The data security problems discussion in application of library service platform

Xiangfei Zhao*, *Yingcong* Chang, *Huizun* Feng, and *Min* Huang

Hebei Normal University, Library, 050024, Shijiazhuang, China

Abstract. With new technology such as big data, cloud computing, artificial intelligence is widely used in the library, library service platform applies in many more libraries. Although big data collection, data analysis and cloud deployment provide a data basis of deploy operation for the platform, but meantime increases risk of data security, so the paper proposes several measures to ensure data security. Based on comparative analysis and literature research method, the paper analyzes qualitatively risk factors for data security from characters of library service platform. It expounds the necessity of data security, and proposes some measures respectively in each stage of data collection, data storage, data access and application, data sharing and updating, data security assessment, in order to promote greatly the level of data security for library.

1 Introduction

New technologies such as big data, cloud computing, Internet of things, and artificial intelligence are widely used in the library, they brought the opportunity of various application requirements with library, and require library to adapt knowledge and wisdom form to get with changes under the situation of information environment. In order to fulfill their duties better, the library management system needs to be adapted constantly to the change on requirements of digital resources management and reader service. Under this trend, the library management system developed from the first generation library automation System which represented by OPAC and focused on the collection Management originally, to the second generation of Library Integrated Management System (ILS) which characterized by collecting and using in a whole and make library service standardization. Although ILS can realize resource navigation and finding, new media and document delivery services, and other functions, but difficult to manage effectively digital resources integration, or cannot carry out effectively knowledge service according to the rule of knowledge flow. At the same time, ILS has many other problems, such as the cross-platform information island, so ILS is hard to meet demand for development with depth of professional knowledge and wisdom in the library.

Based on the trend, library service platform (LSP) arose at the historic moment, Breeding [1] raised it in the 2012 report, soon was acknowledged spread around in library field. He thinks that LSP should contain all media management in unity of print and electronic resources, support the unified management in the whole process of business and

*Corresponding author: xiangfei1401@126.com

service, provide SaaS service in a form of multi-tenant, support library connectivity and system expansion by APIs, it is a platform that ties integrated management system, electronic resources management tools and digital asset management system together [2].

Besides the above characters, Xiao Zheng, etc. [3] think that the main characters of LSP can analyze all kinds of data in library with the help of cloud service, high availability, high scalability and versatility, and provides support decision for development of the library. Zhang Lei, etc. [4] think that LSP can set up big data analysis platform to analyze deeply mass data through visual data analysis tool, in order to help the library with user monitoring analysis and service optimization decision. Shi XiaoHua [5] thinks that LSP needs to consider fully its data security. As it were, LSP is not just a simple system, But is "Internet and cloud and big data" win-win open academic ecological system with data driven, data security is important issues to consider for LSP construction.

2 Data security problem analysis on LSP

Data security is the foundation of library to improve management efficiency and service quality. Data security is defined by ISO in the computer system [6]: Data security divides into data own security, data safeguard security and data storage security, and use the relevant technology and measures to protect data from corruption, modification and leakage because of the accidental and malicious damage. Data safety law [7] was issued in June 2021: “data security means that necessary measures are taken to guarantee data to with protected effectively and used legally, at the same time, the law has been specified that organizations and individuals should undertake corresponding protection obligations and responsibilities with data security in the data activity ". It is a first issue to consider that all kinds of data security guarantee about service intelligent transformation in the library.

2.1 Data security risks in the application process of LSP

LSP will concern inevitably much information of user, privacy and intellectual property issues in the process of application and deployment. As the data center and cloud platform becomes the main target of network attack, libraries are meeting an important problem that how to guarantee the security of data on the cloud platform. Although libraries can improve work performance in information processing and utilization of LSP, but it will still lead inevitably to data security risks. Therefore, data security is crucial to promote service quality of library.

LSP will generate massive data, which has some types as follows: (1) Business management data, such as reader data (identity information, behavior data, etc.); Digital resources data (various databases, e-books), meantime including about copyright data, using scope, utilization rate; Management system data (entry time, metadata, and other bibliographic information). (2) Service data, such as data on reference proposition project, interlibrary loan, citation and novelty retrieval, thematic and so on.

LSP has the characteristic of comprehensiveness, automation, depth and dynamic. First, LSP collects readers' various behavior data a as far as possible, then LSP paints the portrait of the readers in the multi-dimensional form. however these data is sensitive, in order to ensure data security, it is necessary to restrict reasonably access procedure, mode and subsequent use of library data; Second, LSP can interconnect the OPAC system, space facilities management system, access control system and database management system. Centralized management and correlation analysis of various data can maximize the data value, but at the same time increases the risk of data security.

2.2 New technology security risks

Although LSP will bring profound changes to application environment and service mode of the library, but it will cause library exposed to the threat about cloud service security, and thus affects data security and confidentiality. The condition of cloud server will also bring many unknown risks to data security storage of the library. Blockchain technology can improve the security storage of data, but it lacks systematic security protection measures, and its full backup mechanism also encounters storage bottlenecks.

Therefore, based on analyzing safety risk for LSP and combining technology of safety risk analysis, data security strategy is very necessary, which can help library with data security system, framework, policy and service process, and so as to better promote the wisdom of development in library on management, service , construction and so on.

3 Data security scheme of library service platform

Many domestic scholars have conducted intensive research on library data security, and attempted to find some solutions about data security. For example, Zhou Xiuxia, etc. [8] proposed to use Five Safety framework to plan the safe access of library data, and thus improve the safe access level of library data. Liang Junrong [9] designed a library security risk identification and management system to improve the security of library data from the aspect of security management for library. Wan Yinghong, etc. [10] analyzed problems in personal data security of intelligent library and proposed solutions to protect personal data security. Zhang Juan, etc. [11] raised a plan to deal with library security risks by improving information security management and strengthening interlibrary cooperation.

Library data security runs through the whole process of data collection, storage, access and utilization, sharing and updating. Library will have different expectation on data application and security at different stages, Therefore, different stage adopts different measures can make data security solve in more detail (As shown in **Fig.1.**).

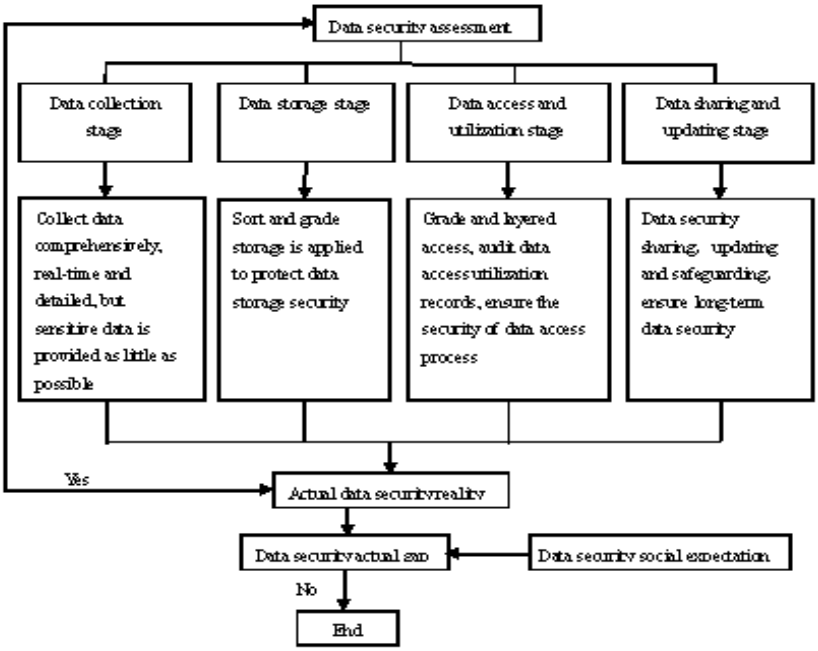

Fig. 1. Data security solution.

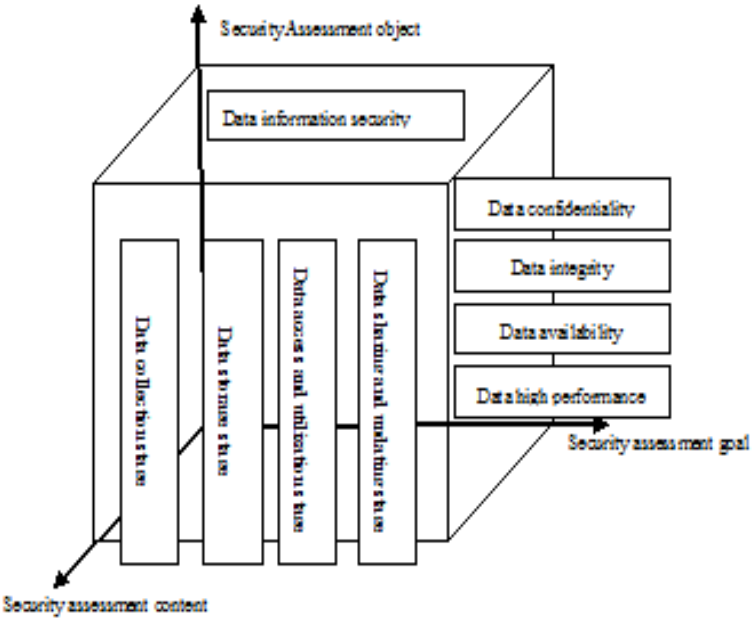

Fig. 2. Security assessment object of library data.

3.1 Data collection stage

LSP can collect data meticulously, comprehensively and dynamically in data collection stage, these data mainly contain two types[12]: First, reader identity data and behavior data, such as reader identity information, entry data, seat reservation data; Second, resource data, mainly includes text data, numerical data, database record data, media files data, retrieval history, and other data. Data collection will provide more reference for the analysis and utilization of the library in the future.

(1) Formulate corresponding data security policies. With the frequent occurrence of network security incidents, readers are more sensitive to the concerns of data security and more will to protect privacy. Readers are becoming more cautious about disclosing sensitive personal information such as ID numbers and student card numbers. According to The Personal Information Protection Law [13], which came into effect in November 1 of 2021, library should follow the principles of legality, legitimacy, necessity and good faith, take corresponding measures to ensure the security of readers data. At present, most libraries at home have not explicitly reached an agreement with readers on personal data collection and security [14], and rarely published management rules with data security on library websites. Meatime, cloud service providers can easily collect sensitive data of libraries through relevant technology, and even have super authority to collect data. Once the authority is out of control, it is easy to bring huge data security risks. Therefore, libraries should unveil corresponding policies and measures to make up loopholes in the above aspects.

(2) Determine different collection method according to the content and purpose. First, we should adopt anonymous and fuzzy methods to handle the data that collected highlight behavior of reader groups. For example, in the library management system, readers' information with borrowing, access and download is shown through the analysis data of borrowing, resource remote access and downloads according to the time axis [15], and the sensitive data of readers is blurred by data desensitization technology. Second, library

should formulate different data security policies according to different data application environment and make data security plan in advance when collect data with obvious identity characteristics from readers.

3.2 Data storage stage

Based on the cloud deployment of LSP, data transferred from the local storage to the cloud storage platform. This means that libraries no longer own the data alone; moreover, the cloud service provider also owns the data. Therefore, the data security of library is greatly threatened.

Although libraries will try their best to ensure data security, they will be affected by various factors in the actual operation process. For example, capacity of data storage device, security technology reserves and so on. Therefore, they can take several measures to ensure data security as follows: First, according to the index such as data importance and security level, adopt different ways to classify and distribute storage. Data related to readers' identities and behaviors, and data of the core assets of the library need to be stored in local server. Second, data with low security level can be stored on the public cloud platform to reduce the storage cost after encryption desensitization. Third, libraries should introduce web firewall software, network intrusion detection and alarm system by referring to “Basic Requirements for Network Level Protection of Information Security Technology” and “Assessment Requirements for Network Level Protection of Information Security Technology” [16]. they should also adopt network isolation to divide domain and rank protection for main business systems that aiming at establishing an integrated network security protection system of "attack and control" to ensure data storage security.

3.3 Data access and utilization stage

Data access and utilization is a process in which libraries and cloud service providers extract and analyze data recommend readers to topics that they are interested in, count and identify readers and analyze their behaviors in management and service, so as to dig deeper into the data value. In this stage, libraries need to make efforts in the following aspects.

(1) Establish a clear data access and utilization institution. In most cases, data security incidents are caused by human activities. Therefore, libraries need to formulate a feasible standard of data access and utilization to protect all kinds of data and form a control mechanism of data security. On the one hand, libraries should standardize behavior of accessing and using of data for staff, and determine the posts and responsibilities of relevant personnel, based on the principle of “who uses who bears the responsibility”. Libraries should implement the responsibilities for relevant personnel according to post demand distribution corresponding privileges, especially standardize management on data security personnel with employed and transferred [17]. On the other hand, robot dynamic security defense technology apply with data security, it shifts from passive defense to active defense, makes data security automatize and instrumentalize, and can be full time, all-round protect data security, thus safeguards data security among libraries and readers, cloud service providers and data vendors in the process of data access and utilization.

(2) Audit accessing and utilization data records. On the one hand, libraries should carry out the corresponding audit procedure according to the privilege of staffs, and restrict the behavior of librarians. Moreover record the time of accessing and using the data, the responsible person, the application and Whereabouts; On the other hand, libraries should sign the corresponding access control agreement with the cloud service provider on the accessing and utilization of data. Only in this way, data can protect from abuse and disclosure.

(3) Grade and distribute to access data. Libraries and cloud service providers can grade and distribute to access data. First, only library staffs with permission can access data with readers' privacy and the core property of libraries, but cloud service providers cannot access. Second, cloud service providers can access and use the data that represent management and service aspects of libraries, this way can ensure the data security to the maximum extent.

3.4 Data sharing and updating stage

Resources, management and service data can provide information support for the intelligent and personalized service of libraries, and provide data basis for third-sector organizations to improve and evaluate the quality of products and services. Therefore, libraries need to establish a data sharing mechanism with readers and third-sector organizations. In the process of data sharing, libraries must follow the principle of rational use of data and use synthetically reliable security management technology to ensure data security. The residual value of data will decrease gradually with the change of data application and reader information, so it is necessary to periodically update and maintain expired and redundant data for reducing the cost of data operation and maintenance.

3.5 Data security assessment stage

Data security assessment analyzes safety risk of data activity to make it visual and controllable from the perspective of data security, so as to enhance the level of protection for library data security, and to ensure data confidentiality, integrity, operability and high performance. Moreover, eventually realize the standardization of the data security and fine management (As shown in **Fig.2.**).

During the data collection stage, libraries should pay attention to risks on environment, behavior, transmission and management of collection, carry out targeted security assessment regularly, and formulate corresponding index on data security assessment. For example, whether the permission or role of the collector is clear and whether collects unauthorized data; whether the collection behavior is standard; whether data transmission encrypts after collection; whether the collected data is processed by security gradation.

During the data storage stage, libraries need to pay attention to security issues such as data storage environment, data storage encryption, data storage space hierarchical distribution, access control of data storage, data disaster backup and recovery, and furthermore periodically assess these security risks.

During the data access and utilization stage, libraries need to evaluate the following risks: (1) the cloud platform manages personnel authentication and permission improperly, resulting in unauthorized users to access data. (2) The cloud platform lacks sensitive data discovery and identification mechanism, resulting in sensitive data leakage after analysis [18]. (3) The lack of security assessment and audit methods leads to the cloud platform be unable to supervise effectively access behaviors of users, this situation can increase the risk of sensitive data leakage, and cause huge economic losses and social impact on data owners.

In the data sharing and updating stage, these risks need to evaluate: (1) In the process of data sharing, cloud service providers directly transfer the unencrypted or undesensitized data to the third-sector organizations for sharing. (2) Library data is used to improve them own services or products and for other purposes, libraries and cloud service providers should specify the purpose with signing the service agreement. (3) After the end of the data life cycle, the data has not been completely updated and maintained, and there are still residual media of sensitive data.

4 Suggestions to promote data security of LSP

4.1 Improve laws and regulations on data security

At present, Legal system is still not perfect for data security protection in our country. In recent years, China has issued “network security law”, “public library law”, “data safety law”, “personal information protection law” and other laws. However, though these laws have given the primary responsibilities for data security, and have make clear the importance of data security to national security, but they not described in detail with how to ensure data security of library and other public welfare units. In order to ensure the smooth implementation of LSP, libraries should formulate data security regulations in accordance with libraries themselves, and avoid unnecessary disputes arising from data leakage or data property rights in the future.

Relevant data security management regulations should concern the following three aspects: (1) Libraries should specify the scope and use of each stage of data activities, the technology and methods for data security, the format of data disclosure, etc., and data security management regulations on the website. (2) Libraries should establish relevant identity authentication, authority, responsibility review mechanism to avoid unclear rights and responsibilities and leading to data leakage; (3) Libraries should sign data security agreements with cloud service providers and other third-sector organizations to clarify the rights and obligations on maintaining data security.

4.2 Apply new technologies to strengthen data security

Network technologies represented by big data, cloud computing and artificial intelligence have changed the academic ecosystem of libraries greatly from theory to practice, and promoted the progress of data security technology. The implementation of many measures and imagines on data security of library needs the support of relevant techniologies. Therefore, libraries need to pay attention to three related technologies: (1) blockchain technology. It is a data structure that connects data blocks in sequence in the form of chain. Two adjacent data blocks are associated with each other. In the case that one data block not modifies, other data blocks can hardly tamper, which can improve the storage security and reliability of library data. Domestic scholars have conducted in-depth studies on blockchain technology in data security [19], data security of user portrait [20] and other aspects. Therefore, blockchain technology will play a great role in the establishment of data security system in libraries. (2) Situational aware technology of network security. Strengthen data security transmission defense measures, predict and monitor possible data security transmission risks in advance, carry out security defense means, and improve the level of data security. (3) Cloud storage technology. In order to ensure the security of data storage, various cloud platforms and technical security frameworks of big data are constantly upgrading. By means of encryption and desensitization, libraries should strengthen security of data access and storage on cloud platforms, and clarify the ownership of data. Libraries should always pay attention to the development trend of new technologies, which must be attached importance to the application of LSP in the future.

4.3 Improve the level of data management in library

At present, library development is rapidly moving towards the datafication, the informatization and the cloud, but at the same time, data is always facing security threat. Therefore, libraries should improve the level of data management from the following two

aspects: (1) Libraries should formulate a data security management system according to the specific situation of themselves, and make corresponding procedures for each stage of data activities [21], and reduce the probability of data security being damaged. (2) Do well in network security grade protection, improve the security physical environment, secure communication network, secure area boundaries and secure computing environment, improve the network and business system security application protection ability, hidden danger detection ability and emergency handling capacity, effectively ensure the data security in libraries.

5 Conclusion

LSP is an open source information service system based on cloud platform and big data technology architecture. It is extremely important for safe and stable operation. How to ensure data security is an important factor for development and service upgrade in libraries. This paper analyzes data security threats in the process of LSP operation from various stages of data activities, and proposes the construction of data security solutions. In this age of enforcing “Network Security Law”, “Data Security Law”and “Personal Information Protection Law”, it is expected to promote the formulation and implementation of library data protection measures. It has certain guiding and practical significance for data security work.

This research was support by social science development research project of Hebei Province (NO.: 20200502011) and College humanities and social science research project of Hebei Province (NO.: SQ201035).

References

1. BREEDING M. Perceptions 2012: An Inter-national Survey of Library Automation[EB/OL]. [2021-05-22]. https://librarytechnology.org/perceptions/2012/.
2. Qian Guo Fu. Research and Development with the Next Generation Library Service Platform [J]. Library Tribune, 2019, **39** (5) : 62-66.
3. Xiao Zheng, Lin Jun Wei. Establishing the Next Generation Library Service Platform with Microservices: FOLIO as an Example [J]. Library Journal, 2018, **37** (11) : 63-69.
4. Zhang Lei, He Chen Zhi, Zhao Liang. Data and Knowledge Service for the Third Generation Library Service Platform [J]. Journal of the National Library of China, 2018, **27** (6) : 40-47.
5. Shi Xiao Hua, Wang Xin, Xu Jing, et al. Research on the Development Status and Characteristics of the NewGeneration Smart Library Service Platform [J]. Journal of Academic Libraries, 2019, **37** (2) : 49-54.
6. Data Security[EB/OL]. [2021-05-12]. https://baike.so.com/doc/6144889-6358066.html.
7. Data Security Law[EB/OL] . [2021-06-17]. https://www.sohu.com/a/471563814_260616.
8. Zhou Xiu Xia, Liu Wan Guo, Sui Hui Min, et al. Five Safes Security Framework and its Enlightenment to access of Sensitive data security in Library field in China [J]. Information Studies:Theory & Application, 2020, **43** (3) : 85-89.
9. Liang Jun Rong. Research on security analysis and management of library information system based on big data decision [J]. Library Theory and Practice, 2017 (3) : 93-98.

10. Wan Ying Hong, Zhang Lu Yue, Wan Li. Research on personal data protection of smart library based on big data Application [J]. Research on Library Science, 2018 (3) : 31-34.
11. Zhang Juan, Li Yi. Security risks and countermeasures of library readers' personal information under cloud computing[J]. Information Studies:Theory & Application, 2017, **40** (5) : 39-43, 49.
12. Huang Guo Bin, Zheng Xia, Wang Ting. Information security risks caused by cloud service protocols and countermeasures of library and information institutions [J]. Library and information service, 2020, **64** (12) : 38-47.
13. Personal Information Protection Law [EB/OL].[2021-11-15]. https://baike.so.com/doc/5469653-5707565.html.
14. Li Yi, Zhang Juan. Research on the setting of personal information rights for library readers -To care for readers' personality [J]. Library Tribune, 2015, **35**(6) : 76-81.
15. Lu Kang. Research on the Construction of Digital Resource Statistics System for University Libraries [J]. Journal of Modern Information, 2015, **35**(9) : 140-145.
16. Zhang Hui. Research on University Network Security System Based on Network Security Level Protection 2.0 [J]. Network Security Technology & Application, 2020 (2) : 83-84.
17. Wang Dan, Sun Yang, Xie Hui, et al. Research on network security system of agricultural academic institutes based on network security level protection 2.0 - Taking Chinese Academy of Agricultural Sciences as an example [J]. Journal of Library and Information Science in Agriculture, 2020, **32** (12): 97-103.
18. Tong Xin, Ren Wang, Feng Yun Bo. Security risk analysis and evaluation method of big data platform [J]. Secrecy Science and Technology, 2018 (2) : 6-14.
19. Liu Ming Da, Chen Zuo Ning, Shi Yi Juan, et al. Research progress of blockchain in data security [J]. Chinese Journal of Computers, 2021, **44** (1) : 1-27.
20. Liu Hai Ou, Yao Su Mei, Huang Wen Na, et al. Dilemma and countermeasures of big data application in mobile library user portrait - Based on blockchain concept [J]. Research on Library Science, 2019 (23) : 26-33.
21. Wang Wei Qiu, Liu Chun Li, Qiu Yu Hong, et al. Research on big data security of smart library [J]. Journal of Library Science, 2020, **42** (8) : 91-94, 106.