COLLEGE OF ENGINEERING TRIVANDRUM

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

CSD 334
V-LIBRA
Software Requirement Specification

Ashwin Jacob Gigo	TVE20CS030
Christin Sibi	TVE20CS128
Don Joshi N	TVE20CS044
Gokul G Menon	TVE20CS046

Table of Contents

1. Introduction

1.1. Purpose

The purpose of this project is to design and build a searchable vector database for a college library that can store and retrieve information about books, journals, and other resources with high accuracy on descriptive search queries. The system will be designed to improve the search experience for users and make it easier to find what they are looking for.
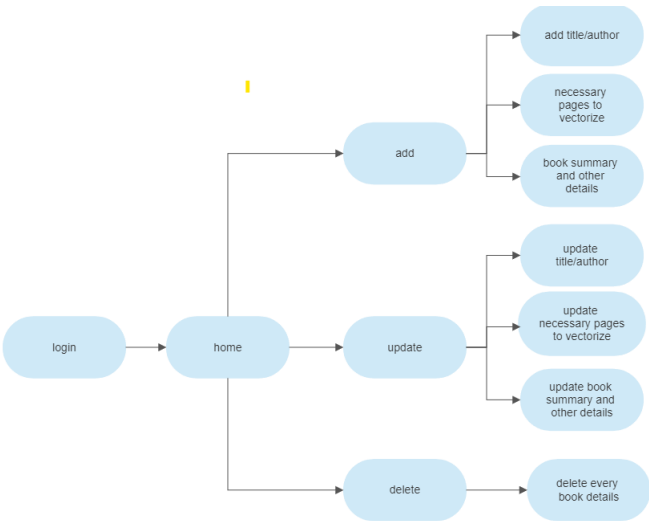
1.2. Intended Audience

The intended audience for this project is students, faculty, and staff at the college who will use the library to find resources for their research and academic work. The system will also be used by librarians and other library staff to manage the collection and keep it up-to-date.

1.3. Project Scope

The project scope includes designing and building a database that can store information about books, journals, and other resources, as well as providing a user-friendly interface for searching and retrieving information from the database with higher accuracy. The system will be built using vector space models to enable efficient searching based on various criteria such as descriptions about information to be obtained in addition to author name, title, subject, and keywords. The system can also be engineered to provide personalised recommendations based on a user's search history and preferences. The project will not include the physical infrastructure for the library, such as shelving and furniture, or the management of borrowing and lending of physical items.

Admin Layout

```
                                                    ┌──────────────────┐
                                              ┌────▶│  add title/author │
                                              │     └──────────────────┘
                                              │     ┌──────────────────┐
                          ┌────────┐          ├────▶│  necessary        │
                     ┌───▶│  add   │──────────┤     │  pages to         │
                     │    └────────┘          │     │  vectorize        │
                     │                        │     └──────────────────┘
                     │                        │     ┌──────────────────┐
                     │                        └────▶│  book summary     │
                     │                              │  and other        │
                     │                              │  details          │
                     │                              └──────────────────┘
                     │                              ┌──────────────────┐
                     │                        ┌────▶│  update           │
                     │                        │     │  title/author     │
┌────────┐   ┌────────┐    ┌────────┐         │     └──────────────────┘
│  login │──▶│  home  │──▶ │ update │─────────┤     ┌──────────────────┐
└────────┘   └────────┘    └────────┘         ├────▶│  update           │
                     │                        │     │  necessary pages  │
                     │                        │     │  to vectorize     │
                     │                        │     └──────────────────┘
                     │                        │     ┌──────────────────┐
                     │                        └────▶│  update book      │
                     │                              │  summary and      │
                     │                              │  other details    │
                     │                              └──────────────────┘
                     │    ┌────────┐                ┌──────────────────┐
                     └───▶│ delete │───────────────▶│  delete every     │
                          └────────┘                │  book details     │
                                                    └──────────────────┘
```

User Layout

```
┌────────┐        ┌────────┐         ┌──────────────────┐
│  login │───────▶│ search │────────▶│  view result as   │
└────────┘        └────────┘         │  per genre        │
    │                  ▲             └──────────────────┘
    │                  │                      │
    ▼                  │                      ▼
┌────────────┐         │             ┌──────────────────┐
│ view home  │◀────────┤             │    pick book     │
└────────────┘         │             └──────────────────┘
    │   ▲              │                      │
    ▼   │              │                      ▼
┌────────────┐         │             ┌──────────────────┐
│   pick     │─────────┘             │  feedback to      │
│ recomended │                       │  improve vector   │
│   book     │                       └──────────────────┘
└────────────┘
```
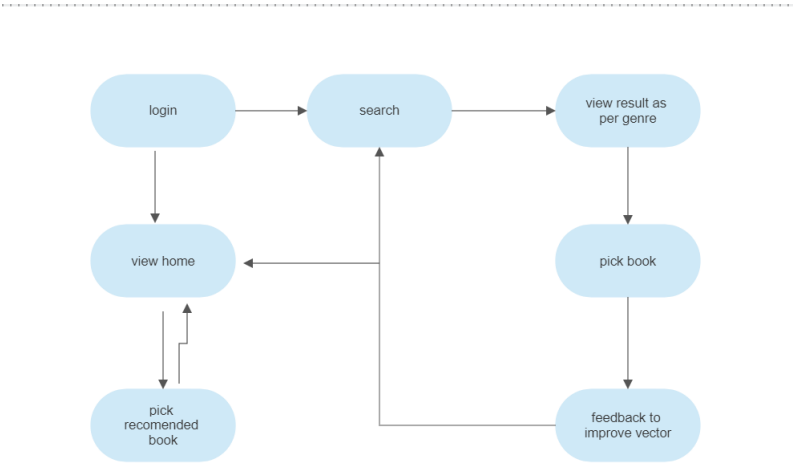
2. Overall Description

2.1. Product Perspective

The searchable vector database for the college library will be a standalone system that will be integrated into the existing library infrastructure. It will act as a complementary tool to the traditional library catalogue and provide an enhanced search experience for users. The database will be designed to be scalable and adaptable, allowing for future expansion and integration with other library systems or third-party applications.

2.2. User Class and Characteristics

Administrators:

- Privileged access to the system

- Responsible for managing the database, including adding, deleting, and updating entries

- Manage user accounts and access controls

Users:

- Students, faculty, and staff of the college

- Utilise the system to search for library resources

- Basic computer literacy and familiarity with online search interfaces expected

- Search based on descriptive queries of content in addition to other criterias such as author name, title, subject, and keywords

- View search results with relevance ranking

- Access relevant recommendations based on search history and preferences.

2.3. Database Design
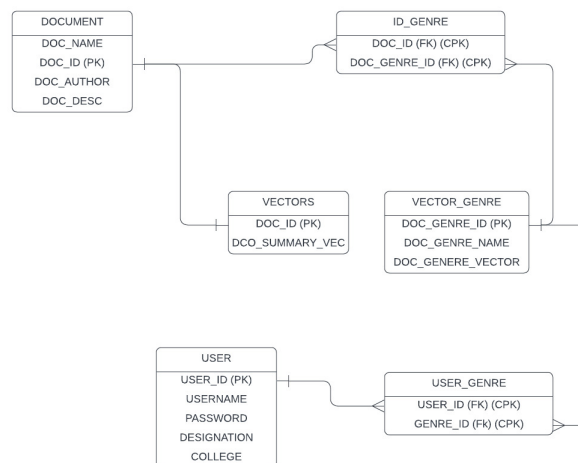
2.4. Operating Environment

The application will be operated as an active webpage

2.5. Design and Implementation Constraints

The design of the system will follow a user-centred approach, focusing on providing a simple and intuitive user interface for both administrators and users. The user interface will be designed to allow users to easily search for resources based on various criteria, view search results with relevance ranking, access relevant recommendations based on their search history and preferences and provide relevant feedback to improve performance . The admin interface will provide easy-to-use functionalities for managing the database, including adding, deleting, and updating entries, as well as managing user accounts and access controls. The system will also incorporate error handling mechanisms and robust data validation to ensure data integrity and system reliability. The design will be modular and extensible, allowing for future enhancements and modifications as needed. In conclusion, the searchable vector database for the college library will be a scalable and adaptable system that aims to enhance the information retrieval capabilities of the library. With a user-friendly interface, efficient database design, an d suitable operating environment, the system will provide faster search times, more accurate results, and increased flexibility in handling complex queries, benefiting both administrators and users alike.

3. System Features

3.1. Description and Priority

The priority of the system is to build a highly efficient and user-friendly searchable vector database for the college library. The system will be designed to store and retrieve information about books, journals, and other resources available in the library, using vector space models (VSMs) for document representation and cosine similarity measure for document similarity calculation. The system will allow users to search for resources based on various criteria such as author name, title, subject, and keywords. On the admin side, options will be provided to add, delete, and update entries in the database as new books arrive and old ones are removed. The system will also include features such as including only certain pages of entries to be vectorized if necessary, and providing relevant recommendations to users based on their search history and preferences.

- Accuracy and Relevance: The system should provide accurate search results based on user queries and rank them based on relevance to the user's query. The system should also provide relevant recommendations to users based on their search history and preferences.

- Efficiency: The system should be designed to provide fast search times and minimize processing overhead in order to provide a smooth and efficient user experience.

- Flexibility: The system should be flexible and able to handle complex queries involving multiple criteria, providing users with a versatile and powerful search functionality.

- User-Friendliness: The system should be designed with a user-friendly interface, making it easy for users to search for resources and perform administrative tasks such as adding, deleting, and updating entries in the database.

- Security: The system should implement appropriate security measures, including user authentication and access controls, to protect the confidentiality and integrity of the library's data.

- Scalability: The system should be scalable to accommodate future growth in the library's collection and user base, without compromising performance or functionality.

- Maintainability: The system should be designed in a modular and maintainable manner, making it easy to update and enhance in the future as needed. Proper documentation and commenting should be provided for ease of maintenance by the IT team.

3.2. Functional Requirements

- Search Functionality: The system should allow users to search for resources such as books, journals, and other materials based on various criteria such as author name, title, subject, and keywords. The search results should be ranked based on relevance to the user's query and should provide accurate and relevant results.

- Document Representation: The system should use vector space models (VSMs) to represent documents as vectors in a high-dimensional space, with each dimension corresponding to a term in the document. The system should implement cosine similarity measure to calculate the similarity between two documents based on the angle between their corresponding vectors.

- User Authentication and Access Controls: The system should implement user authentication to ensure that only authorized users have access to the database. Different user roles such as admin and regular users may be defined with appropriate access controls to perform actions such as adding, deleting, and updating entries in the database.

- Administrative Functionality: The system should provide options for the admin to add, delete, and update entries in the database as new resources arrive and old ones are removed. The system should also allow the admin to include only certain pages of entries to be vectorized if deemed necessary.

- User History and Recommendations: The system should track users' search history and preferences to provide relevant recommendations for resources based on their past searches. This feature should be designed to enhance user experience and provide personalized recommendations.

- Advanced Search Queries: The system should be able to handle complex queries involving multiple criteria, allowing users to perform advanced searches based on combinations of author name, title, subject, keywords, and other relevant attributes.

- Error Handling: The system should implement appropriate error handling mechanisms to handle invalid queries, input validation, and other potential errors, providing meaningful error messages to users when necessary.

- Reporting and Analytics: The system may include reporting and analytics features to generate insights and statistics on resource usage, popular resources, and other relevant metrics to help library management make informed decisions.

- Data Backup and Recovery: The system should implement regular data backup and recovery mechanisms to ensure data integrity and availability in case of system failures or other unforeseen events.

- User-Friendly Interface: The system should have a user-friendly interface that is easy to navigate and use, with clear instructions and prompts to guide users in performing their tasks efficiently.

- Scalability: The system should be designed to handle a growing number of resources and users, without compromising performance or functionality, to accommodate future growth in the library's collection and user base.

- Documentation: The system should be well-documented with clear instructions for users and administrators on how to use and maintain the system, including user manuals, system documentation, and technical guides as needed.

3.3. Non-Functional Requirements

- Performance: The system should have fast response times and be able to handle a large number of concurrent users and queries, providing efficient search results and recommendations to users.

- Scalability: The system should be scalable to accommodate future growth in the library's collection and user base, without compromising performance or functionality.

- Reliability: The system should be reliable, with minimal downtime and high availability, to ensure uninterrupted access to the library's resources.

- Security: The system should implement appropriate security measures, such as user authentication, access controls and others feasible to protect the integrity, confidentiality, and availability of the data stored in the database.

- Usability: The system should be easy to use, with a user-friendly interface that requires minimal training for users to navigate and perform their tasks efficiently.

- Accessibility: The system should be designed to be accessible to all users with required permissions to ensure inclusivity and equal access to all users.

- Interoperability: The system should be designed to integrate with other existing systems or future systems, allowing for data exchange and interoperability with other library or information management systems.

- Compliance: The system should comply with relevant laws, regulations, and industry standards, such as data privacy regulations, copyright laws, and other regulations concerned with the college library.

- Maintenance and Support: The system should be easy to maintain with mechanisms to assist users and administrators in resolving issues

- Technology Stack: The system should be built on a reliable and scalable technology stack that meets the library's requirements for performance, security, and scalability, with consideration for future technology advancements and updates.