

# **A Natural Language Interface for Movie Database Queries**

## **1. Team details**

Group Name: ChatDB 81

Team Size: 1 member

Member Name: Jinyi Wang

The project is being undertaken individually by Jinyi Wang under the team name ChatDB 81.

## **2. Team Member Background and Skills**

Jinyi Wang specializes in data analysis and has experience working with Python, NoSQL databases, and API-based data extraction. He has a solid understanding of MongoDB, including document-based storage, aggregation queries, and indexing. Additionally, he has experience in backend scripting and data processing, allowing him to build a command-line interface (CLI) for interacting with a NoSQL database. With a strong focus on data handling and query optimization, he is well-equipped to develop a natural language interface for querying MongoDB.

## **3. Project Requirements**

This project aims to develop a Natural Language Interface (NLI) for querying MongoDB, enabling users to interact with a NoSQL database using natural language commands via a command-line interface (CLI). The system will allow users to search for documents, filter data, and modify records without needing to write MongoDB queries manually.

The system will support three key functionalities.

First, it will allow users to explore database collections and documents, enabling them to retrieve available collections, view sample documents, and understand data structures.

Second, it will support query execution, where natural language inputs will be converted into MongoDB queries and executed against a MongoDB database. Users will be able to search for records using filters, sorting, grouping, and aggregations.

Third, the system will allow users to modify data, supporting insert, update, and delete operations through natural language input.

Security measures will be implemented to ensure safe database access and prevent unauthorized modifications.

## **4. Planned Implementation**

The implementation is divided into six phases.

The first phase, spanning the first two weeks, involves setting up the MongoDB database, designing appropriate collections, and populating the database with sample data.

The second phase, in the third week, focuses on implementing schema

exploration functionality, allowing users to retrieve available collections, document structures, and sample data through natural language queries.

The third phase, during weeks four to six, involves integrating an NLP model, such as DeepSeek API, to interpret natural language input and convert it into MongoDB queries for execution.

The fourth phase, in weeks seven and eight, includes implementing data modification capabilities, enabling users to insert, update, and delete documents using natural language commands.

The fifth phase, scheduled for week nine, is dedicated to testing and debugging, ensuring query accuracy, system efficiency, and security.

The final phase, spanning weeks ten and eleven, focuses on refining the CLI interface for better user interaction and making final optimizations to the system.

## 5. Team Responsibilities

As the sole team member, Jinyi Wang is responsible for all aspects of the project. He will design and populate the MongoDB database, develop backend logic for natural language processing, integrate NLP models to interpret user queries, implement security measures, and conduct extensive testing and debugging to ensure system reliability.

## 6. Timeline

Week	Dates	Tasks
Week 1	Feb 7 - Feb 14	Set up MySQL database and retrieve initial movie data from TMDb API
Week 2	Feb 15-Feb 21	Implement schema exploration functionality
Week 3	Feb 22 - Mar 1	Develop natural language query execution for basic movie searches
Week 4	Mar 2 - Mar 8	Improve NLP model accuracy for handling complex queries
Week 5	Mar 9 - Mar 15	Implement filtering and ranking features (e.g., sorting by rating, genre)
Week 6	Mar 16 - Mar 22	Conduct initial testing and debugging
Week 7	Mar 23 - Mar 29	Optimize system performance and security
Week 8	Mar 30 - Apr 5	Final refinements and optimizations
Week 9	Apr 6 - Apr 12	Prepare for final project demonstration
Week 10	Apr 13 - Apr 19	Final project demonstration preparation
Week 11	Apr 20 - Apr 26	Buffer time for unexpected issues

---

<b>Week 12</b>	Apr 27 - May 3	Final documentation and report writing
<b>Week 13</b>	May 4 - May 9	Submit final project report

---