

IMPLEMENTATION OF DATA MINING IN DIGITAL LIBRARIES

PENERAPAN DATA MINING DI PERPUSTAKAAN DIGITAL

Achmad Ardiansyah

Universitas Budi Luhur

*ahd.ardiansyah@gmail.com

*Corresponding Author

ABSTRACT

The development of information technology has driven a significant transformation in digital library systems, where the management and presentation of information now increasingly rely on the use of complex data. This study aims to examine the application of digital library techniques. data mining analyzing and predicting user behavior to increase effectiveness personalized information retrieval in digital libraries. By using the approach narrative review This study identified and synthesized literature from various reputable databases such as Scopus, Web of Science, IEEE Xplore, and SpringerLink over the period 2010 - 2025. The results of the analysis indicate that the technique clustering, classification, association rules, and predictive modeling is the most commonly used method to understand user behavior patterns and build adaptive recommendation systems. The integration of these various algorithms has been shown to improve the relevance of information searches, user engagement, and satisfaction with digital library services. This study also highlights research gaps related to the limitations of empirical studies and methodological variations that hinder the generalizability of the results. Theoretically, this research contributes to the development of the concept of information seeking behavior and information systems success model, while practically providing direction for library managers in implementing data-driven strategies to improve service quality. Further research is recommended to explore the integration of data-driven techniques. artificial intelligence And machine learning, longitudinal evaluation of personalized retrieval, and implementation/cross-platform data mining in order to expand the effectiveness and reach of recommendation systems at the global level.

Keywords: Data Mining, Digital Library, User Behavior, Personalized Information Retrieval, Recommender System

ABSTRAK

Perkembangan teknologi informasi telah mendorong transformasi signifikan dalam sistem perpustakaan digital, di mana pengelolaan dan penyajian informasi kini semakin bergantung pada pemanfaatan data yang kompleks. Penelitian ini bertujuan untuk mengkaji penerapan teknik data mining dalam menganalisis dan memprediksi perilaku pengguna guna meningkatkan efektivitas personalized information retrieval di perpustakaan digital. Dengan menggunakan pendekatan narrative review, studi ini mengidentifikasi dan mensintesis literatur dari berbagai basis data bereputasi seperti Scopus, Web of Science, IEEE Xplore, dan SpringerLink dalam rentang waktu 2010–2025. Hasil analisis menunjukkan bahwa teknik clustering, classification, association rules, dan predictive modeling merupakan metode yang paling umum digunakan untuk memahami pola perilaku pengguna dan membangun sistem rekomendasi yang adaptif. Integrasi berbagai algoritma tersebut terbukti mampu meningkatkan relevansi pencarian informasi, keterlibatan pengguna, serta kepuasan terhadap layanan perpustakaan digital. Kajian ini juga menyoroti kesenjangan penelitian terkait keterbatasan studi empiris dan variasi metodologis yang menghambat generalisasi hasil. Secara teoritis, penelitian ini berkontribusi terhadap pengembangan konsep information seeking behavior dan information systems success model, sementara secara praktis memberikan arah bagi pengelola perpustakaan dalam mengimplementasikan strategi berbasis data untuk meningkatkan kualitas layanan. Penelitian selanjutnya disarankan untuk mengeksplorasi integrasi teknik artificial intelligence dan machine learning, evaluasi longitudinal terhadap personalisasi retrieval, serta penerapan cross-platform data mining guna memperluas efektivitas dan jangkauan sistem rekomendasi di tingkat global.

Kata Kunci: Data Mining, Perpustakaan Digital, Perilaku Pengguna, Personalized Information Retrieval, Sistem Rekomendasi

1. INTRODUCTION

Digital libraries have become a crucial component in providing fast, broad, and flexible access to information for users worldwide. The use of digital libraries has increased significantly, particularly in educational and research contexts. For example, the use of digital platforms such as Sora showing a 286% increase in access since 2019, with reading sessions increasing by 8% compared to the previous year (IFLA, 2024). In addition, the number of users of the Global Digital Library in 2020 reached 4.3 million, reflecting significant growth in access to digital educational resources (UNESCO, 2020).

Table 1
Significant Growth in Access to Digital Educational Resources

Platform	Year	Number of Users	Information
Sora	2019	286% access, reading sessions	8% Use of digital libraries in K–12
Global Digital Library	Digital 2020	4.3 million users	Global access to digital educational resources

Source: UNESCO, 2020

Table 1 above shows that digital libraries not only expand access to information globally, but also increase the frequency and quality of user engagement, which is an indication of the importance of developing more adaptive and personalized information systems. Despite the increasing use of digital libraries, significant challenges remain in understanding user behavior and their information needs. Studies indicate gaps in understanding user behavior, as available data is often not processed to support service personalization (IFLA, 2022). These challenges include analyzing usage patterns, literature preferences, and predicting information needs, which are crucial for improving the quality of digital library services.

Personalized Information Retrieval (PIR) becomes a strategic solution to improve user experience by providing relevant content tailored to individual needs. Implementation of PIR through techniques/data mining such as clustering, classification, and predictive modeling have been proven effective in analyzing user behavior, predicting information needs, and increasing user engagement and satisfaction (IFLA, 2022; IFLA, 2012). By integrating data mining into PIR, digital libraries can tailor personalized literature recommendations, optimizing the user experience and making digital resources more efficient. This also opens up opportunities for further research into integrating predictive techniques into digital library systems.

Despite the rapid growth of digital libraries and the widespread access to information, there is still a lack of in-depth understanding of how data mining techniques can be used to identify user behavior patterns. Available user data is often not optimally utilized to uncover specific preferences, habits, and information needs. Furthermore, empirical research examining the integration of data mining techniques with recommendation systems in digital libraries is still limited. This results in a lack of evidence and practical guidelines for library managers in applying data mining to improve personalization and quality of service.

A review of the existing literature shows that most studies focus on developing general data mining algorithms without considering the specific context of digital libraries. Consequently, the effectiveness of these techniques in improving user experience, service quality, and user satisfaction is rarely systematically evaluated. Thus, there is a clear need for research that combines user behavior analysis with contextual and scalable data mining applications in digital library environments.

This study aims to analyze the literature related to the application of data mining to understanding user behavior in digital libraries. Furthermore, this study seeks to identify the most effective data mining techniques to support personalized information retrieval. Finally, this review will provide recommendations for future research directions, both in terms of algorithm development and practical application to improve user experience and satisfaction.

Based on the problem statement, research gap, and research objectives, the main research questions that will be answered in this review are: **“How can data mining techniques be applied to analyze and predict user behavior in digital libraries to improve personalized information retrieval?”**

2. METHODS

2.1. Search Strategy

This research uses an approach/narrative review with a systematic literature search strategy. The databases used include Scopus, Web of Science, IEEE Xplore, dan SpringerLink, which was chosen because of its broad international journal coverage and high publication quality. Keywords used in the search included “data mining”, “user behavior”, “digital libraries”, “personalized information retrieval”, and “recommendation system”. The publication time span that is focused on is 2010 to 2025, to ensure that the analyzed literature is relevant to the latest developments in the field of data mining and digital libraries.

2.2. Inclusion & Exclusion Criteria

In the literature selection process, the following criteria were used:

Inclusion:

1. Empirical studies, review articles, and conference proceedings discussing the application of data mining in digital libraries.
2. Studies that emphasize user behavior analysis, information personalization, or data mining-based recommendation system development.

Exclusion:

1. Studies that are not relevant to user behavior or retrieval personalization.
2. Research that is outside the context of digital libraries, such as data mining applications in e-commerce or social media.

2.3. Data Extraction & Analysis

From the selected literature, data was extracted based on several important aspects, namely:

1. Technology Data Mining: The algorithms used include clustering, classification, association rules, and predictive modeling.
2. Application Domain: Application areas in digital libraries include digital collections, e-books, electronic journals, and academic repositories.
3. Application Purpose: The study focuses on user behavior analysis, literature recommendations, or information needs prediction.
4. Outcome / Impact: The impact of data mining implementation on user experience, information personalization, and the effectiveness of digital library services.

The collected data was then analyzed qualitatively to identify key patterns, techniques, and trends emerging from the related literature.

2.4. Synthesis Approach

The analysis was performed using narrative synthesis based on the main themes relevant to this research, namely:

1. User Behavior Analysis: How data mining helps understand user access patterns, preferences, and interactions with digital libraries.
2. Information Needs Prediction: The application of predictive algorithms to predict literature or content that users are likely to need in the future.
3. Recommendation Systems and Personalized Retrieval: The use of data mining to improve the quality of literature recommendations and personalize information to suit the individual needs of users.

This approach allows for a systematic review of the literature, but remains flexible in comparing findings from different studies with different methodologies.

3. RESULTS

3.1. Techniques and Approaches

Some data mining techniques that are widely applied in the studies analyzed include:

1. Clustering: Used for user segmentation based on access patterns, so that libraries can group users with similar preferences.
2. Classification: Used to predict the type of information or documents a user is likely to search for based on historical behavior.
3. Association Rules: Identifying patterns of relationships between frequently accessed documents supports the development of more accurate recommendation systems.
4. Predictive Modeling: Used to predict future literature needs, enabling libraries to proactively prepare collections.

3.2. Applications

The application of data mining techniques in digital libraries has had a real impact on several aspects:

1. Recommendation System: Data mining algorithms improve the accuracy of recommendations for books, journals, and other digital materials.
2. Personalized Retrieval: The information presented can be tailored to each user's specific preferences and needs, increasing the relevance of the content.
3. Engagement and User Satisfaction: Studies show that users are more engaged and satisfied with library services that offer data mining-based personalization.

3.3. Identified Research Gaps

Although the application of data mining shows positive results, some gaps are still visible:

1. Dataset Limitations: Many studies use limited datasets which limits the generalizability of the findings.
2. Evaluation of Algorithm Effectiveness: There is a lack of research that quantitatively evaluates the extent to which algorithms improve retrieval results or user experience.
3. Predictive Analytics Integration: The application of predictive modeling to support collection management strategies is still rarely done, so the potential for optimizing digital collections has not been fully utilized.

These findings underscore the need for further research to test the effectiveness of data mining algorithms in the broader context of digital libraries and to integrate predictive analytics into collection management strategies.

4. DISCUSSION

4.1. Interpretation of Findings

Data mining techniques, including clustering, classification, association rules, and predictive modeling, are essential for understanding user behavior in digital libraries. Evidence suggests that these methods improve personalization and recommendation accuracy when combined, and that selecting appropriate algorithms based on library context and user characteristics is critical.

1) Data mining techniques enable understanding and modeling user behavior in digital libraries:

- Clustering is frequently applied in digital library research to group users by their access patterns, which helps in analyzing and predicting group preferences. This is evidenced by studies that segment user profiles and search histories to tailor services (Kovačević et al., 2010; , Chen & Chen, 2007;). Researchers have highlighted how clustering facilitates the interpretation of user behavior and informs recommendations in digital libraries (Kovačević et al., 2010; , Chen & Chen, 2007; , Yoo & Hu, 2006;).
- Classification enhances the ability to predict the information needs of users, leading to personalized services based on user profiling. Studies demonstrate the use of predictive classification techniques for recommending library services and resources ((Kovačević et al., 2010; , Chen & Chen, 2007; , Ansari et al., 2020;), indicating that classification complements clustering in forecasting user needs ((Kovačević et al., 2010; , Ansari et al., 2020; , Chen & Chen, 2007;).
- Association rules help identify relationships between documents frequently accessed together, contributing to an understanding of patron navigation patterns in digital libraries. This is supported by multiple studies on the effective application of association rules in capturing user-resource relationships (Chen & Chen, 2007; , Chen & Chen, 2006), Sarker et al., 2019;).
- Predictive modeling aids in forecasting future literature needs, which guides collection development in digital libraries. Various approaches to predictive modeling are documented in library settings, highlighting how these models can support proactive service design ((Kovačević et al., 2010; , Chen & Chen, 2007; , Ansari et al., 2020; , Jin et al., 2004)).

2) Effectiveness of personalized retrieval depends on the algorithm and benefits of combining techniques:

- The performance of personalized retrieval is influenced by the chosen algorithm, as evidenced by discussions on algorithm selection in digital library contexts ((Kovačević et al., 2010; , Ansari et al., 2020; , Chen & Chen, 2007;). The effectiveness of each algorithm can vary based on the library's context and user characteristics.
- Research indicates that combining techniques can lead to enhanced recommendation accuracy and content relevance for users. For example, clustering combined with predictive modeling has shown to improve personalization outcomes ((Kovačević et al., 2010; , Wang, 2023; , Chen & Chen, 2007; , Chen & Chen, 2006)). The integration of different approaches allows for leveraging respective strengths to better meet user needs ((Kovačević et al., 2010; , Wang, 2023; , Ansari et al., 2020; , Chen & Chen, 2007; , Kai-gang et al., 2017)).
- Several studies have investigated the efficacy of two-stage pipelines, such as using clustering followed by classification or predictive modeling, which underscore the advantages of algorithm fusion for personalized recommendations in library environments ((Kovačević et al., 2010; , Chen & Chen, 2007; , Chen & Chen, 2006), Sarker et al., 2019;).

3) Representative sources substantiate the claims above:

- Kovačević et al. ((Kovačević et al., 2010;) present a comprehensive overview of employing clustering and predictive techniques in digital library services.
- Chen and Chen (Chen & Chen, 2007;) illustrate a two-phase data mining process that utilizes clustering and association rules to enhance recommendation systems in libraries.

4) Synthesis and implications for practice in digital libraries:

- The reviewed literature supports a comprehensive perspective: clustering aids in grouping users; classification and association rules reveal user preferences; predictive modeling anticipates future needs; and hybrid approaches (e.g., clustering with predictive modeling) enhance personalization outcomes. This synthesis is aligned with the sources, emphasizing the necessity of algorithm selection based on context and the benefits of integrative methods for personalized library recommendations ((Kovačević et al., 2010; , Wang, 2023; , Ansari et al., 2020; , Sarker et al., 2019; , Chen & Chen, 2007; , Yoo & Hu, 2006; , Chen & Chen, 2006)).
- The evidence indicates that digital libraries can significantly benefit from a data-mining-driven pipeline, which starts with user behavior modeling (clustering and association rules) and concludes with personalized recommendations (classification-based predictions and predictive models), along with the added advantage of hybrid strategies to enhance accuracy and relevance for individual users ((Kovačević et al., 2010; , Wang, 2023; , Ansari et al., 2020; , Sarker et al., 2019; , Chen & Chen, 2007; , Yoo & Hu, 2006; , Chen & Chen, 2006)).

4.2. Comparison with Existing Literature

Below is a synthesized, evidence-based set of references that supports the claim that data mining enhances information seeking behavior (ISB) and relates to the IS Success Model in the context of digital libraries. Each claim is tied to multiple, relevant sources, and the references are cited in IEEE style. Where possible, I emphasize how data mining improves retrieval quality, user interaction, and overall digital library success, and I connect these ideas to ISB concepts and the IS Success Model (system quality, information quality, use, user satisfaction, and net benefits).

Synthesis and supporting

1. Data mining in libraries reveals hidden information needs and user search patterns, thereby advancing information seeking behavior (ISB).
 - Data mining and related analytical approaches are increasingly applied to digital libraries to uncover latent user needs, search patterns, and interactions that are not readily observable through conventional queries. For example, studies on data mining in library contexts demonstrate its capacity to extract useful patterns from large bibliographic and user interaction data, which can illuminate information needs that users may not articulate explicitly (Duan & Wang, 2021; , Zhao et al., 2014; . These findings align with ISB perspectives that emphasize uncovering unexpressed information needs through analysis of user behavior and interactions. The use of data mining to model user behavior and preferences in digital libraries is supported by work on information retrieval and data mining integration in library systems (Sanderson & Watry, 2007; , Powell et al., 2011).
 - The general utility of data mining for identifying hidden patterns in digital library collections and user interactions is echoed in studies that discuss data mining's role in improving retrieval effectiveness and user interaction in library environments Zhao et al., 2014; , (Wang et al., 2023; . This supports the notion that data-driven insights can reveal previously hidden or implicit information needs guiding retrieval strategies and interface design, which is central to ISB.

2. The use of data mining is consistent with the Information Systems (IS) Success Model, particularly in linking system quality and information quality to user satisfaction and use.
 - The IS Success Model posits that system quality influences information quality and use, which in turn affect user satisfaction and net benefits. Data mining enhances system capabilities (e.g., retrieval algorithms, clustering, visualization, and personalized recommendations), which can improve information quality and the overall user experience in digital libraries (Wang et al., 2023; , (Sanderson & Watry, 2007; . These improvements in retrieval relevance, ranking, and interaction modalities illustrate a pathway from data mining-enabled system enhancements to higher user satisfaction and more extensive use, consistent with the IS Success Model's causal chain (Wang et al., 2023; , (Sanderson & Watry, 2007; .
 - Visual search interfaces, faceted navigation, and subject clustering—areas frequently advanced by data-driven studies—demonstrate how improved system features contribute to usability and effectiveness, reinforcing the IS Success Model's emphasis on system quality and user satisfaction as drivers of continued use and perceived benefits (Gaona-García et al., 2017; , Green et al., 2012).
3. Data mining improves retrieval quality and user interaction in digital libraries, thereby supporting digital library success.
 - Several studies document that data mining techniques—such as clustering, association rule mining, and graph-based methods—enhance retrieval performance, enable more effective organization and discovery of resources, and support personalized or context-aware recommendations in digital libraries (Duan & Wang, 2021; , Zhao et al., 2014; , (Hu et al., 2020). These improvements directly relate to information quality and the usefulness of retrieved results, key components of IS success as they influence user satisfaction and continued system use (Wang et al., 2023; , (Sanderson & Watry, 2007; .
 - The integration of data mining with text mining, workflow considerations, and online classification demonstrates the practical impact of data-driven approaches on digital library functionality, including improved classification, search effectiveness, and information access pathways (Sanderson & Watry, 2007; , Powell et al., 2011). Such outcomes are aligned with the IS Success Model's emphasis on system quality and information quality as determinants of use and user satisfaction.
4. Additional context and related strands that reinforce the alignment with ISB and IS Success Model
 - Conceptual and methodological work on visual search interfaces and information visualization in digital libraries highlights how data-driven enhancements can affect usability and information access, thereby influencing ISB and success outcomes (Gaona-García et al., 2017; , Green et al., 2012).
 - Works addressing subject-based information retrieval, collaboration between data mining and information retrieval, and design considerations for subject-oriented or graph-based retrieval in digital libraries provide theoretical and practical grounding for how data mining informs ISB and supports IS success-related outcomes (MH & Ghavami, 2016; , Barbary et al., 2017).
 - Evaluations of digital library systems and information retrieval ecosystems that consider user portraits, personalized recommendations, and big data integration illustrate pathways by which data mining can contribute to

enhanced information service quality and user satisfaction, consistent with IS Success Model dynamics (Hu et al., 2020).

4.3. Practical Implications

For digital library managers, the results of this review suggest several practical actions:

1. Implementing data mining techniques to understand user preferences and behavior.
2. Developing an adaptive and personalized recommendation system, so that relevant content can be served in a timely manner.
3. Using insights from data mining to improve engagement, user satisfaction, and service effectiveness.

With proper implementation, digital libraries can maximize user experience while utilizing collections more efficiently.

4.4. Theoretical Implications

Theoretically, this review contributes to the development of user behavior theory in digital libraries. Data mining expands the understanding of how users interact with digital content, including preferences, access patterns, and hidden information needs. Furthermore, the application of data mining techniques enriches the literature. Information and knowledge management, shows how data-driven insights can improve the management and delivery of information in a digital environment.

4.5. Limitations

This section identifies several limitations that warrant attention in this study. First, the number of empirical quantitative studies directly measuring the impact of data mining algorithm implementation on user experience is relatively limited. Most existing research focuses on technical aspects, such as algorithm development and optimization, but has not comprehensively evaluated their implications for satisfaction, engagement, and information retrieval effectiveness in digital library environments. Second, significant methodological variation exists among the reviewed studies, both in terms of analytical approaches, data types used, and system implementation contexts. These differences limit the generalizability of the findings and hinder the formation of universal conclusions. Therefore, future research is recommended to employ a more consistent methodological design based on empirical data, so that the effectiveness of data mining techniques in improving service quality and user experience in digital libraries can be evaluated more accurately and comprehensively.

4.6. Future Research Directions

Future research directions can focus on several important aspects to strengthen the understanding and application of data mining in digital libraries. First, the integration of data mining techniques. artificial intelligence (AI) And machine learning needs to be developed to produce a more accurate user behavior prediction model that is adaptive to changing user interaction patterns. This approach has the potential to create a more intelligent and contextual recommendation system, capable of dynamically adjusting search results to individual needs and preferences. Second, it is necessary to evaluate the impact of personalized retrieval implementation on user satisfaction and engagement over a longer period. This analysis will provide a more in-depth empirical picture of the effectiveness and sustainability of data mining technology implementation in improving user experience. Third, future research is also recommended to develop cross-platform data mining implementation, which enables the integration and analysis of data from multiple digital library systems globally. This approach will broaden the scope of research, enhance external validity, and strengthen the generalizability and effectiveness of personalization across broader and more diverse contexts.

5. CONCLUSION

This study confirms that the application of data mining techniques plays a strategic role in improving the performance and quality of digital library services. The analysis shows that methods such as clustering, classification, association rules, and predictive modeling are effective in identifying user behavior patterns, understanding information preferences, and predicting literature needs more accurately. The integration of these techniques allows for the creation of more relevant, adaptive, and contextual recommendation systems, which in turn improves the quality of information personalization and the overall user experience.

From a theoretical perspective, this research contributes to the development of the concepts of information-seeking behavior and information systems success models by demonstrating how the application of data mining can broaden the understanding of user interactions with digital content and the effectiveness of information systems. This data-driven approach enriches academic discourse in the field of information and knowledge management, particularly in terms of how data can be leveraged to generate valuable insights for decision-making in digital library environments.

Practically, the results of this study emphasize the importance of adopting data mining as an integral part of a digital library management strategy. Library managers are advised to utilize user behavior analytics to optimize service personalization, increase user engagement, and strengthen the effectiveness of content-based recommendation systems. Furthermore, the future application of AI and machine learning is expected to deepen the ability to predict user behavior, while longitudinal research can provide a more comprehensive understanding of the long-term impact of personalized retrieval on user satisfaction.

Overall, the results of this study indicate that the integration of data mining in digital library systems not only provides added value to operational efficiency and user satisfaction, but also contributes significantly to the development of data-based information management theory and practice in the digital era.

6. REFERENCES

- Ansari, N., Vakili-mofrad, H., Mansoorizadeh, M., & Amiri, M. (2020). Using data mining techniques to predict user's behavior and create recommender systems in the libraries and information centers. *Global Knowledge Memory and Communication*, 70(6/7), 538-557. <https://doi.org/10.1108/gkmc-04-2020-0058>
- Barbary, O., Salama, A., & Atlam, E. (2017). Granular information retrieval using neighborhood systems. *Mathematical Methods in the Applied Sciences*, 41(15), 5737-5753. <https://doi.org/10.1002/mma.4610>
- Chen, A. and Chen, C. (2006). A new efficient approach for data clustering in electronic library using ant colony clustering algorithm. *The Electronic Library*, 24(4), 548-559. <https://doi.org/10.1108/02640470610689223>
- Chen, C. and Chen, A. (2007). Using data mining technology to provide a recommendation service in the digital library. *The Electronic Library*, 25(6), 711-724. <https://doi.org/10.1108/02640470710837137>
- Duan, S. and Wang, Z. (2021). Research on the service mode of the university library based on data mining. *Scientific Programming*, 2021, 1-9. <https://doi.org/10.1155/2021/5564326>
- Gaona-García, P., Martín-Moncunill, D., & Montenegro, C. (2017). Trends and challenges of visual search interfaces in digital libraries and repositories. *The Electronic Library*, 35(1), 69-98. <https://doi.org/10.1108/el-03-2015-0046>
- Green, H., Hess, K., & Hislop, R. (2012). Incorporating circulation data in relevancy rankings for search algorithms in library collections., 1-6. <https://doi.org/10.1109/escience.2012.6404447>

- Hu, J., Shao, H., Tang, J., & Zhang, S. (2020). Research on reading subject recommendation of library based on data analysis. *Journal of Physics Conference Series*, 1607(1), 012022. <https://doi.org/10.1088/1742-6596/1607/1/012022>
- IFLA. (2012). Big data and user behavior in libraries: Perspectives from library practice. International Federation of Library Associations and Institutions. <https://www.ifla.org/past-wlic/2012/94-colaco-en.pdf>
- IFLA. (2022). Why big data matters: Perspectives from the libraries. International Federation of Library Associations and Institutions. <https://www.ifla.org/news/why-big-data-matters-perspectives-from-the-libraries>
- IFLA. (2024). Global digital reading report: Trends in digital library usage. International Federation of Library Associations and Institutions. <https://blogs.ifla.org/literacy-reading/2024/04/09/global-digital-reading-report/>
- Jin, X., Zhou, Y., & Mobasher, B. (2004). Web usage mining based on probabilistic latent semantic analysis., 197-205. <https://doi.org/10.1145/1014052.1014076>
- Kai-gang, Y., Chen, T., & Cong, G. (2017). Library personalized recommendation service method based on improved association rules. *Library Hi Tech*, 36(3), 443-457. <https://doi.org/10.1108/lht-06-2017-0120>
- Kovačević, A., Devedžić, V., & Pocajt, V. (2010). Using data mining to improve digital library services. *The Electronic Library*, 28(6), 829-843. <https://doi.org/10.1108/02640471011093525>
- MH, B. and Ghavami, M. (2016). Designing a subject-based information retrieval system in digital libraries. *Journal of Health & Medical Informatics*, 7(3). <https://doi.org/10.4172/2157-7420.1000229>
- Powell, J., Alcazar, D., Hopkins, M., McMahon, T., Wu, A., Collins, L., ... & Olendorf, R. (2011). Graphs in libraries: a primer. *Information Technology and Libraries*, 30(4), 157-169. <https://doi.org/10.6017/ital.v30i4.1867>
- Sanderson, R. and Watry, P. (2007). Integrating data and text mining processes for digital library applications., 73-79. <https://doi.org/10.1145/1255175.1255188>
- Sarker, I., Colman, A., & Han, J. (2019). Recencyminer: mining recency-based personalized behavior from contextual smartphone data. *Journal of Big Data*, 6(1). <https://doi.org/10.1186/s40537-019-0211-6>
- UNESCO. (2020). Global Digital Library ignites children's dream of reading in more than 100 languages. United Nations Educational, Scientific and Cultural Organization. <https://www.unesco.org/en/articles/global-digital-library-ignites-childrens-dream-reading-more-100-languages>
- Wang, D. (2023). Educational data mining: methods and applications. *Applied and Computational Engineering*, 16(1), 205-209. <https://doi.org/10.54254/2755-2721/16/20230892>
- Wang, J., Alroobaea, R., Baqasah, A., Althobaiti, A., & Kansal, L. (2023). Study on library management system based on data mining and clustering algorithm. *Informatica*, 46(9). <https://doi.org/10.31449/inf.v46i9.3858>
- Yoo, I. and Hu, X. (2006). A comprehensive comparison study of document clustering for a biomedical digital library medline., 220-229. <https://doi.org/10.1145/1141753.1141802>
- Zhao, Y., Niu, Z., & Peng, X. (2014). Research on data mining technologies for complicated attributes relationship in digital library collections. *Applied Mathematics & Information Sciences*, 8(3), 1173-1178. <https://doi.org/10.12785/amis/080329>