Design and Implementation of Mass Spectrometer Database

L.F. Liu, M. Li, R. S. Mao
Institute of Modern Physics, Lanzhou, China

WEP69
IBIC2024
Sept. 9-13, 2024 · Beijing, China

Abstract

A mass spectrometer, a beam instrument, measures and analyzes the mass and charge of molecules and ions, identifying particle types. Advances in mass spectrometry technology have made data management crucial in fields like biomedicine and environmental science. Mass spectrometer database software, an integral component, handles data storage, management, sharing, and analysis. Effective database systems and retrieval techniques enhance rapid compound identification, offering efficient, accurate substance detection solutions. This paper presents a mass spectrometry database management system that simplifies operations from data collection to analysis, providing a fast, precise tool for research. The system aiding in the identification and analysis of unknown compounds, thus supporting various applications.

INTRODUCTION

Since its introduction in the early 1970s, the mass spectrometry library search technique has become crucial in analyzing complex samples by identifying chemical components through ionization and separation. It's increasingly popular in fields like environmental monitoring, proteomics, and food testing due to its high resolution and speed. Modern mass spectrometry databases, like those from NIST and Wiley, along with software like Agilent's Mass Hunter and Thermo Fisher's Xcalibur, support this analysis, although costs and customization limits are concerns. This study proposes developing an open-source, customizable mass spectrometry data management system to enhance data processing, integration, and sharing, accelerating scientific research and discovery.

Database Design

This system uses the MySQL database management system, which is widely used for data storage and management. For mass spectrometry data, using MySQL can effectively manage a large amount of structured data, support efficient data query and multi-user access, ensure data integrity, and have good scalability and maintainability. The Entity-Relationship (E-R) diagram provides a clear visual overview for database design, aiding in understanding the data flow and relationships between different tables. The E-R diagram of this system is shown in Figure 1.

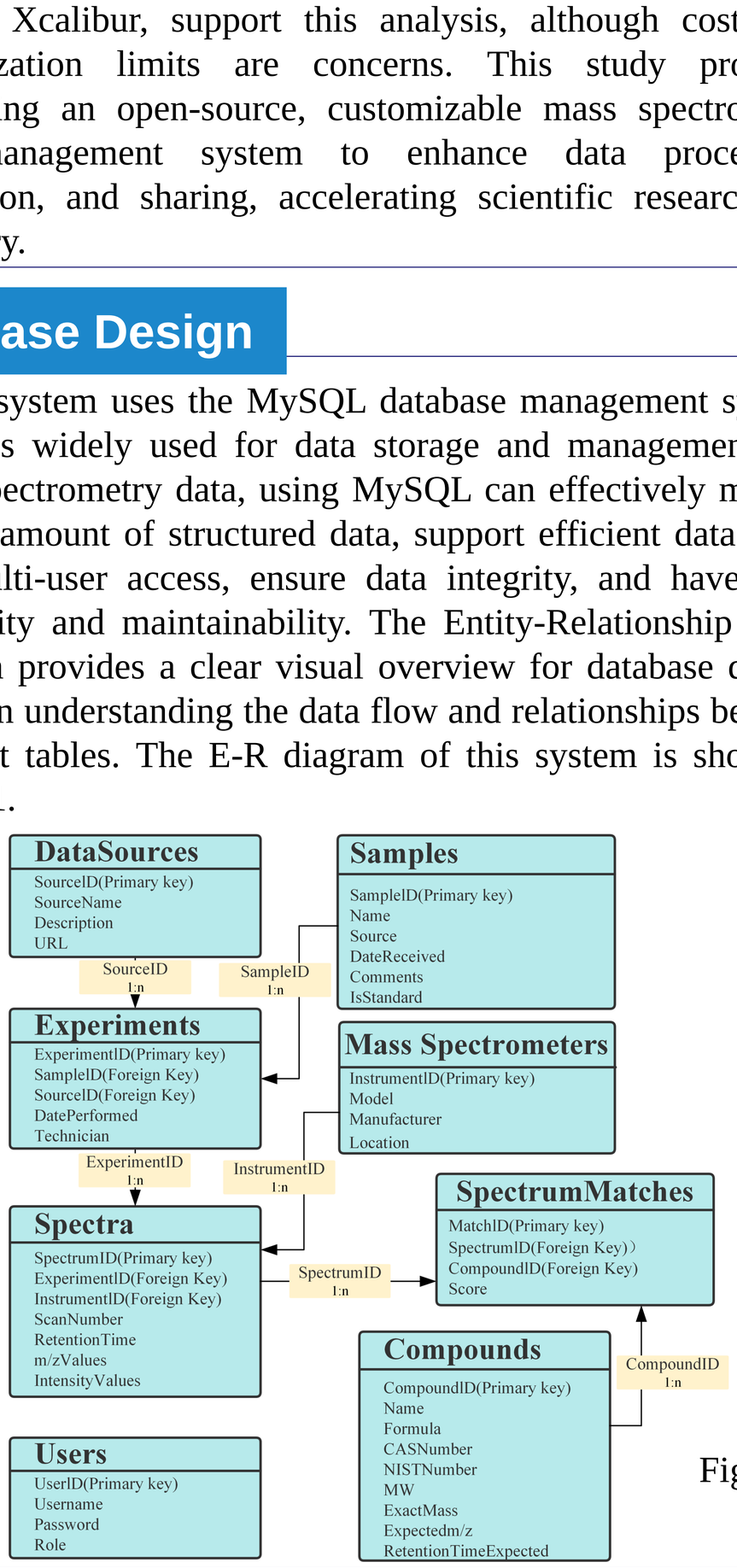

Figure 1

System Implementation

· Technical Implementation

The system uses MySQL database management system for database construction, Python programming language for development, and PyQt for graphical user interface design to create a full-featured, user-friendly, and high-performance mass spectrometry database management system.

· System Architecture

MVC (Model-View-Controller) architecture is chosen as the overall architecture for designing this program, which helps to separate data processing, user interface and business logic, and helps to structure the code to improve the maintainability and extensibility of the code. A simple MVC architecture diagram is shown in Figure 2.

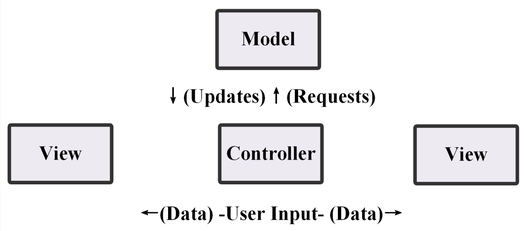

Figure 2

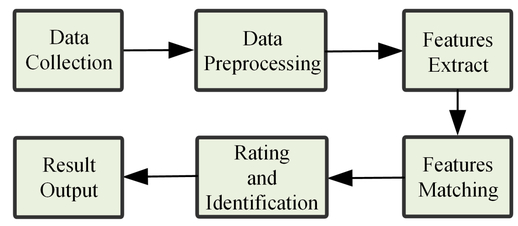

Figure 3

·Mass spectral fingerprint matching

Mass spectral fingerprint matching is a technique for identifying unknown compounds by comparing the characteristic peaks of a sample's mass spectrum with records in a database of known mass spectra. This method relies heavily on features such as accurate mass measurements, peak intensities, and peak patterns. The key steps and techniques for mass spectrometry fingerprint matching are shown in Figure 3.

RESULTS AND DISCUSSION

This paper discusses the development of a mass spectrometry data management system, offering various data input and query options like compound name, molecular formula, and CAS number. It features a graphical user interface, including a dynamic mass spectrometry diagram display, enhancing user understanding and facilitating detailed analysis. The system also supports extensive data management capabilities such as import, export, editing, and deletion, ensuring database reliability and security. Additionally, the integration of Tandem Mass Spectrometry (MS/MS) data is highlighted as crucial for improving the accuracy of compound identification and supporting complex structural analysis. Future enhancements include adding more advanced data analysis and interpretation features related to secondary mass spectrometry.

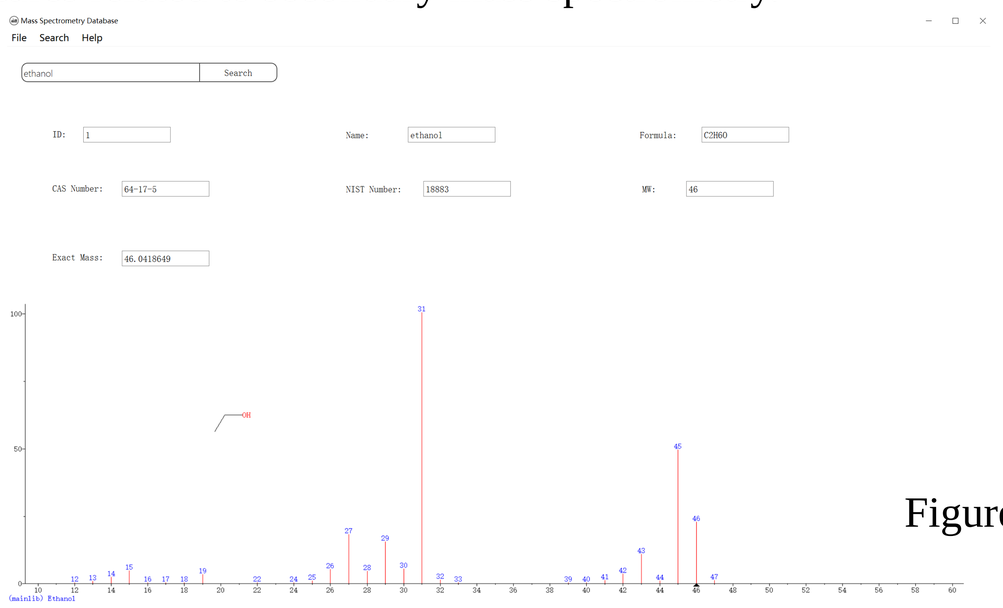

Figure 4

References

[1] Y. Zhou, "The research and application of the algorithm to Mass Spectral Data", M.D. thesis, Fac. of Elect. Eng. and Comp. Sci., Ningbo University, Ningbo, China, 2017.
[2] C. Shou, G. Naren, S. Gao, J. Liu and F. Zhao, "Establishment and application of mass spectral database for natural dyes", J. of Text. Res., vol. 44, no. 11, pp. 120-131, Nov. 2023.