., the exam patterns, slots, blueprints, and generated papers.

4) Student Interaction Layer:

This layer consists of the user-related data, i.e., the study plan and the student's performance.

Contrary to the conventional approaches that treat the generation of an examination as a procedural process, the proposed architecture models the structure of the examination in the database using slot representations, historical analysis, and blueprint constraints. This allows the database to become an active component in the decision-making process rather than a passive component.

C. Conceptual Data Model

The system attempts to model complex relationships between academic content, questions, and structure of examinations through an interconnected relational schema as follows:



- Subjects are broken down into Units and LOs
- Questions are mapped to Units and TLOs based on similarity
- Questions are classified based on Bloom's taxonomy
- Exam patterns represent structural constraints of exams
- Blueprints represent slot-level constraints of exams

This unified representation allows for efficient querying, searching, and generation of examination papers based on constraints.

D. Schema Design Rationale

The database has a normalized relational structure to ensure data consistency and prevent data redundancy. The use of mapping tables, such as the `question_tlo_map`, facilitates flexible many-to-many relationships. The separation of academic data, question intelligence, and examination modeling promotes modularity. The use of JSON fields for storing blueprint and slot analysis facilitates flexible data representation without compromising relational data integrity.

E. Key Tables and Relationships

TABLE I CORE DATABASE TABLES AND THEIR FUNCTIONAL ROLES

Table Name	Role	Key Purpose
subjects	Academic structure	Stores subject details
units	Academic structure	Defines unit-level breakdown
tlos	Academic mapping	Stores learning outcomes
pyq	Question repository	Stores previous year questions
question_tlo_map	Mapping	Links questions to units & TLOs
pyq_bloom_mapping	Classification	Stores Bloom levels
exam_patterns	Exam modelling	Defines paper structure
slots	Exam structure	Defines question positions
blueprints	Generation rules	Stores paper constraints
practice_papers	Output	Stores generated papers

F. Database Workflow

The database also supports a multi-stage workflow for transforming raw academic data into structured knowledge:

- Syllabus data is structured into hierarchical entities
- Previous year questions are stored and enriched
- Questions are mapped onto units and learning outcomes
- Questions are classified using Bloom's taxonomy
- Examination structures are modeled using slots and patterns
- Blueprints are generated through historical analysis
- Practice papers are constructed through constraint-based selection

The above workflow shows the database's capability in facilitating intelligent academic operations.

V. QUESTION MAPPING AND PROCESSING

The system enriches the raw data of the questions through a series of processes of structured mapping and classification. The questions are mapped to relevant syllabus units and learning outcomes through similarity-based mapping techniques.

The mapping of the questions is performed by assigning status values such as AUTO, REVIEW, and NO_MATCH, thereby making it more reliable and traceable. Further, the questions are classified according to Bloom's taxonomy through verb-based analysis.

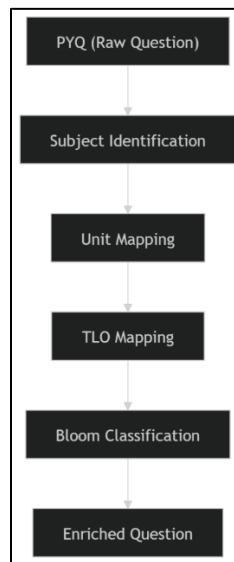


Fig. 3 Question Processing and Enrichment Pipeline

This enriches the raw data of the questions into a structured academic knowledge base, making it possible to retrieve them exactly while generating a balanced practice paper.

VI. PRACTICE PAPER GENERATION

The system uses a constraint-driven approach to generate practice papers with the help of a database architecture.

The examination patterns describe the structure and mark distribution of the paper. The slots describe a particular position in the paper with a set of marks allocated to that slot. The historical slot analysis determines the probabilistic distribution of units and Bloom levels for a particular slot based on previous year question papers.

The blueprint combines examination patterns and curriculum constraints to specify unit-wise and cognitive level requirements for a generated paper.

To generate a paper, questions are selected from the database according to mapping, classification, and slot constraints. This selection occurs through a weighted and constraint-driven strategy.

The above explanation illustrates the examination structure modeling within a database and its role in generating examination papers systematically and according to a particular pattern. This separates the proposed system from other conventional rule-based approaches.

VII. CONCLUSION

This research introduces a database-centric method for academic data modeling and exam preparation. The proposed system combines the concepts of hierarchy in the syllabus, question intelligence, and examination patterns in a single relational framework.

This research aims to show how the proposed architecture can help databases transcend the traditional boundaries of data storage and play a more active role in supporting the academic decision-making process. This proposed system provides a foundation for the development of a new generation of academic support systems and intelligent learning tools.

REFERENCES

- [1]. Thing Thing Goh, Nor Azliana Akmal Jamaludin, Hassan Mohamed, and Mohd Nazri Ismail, "Semantic Similarity Analysis for Examination Questions Classification Using WordNet", *Appl. Sci.* 2023, 13, 8323.
- [2]. M. Azuddin, M. Yusup, H. Setiyowati, S. Wibowo, J. Suwita, S. Basuki and E. D. Astuti, "Design and Implementation of a Relational Database for an Academic Information System," *International Transactions on Artificial Intelligence (ITALIC)*, vol. 3, no. 2, pp. 171–180, May 2025.



- [3]. H. F. El-Sofany, N. Al-Jaidah, S. Ibrahim and S. Al-Kubaisi, "Web-based 'Questions-Bank' System to Improve E-Learning Education in Qatari School," Journal of Computer Science, vol. 5, no. 2, pp. 97–108, 2009.
- [4]. Tatiana A. Kustitskaya, Roman V. Esin, Alexey A. Kytmanov, and Tatiana V. Zykova, "Designing an Education Database in a Higher Education Institution for the Data-Driven Management of the Educational Process", Educ. Sci. 2023, 13, 947.
- [5]. Y. Song, "Design and Implementation of an Online Exam Database System," Bachelor's Thesis, Lappeenranta-Lahti University of Technology (LUT), Finland, 2025.
- [6]. Maharashtra State Board of Technical Education (MSBTE), "MSBTE Official Website," [Online]. Available: <https://msbte.org.in>. [Accessed: 2026].