Exploration of Optimization Strategies for Library and Information Work Management Based on Big Data Technology

Jin Wu

Wuhan Railway Vocational College of Technology, Wuhan 430205, Hubei, China

Abstract: With the rise of information technology, big data has penetrated into all aspects of social life and gradually become an important means to improve management efficiency and decision-making level. As an important aspect of information resource management, it will inevitably be impacted and promoted by big data technology. Traditional libraries have rich experience in resource management, information retrieval, and reader services, but in the face of increasing information volume and diverse user needs, traditional library information management methods have shown various drawbacks. The introduction of big data technology provides a new way and possibility for optimizing library information services. Through the collection, storage, analysis, and application of massive data, it is possible to more accurately understand user needs, optimize resource allocation, and improve service quality and efficiency. Therefore, exploring optimization strategies for library and information work management based on big data technology is of great significance.

Keywords: big data technology, library and information work, management optimization

1. Introduction

Exploring the application of big data technology in the management of library and information work has both theoretical significance and practical application value, which helps to expand and deepen the research field of library and information science, promote its deep integration with big data technology, and enrich the theoretical system of big data application. By utilizing big data technology, it is possible to achieve intelligent transformation of library and information work, improve management efficiency, enhance service levels, and provide strong data support for library decision-making and operation. Therefore, exploring optimization strategies for library information management in the big data environment is not only a need for disciplinary development, but also a practical need, and is a hot topic in the current related field.

2. The significance of optimizing the management of library and information work based on big data technology

2.1 Improve information acquisition efficiency

Traditional library information collection mainly relies on manual work, which is not only inefficient but also prone to missed detections. The introduction of big data technology can quickly process massive amounts of data and perform intelligent analysis, greatly shortening the time to obtain information. By utilizing big data technology, a huge database can be constructed, and intelligent algorithms can be used for fast retrieval. Users only need to input keywords or relevant information to quickly obtain the required literature, data, and intelligence. At the same time, utilizing machine learning, natural language processing and other technologies to automatically process unstructured data, further improving the efficiency and accuracy of information acquisition.

2.2 Improving user experience

User experience is very important for library and information work. A good user experience can not only improve user satisfaction, but also enhance user stickiness and loyalty. The application of big data technology enables library and information workers to more accurately understand user needs and preferences, and provide personalized services for readers. By analyzing the borrowing records, browsing history, and retrieval behavior of users, we recommend books, articles, and intelligent information that meet their interests. This not only improves the user experience, but also saves them time and energy. In addition, big data technology can also monitor and analyze user feedback in real-time, in order to timely discover and solve problems encountered by users during use, further improving user satisfaction and experience.

2.3 Optimize resource management

Library information resource management is an important component of library management, which involves various aspects such as resource procurement, classification, storage, maintenance, and use. By using big data technology, the entire process management and fine control of library and information resources are carried out, greatly improving the management efficiency of library and information resources. By utilizing big data analysis technology, accurate predictions can be made on user needs and market trends, which can guide libraries to scientifically purchase and allocate resources, avoiding resource waste and backlog. In addition, big data technology can also monitor and analyze the usage of resources in real time, timely discover and solve problems in resource management, and ensure the effective use and rational allocation of resources.

2.4 Enhance intelligence analysis capabilities

Intelligence analysis, as an important part of library and information work, aims to collect, organize, analyze, and interpret information in order to provide scientific basis and support for decision-making. Traditional intelligent analysis methods mainly rely on human experience judgment, with low efficiency and a certain degree of subjectivity. By introducing big data technology, massive amounts of information can be automatically processed and intelligently analyzed, greatly improving analysis efficiency and accuracy. By utilizing big data technology, it is possible to comprehensively analyze multiple sources and types of information, reveal their inherent connections and patterns, and mine potential intelligence information and decision-making basis. In addition, big data can also use machine learning, data mining and other technologies to deeply analyze and predict historical data, providing more scientific and accurate support for decision-making.

3. Optimization Strategy for Library and Information Work Management Based on Big Data Technology

3.1 Data collection and integration

By utilizing big data technology, data can be effectively collected and integrated, improving the management and service level of libraries. In the application of big data, data collection and integration are important links, including multi-channel data collection, data cleaning, and preprocessing. Multi channel data collection is an important means of applying big data technology to library and intelligence work. It fully utilizes various information resources such as library management systems (LMS), online databases, social media, and user feedback, providing users with rich information resources. LMS can record the borrowing frequency of books, the borrowing history of readers, etc; The online database provides a large number of academic and journal papers; Social media can reflect the enthusiasm of readers for comments and discussions on a book; Reader feedback is the most direct reflection of information by readers. A unified data collection interface can be used to summarize and organize data in multiple formats, forming a comprehensive data pool to achieve multi-channel data collection. Taking a large library as an example, data collection was conducted through multiple channels such as LMS, Google Scholar, Weibo, and WeChat, resulting in a database containing millions of pieces of data. Data purification and preprocessing are important steps to ensure data quality. Due to the diversity of data sources, there is inevitably noise and redundant data in the processing process. Information collected from social media may contain many irrelevant discussions and advertising information; There may be duplicate borrowing records in the data obtained from LMS, and data cleaning techniques can effectively eliminate irrelevant and duplicate data. For example, a cleaning script written in Python can remove irrelevant social media discussions and filter out social media data from a certain library through keyword filtering and deduplication methods. It can ultimately retain valuable discussion data and remove 80% of irrelevant information. Data preprocessing is the standardization and normalization of data to prepare for future analysis. The format of data from different sources varies. For example, borrowing records from LMS libraries can be in CSV format, while borrowing records from online libraries can be in XML or JSON format. Data preprocessing technology is used to achieve unified conversion of data in different formats. Data such as CSV, XML, JSON, etc. can be converted into a unified data framework through the Pandas library and standardized, such as unifying date formats to YYYY-MM-DD, capitalization of text content, etc.

3.2 Data storage and management

The storage and management of data is an important component of library work. With the development of big data technology, it has become possible to effectively store, manage, and use this data. Storage systems represented by distributed storage systems such as Hadoop and HBase can effectively improve data storage efficiency and reliability. Like the Hadoop distributed file system HDFS, it provides a high error tolerance and high throughput data storage solution. HDFS can divide larger files into many small blocks, and then store these blocks in different nodes. Each data block has a copy of multiple nodes, so that when a node fails, the data is still available. This design ensures high availability of the system and improves

scalability, allowing it to increase storage capacity as needed at any time. In practical applications, if a library wants to save electronic data and borrowing records of all its paper books, assuming that the library has millions of borrowing records in a year, the traditional single storage method will be difficult to meet this demand. Using HDFS technology, each borrowing record is stored as a data block in different nodes to ensure data security and reliability. In this way, even if one node fails, copies of other nodes can still ensure data integrity. In addition, to improve retrieval speed and accuracy, an effective indexing system can be established. For example, Elasticsearch, a distributed search engine that can search and analyze large amounts of data in real time, can use inverted indexing technology to associate each item in each document with its unique identifier, in order to achieve rapid retrieval. For libraries, they can establish metadata indexes such as book titles, authors, publication dates, titles, etc. When users input query criteria into the system, the system can quickly retrieve relevant information about books and provide feedback to users. This efficient query mechanism greatly improves the user's query experience and improves the service efficiency of the library.

3.3 Data analysis and mining

User behavior analysis is an important component of data mining, and system analysis can combine user borrowing records, query logs, and access behavior to understand user reading preferences and demand trends. The library can analyze the borrowing situation of the past year through big data analysis to identify which book is the most popular and which subject is the most popular. For example, a certain library borrowed 500000 books throughout the year last year. Through big data analysis, it was found that the borrowing rate of books was highest in literature, accounting for 30%, followed by science books, accounting for 25%. On this basis, the library can adjust its collection structure accordingly, increase popular books, reduce the purchase of unpopular books, and optimize the allocation of collection resources. At the same time, by analyzing user query logs, we can identify keywords that are frequently used by users, thereby improving the book classification system and helping users quickly find the information they need. In addition, topic analysis and hotspot discovery can be used for data mining. Through techniques such as data mining and natural language processing, topic information and hot topics can be extracted from massive literature data [5]. Libraries can discover current academic hot topics by mining literature and materials published in recent years. Assuming that in the past five years, the number of papers related to artificial intelligence has surged, especially in areas such as deep learning and natural language processing, significant progress has been made. At the same time, tracking and analyzing hot topics can provide targeted literature resource support for researchers, enabling them to better grasp the forefront of academic dynamics.

3.4 Information Services and Decision Support

In the field of information services, big data technology can greatly improve the accuracy and personalization level of book recommendations, among which intelligent recommendation systems are an important tool. The system can establish a unique interest model for each user based on their reading history, borrowing records, book reviews, and search behavior. If a user frequently borrows books on artificial intelligence and machine learning in the past six months and gives a higher rating, based on this, the intelligent recommendation system can infer that the user has a great demand for information in the field and recommend relevant new books or high rated classic books to them. In addition, collaborative filtering, content filtering, and mixed recommendation methods can be applied. Collaborative filtering algorithms recommend books to users who have similar reading preferences to the current user by searching for them. The content filtering algorithm can recommend books with similar content to users based on the characteristics of the books they have read. It can organically combine the advantages of the first two algorithms to further improve the accuracy and diversity of recommendations. In addition, big data can provide managers with comprehensive and in-depth analysis reports, helping them make scientific decisions. On the basis of comprehensive analysis of library and information data, various statistical reports such as book borrowing rates, reader satisfaction, and book bestsellers can be generated. Managers can use this information to understand changes in library usage and user needs, and then allocate resources and services.

4. Conclusion

In summary, the optimization strategy for library and intelligence work management based on big data will greatly improve resource utilization and user experience while improving information acquisition efficiency and accuracy. This not only maintains a competitive advantage for libraries and intelligence agencies in the information age, but also provides more convenient and effective information services for researchers and readers. In the future, with the continuous development and innovation of big data technology, library information management is facing unprecedented opportunities and challenges. Based on continuous technological updates and management optimization, the library information field will achieve more intelligent and personalized services, promote the dissemination and sharing of knowledge, and promote social progress and

cultural inheritance.

References

[1] Guo Wei. The Necessity and Approaches of Informationization Construction in Library and Information Management [J]. Science and Technology Wind, 2022, (26): 161-163.

[2] Yang Fan. Exploration of Optimizing Library and Information Management Models in the New Era [J]. Journal of Harbin Vocational and Technical College, 2022, (01): 108-110.

[3] Chen Meina. Exploration of Information Construction in Digital Libraries of Universities [J]. Wireless Internet Technology, 2021,18 (21): 95-96.

[4] Wang Juan, Li Hui, Yang Qing. Optimization Methods for Book Information Management in Modern Libraries [J]. Inner Mongolia Science and Technology and Economy, 2021, (15): 141-142.

[5] Gao Yan. Exploration of Work in the Field of Library and Information Science under the Background of Big Data [J]. Journal of Xuchang University, 2021,40 (02): 147-149.

[6] Jing Rongjuan. Research on Innovative Models of Library and Information Work under the Background of Big Data [J]. Comparative Study of Cultural Innovation, 2020,4 (30): 124-126.