Information and Knowledge Management (2025)
Clausius Scientific Press, Canada

DOI: 10.23977/infkm.2025.060213
ISSN 2523-5893 Vol. 6 Num. 2

Research on User Privacy Protection Strategies in the Context of Smart Libraries

Tianling Guan[1,a,*]

[1]*Library, Guangdong Ocean University, Yangjiang, China*
[a]*gtl@gdou.edu.cn*
**Corresponding author*

Keywords: Smart Libraries; User Privacy; Privacy Protection; Data Security; Optimization Strategies

Abstract: Against the backdrop of the transformation of smart libraries driven by artificial intelligence and the Internet of Things, the expanded scope of user data collection has raised the risk of privacy breaches. Balancing service quality and privacy protection has become a critical issue for the industry. Utilizing literature review and case analysis, this study examines the core forms and emerging characteristics of user privacy in smart libraries, and analyzes key risks such as excessive data collection, insufficient storage and transmission security, algorithmic tracking misuse, and regulatory gaps. Grounded in data security theory and the principles of personal information protection laws, the research constructs a four-pronged optimization framework encompassing "legal systems, technical safeguards, management services, and user participation." This framework includes measures such as developing specialized regulations, standardizing informed consent procedures, enhancing data security technologies, optimizing algorithms, conducting librarian training, and promoting user privacy education. The findings contribute to the theoretical discourse on privacy protection in smart libraries and offer actionable recommendations for libraries at various levels to mitigate privacy risks and achieve sustainable, intelligent transformation.

1. Introduction

1.1. Research Background

Driven by the deep integration of digital technologies such as artificial intelligence, the Internet of Things, and big data, the library sector is accelerating its transformation from a "traditional document repository" into an "intelligent service platform." Smart libraries, leveraging technologies like AI retrieval systems, personal recommendation algorithms, self-service loan devices, and spatial awareness terminals, have achieved precision, efficiency, and contextualization in services. Users can complete book reservations, seat bookings, and literature searches via mobile apps, while libraries optimize service delivery by analyzing user loan records, search behaviors, and spatial activity data. However, this technological empowerment is accompanied by a continuous expansion in the scope of user data collection. It has extended beyond basic personal information and loan history to include granular behavioral data such as search keywords, dwell time, and device interaction trails,

significantly widening the risk exposure for user privacy breaches[1].In recent years, privacy complaints have arisen in multiple regions due to issues like excessive data permissions and non-compliant data storage practices in library apps. Consequently, balancing service quality with privacy protection during this intelligent transformation has become a pressing practical challenge for the library industry.

1.2. Research Significance

Current domestic research on smart libraries predominantly focuses on technological application and service innovation. Discussions on user privacy protection are often fragmented and attached as subsections within broader thematic studies, lacking an independent and systematic framework. Comprehensive analysis and targeted empirical research in this area remain notably underdeveloped. By delineating the core manifestations and risk characteristics of user privacy within the smart library context and constructing a corresponding "risk-strategy" framework, this study aims to enrich the theoretical discourse in this field, refine the theoretical system of library science in the digital era, and provide a foundational paradigm for subsequent research.

Privacy protection in smart libraries is crucial for maintaining user trust and ensuring service sustainability. Integrating industry case studies with policy requirements, this study proposes optimization strategies from legal, technical, managerial, and user-participation dimensions. It offers actionable guidance for libraries at various levels to mitigate privacy risks and standardize data management practices. This approach assists libraries in achieving their intelligent transformation while safeguarding user rights, thereby enhancing service credibility and user satisfaction.

1.3. Research Approach and Methodology

This study follows a core logic of "conceptual definition to analysis of current status and risks to strategy proposal." First, it clarifies the core concepts of smart libraries and user privacy, and reviews relevant theoretical foundations. Subsequently, employing literature review and case analysis methods, it synthesizes existing measures and identifies core risks in current smart library privacy protection practices. Finally, targeting the root causes of these risks, it proposes optimization strategies across four dimensions: legal systems, technological application, management services, and user participation.

During the research process, the literature review method is adopted. Chinese databases such as CNKI and WanFang, alongside international journals like *Library Quarterly* and *Journal of Academic Librarianship*, serve as primary sources. This enables a systematic review of research findings, industry reports, and policy documents in the fields of smart libraries and user privacy protection, thereby solidifying the theoretical foundation. The case analysis method is employed by selecting typical domestic and international cases, such as privacy compliance practices in European and American libraries and data breach incidents involving university library apps in China. This facilitates the extraction of experiential insights and lessons learned regarding privacy protection, enhancing the practical relevance of the research [2].

2. Definition of Core Concepts and Theoretical Foundations

2.1. Core Characteristics of Smart Libraries

A smart library represents an advanced form of digital library, characterized by "data-driven operations, technology-enabled functions, and intelligent services." In terms of technological support, it integrates digital technologies such as AI, the Internet of Things (IoT), big data, and cloud

computing to achieve the digitization of collections, the automation of service processes, and the precise identification of user needs. Regarding service models, it transcends traditional physical space limitations, forming an integrated "online + offline" service system that encompasses diverse functions like intelligent retrieval, personalized recommendations, self-service lending/return, and remote consultation. In data applications, it enables the dynamic optimization of service delivery through the collection and analysis of user behavior data across the entire service journey. Its essence lies in the deep integration of "technological tools, data resources, and service scenarios."

2.2. Connotation and Scope of User Privacy

User privacy in the context of smart libraries encompasses three core categories of information: Firstly, basic personal information, including identifying details such as name, ID number, contact information, and identification photos. Secondly, behavioral data, including loan records, search keywords, seat usage logs, online operation trails, and duration of stay in specific spaces. Thirdly, associated information, comprising auxiliary data related to individuals, such as device models, IP addresses, and payment information.

Compared to traditional libraries, privacy in smart libraries exhibits new characteristics. On one hand, the "stealthiness" of data collection increases; IoT devices and algorithmic systems can gather behavioral data without users' conscious awareness, heightening the covert nature of potential privacy breaches. On the other hand, the "chain-reaction" effect of privacy leaks becomes more pronounced; fragmented user behavioral data, when integrated by algorithms, can form precise behavioral profiles. Once leaked, this may lead to cascading risks such as targeted fraud and infringement of rights. Furthermore, the "ambiguity" of privacy boundaries intensifies. While enjoying personalized services, users often find it difficult to clearly distinguish between "necessary information" and "excessive information," making their rights to informed consent and choice vulnerable to infringement.

2.3. Theoretical Foundation

Data Security Theory emphasizes the confidentiality, integrity, and availability of data throughout its entire lifecycle—collection, storage, transmission, and use. Its core requirement is to prevent data leakage, tampering, and misuse through technical and managerial measures. This theory provides the fundamental logic for privacy protection in smart libraries, mandating that libraries establish security safeguards across the entire data process to protect user privacy from harm.

The principles of "Data Minimization," "Informed Consent," and "Transparency" constitute the legal framework for library privacy protection. "Data Minimization" requires libraries to collect only user information essential for service provision, avoiding excessive collection. "Informed Consent" requires libraries to clearly inform users of the purpose, scope, and use of information collection, obtaining their voluntary authorization. "Transparency" requires libraries to publicly disclose their privacy protection rules, safeguarding users' right to know and to supervise [3].

3. Analysis of Current Status and Risks in User Privacy Protection in Smart Libraries

3.1. Overview of Existing Privacy Protection Measures in Smart Libraries

Currently, most smart libraries have implemented basic privacy protection technologies. In the data storage phase, methods like encrypted storage and cloud backups are employed to prevent data breaches. For access control, tiered permissions are established, granting only authorized librarians' access to sensitive user information. On service platforms (apps, official websites), privacy policies are published, clearly outlining the scope of information collection and its usage. Some advanced

libraries have also adopted data anonymization techniques, desensitizing user behavioral data before using it for service optimization.

Management protective measures primarily include formulating internal data usage guidelines that clarify the procedures and responsible parties for accessing user data; conducting privacy protection training for librarians to enhance their understanding of data security regulations and operational standards; and establishing data security inspection mechanisms to regularly identify potential privacy risks. Furthermore, some libraries restrict third-party access to user data by signing confidentiality agreements with technology vendors.

3.2. Analysis of Core Privacy Risk Points

Smart library service platforms, while enhancing user experience, introduce significant data protection challenges across several dimensions. A primary concern is the practice of excessive data collection, where applications compel users to grant permissions for non-essential information such as location data or contact lists as a prerequisite for accessing core services like book searches or reservations. This is often facilitated through a "blanket authorization" model during registration, which fails to provide granular consent options. Consequently, users cannot autonomously choose the scope of information they provide, which undermines the principles of informed consent and user choice, significantly increasing the risk of unnecessary information disclosure. This practice contradicts the established legal principles of "informed consent" and "minimum necessity".

In the realm of data storage and transmission, substantial vulnerabilities persist. Technically constrained small and medium-sized libraries often lack encrypted storage systems, leaving user data in plain text vulnerable to cyber-attacks. For libraries relying on third-party cloud services, the risk extends to the potential inadequacy of the providers' own security measures. During data transmission, some platforms fail to employ encrypted protocols, making user behavioral data susceptible to interception, especially over public Wi-Fi networks. Furthermore, when libraries collaborate with technology vendors or advertisers, a failure to clearly delineate data usage boundaries can easily lead to the illegal collection or misuse of user data by these third parties.

The application of algorithms introduces another layer of risk, primarily concerning behavioral tracking and profile misuse. Personalized recommendation systems require the continuous collection of user data, such as search keywords and loan preferences, to build detailed user profiles. The leakage or misuse of this data constitutes a direct infringement on user privacy. Some algorithms exhibit "over-tracking" vulnerabilities, continuing to collect behavioral data even after personalized recommendations have been disabled by the user. The inherent opacity of these algorithms can also lead to "privacy discrimination," where services are restricted based on user profiles, thereby infringing on users' right to equitable access.

Underpinning these technical risks are often inadequate legal and management frameworks. Many libraries lack specialized privacy protection regulations, resulting in blurred boundaries for data collection, use, and storage, with unclear accountability. Privacy policies are frequently drafted in professional jargon that is obscure and difficult for users to understand, further hindering genuine informed consent. Simultaneously, insufficient privacy protection awareness among staff can lead to unauthorized access or dissemination of user data. The frequent absence of comprehensive emergency response mechanisms for privacy breaches hinders timely remediation, potentially amplifying the damage when a leak occurs.

3.3. Typical Case Studies

This study selects representative cases from both positive and negative perspectives to provide practical references for constructing privacy protection strategies. As a negative case, in 2022, a

university library's smart service app in China suffered a data breach due to inadequate backend access controls and unencrypted data storage. The incident led to the leakage of sensitive user information including names, student ID numbers, borrowing records, and the last six digits of national identification numbers triggering collective student complaints. The app was subsequently forced to suspend operations for rectification, severely undermining the library's credibility and public trust. On the positive side, examples from the European Union and the United States illustrate proactive regulatory and practical approaches. In the EU, the implementation of the General Data Protection Regulation (GDPR) has driven institutions such as the German National Library to adhere strictly to the "data minimization" principle, collecting only information essential for services and imposing strict limits on data retention periods[4]. In the United States, ongoing legislative improvements often spurred by industry advocacy have resulted in milestones such as California's 2004 Senate Bill 1834, the world's first legislation explicitly prohibiting libraries from using RFID technology to identify patrons. Later laws, including Washington State's HB1011 and Nevada's Senate Bill No. 264, further clarified that libraries must promptly delete data retrieved from RFID tags and are prohibited from reading personal information without prior user consent[5][6].

4. Optimization Strategies for User Privacy Protection in Smart Libraries

4.1. Legal and Regulatory

Libraries should develop tailored privacy protection protocols, such as Smart Library Reader Privacy Management Guidelines, grounded in national legislation like the Personal Information Protection Law. These protocols must delineate clear standards for the entire data lifecycle—collection, storage, usage, and transmission. This includes defining the boundary between "essential" and "non-essential" information to prohibit the collection of data irrelevant to service delivery; stipulating time limits for data retention and ensuring prompt anonymization or deletion upon service termination; and clarifying privacy responsibilities across departments with a "whoever collects, is responsible" accountability mechanism. Such measures address current regulatory ambiguities and mitigate risks of unauthorized data exploitation.

Privacy policies should be communicated in clear, accessible language, avoiding technical or legal jargon. Libraries ought to adopt visual aids (e.g., icons or flowcharts) to transparently disclose how reader data will be processed, including purposes, scope, methods, and recourse mechanisms. Moreover, a granular consent model should replace blanket authorization—enabling users to toggle permissions for distinct data types (e.g., location, borrowing history) through independent options in service interfaces. Crucially, declining non-essential permissions must not impede access to core services like book searches or reservations. This approach aligns with the principle of minimal necessity and empowers readers to exercise autonomous control over their data.

4.2. Technical Dimension

It is recommended to embed a "Data Collection Scope Control Module" within service systems to ensure only information essential for specific services is collected. For instance, basic borrowing services should require only identity information such as name and ID number, without mandating access to contacts or photo albums. For personalized recommendation services, an "anonymous participation" option should be provided, allowing users to contribute behavioral data for algorithm optimization without linking it to their personal identity. Furthermore, a "dynamic collection" mode can be implemented, where data is gathered temporarily based on the specific service context and automatically purged upon service completion.

For data storage, robust encryption technologies like AES-256 should be adopted to encrypt

sensitive user information. Regular data backups and security audits are essential to identify vulnerabilities. Data anonymization techniques should be applied to loan records and behavioral trails, removing personally identifiable information before any data analysis. During data transmission, encrypted protocols such as HTTPS must be utilized to prevent interception. When leveraging cloud storage, selecting providers with strong compliance credentials and proven security capabilities is crucial, and these should be underpinned by strict confidentiality agreements in contracts.

Algorithmic transparency is paramount. Libraries should clearly inform users about how recommendation algorithms work and the scope of data used. Providing an explicit "algorithm opt-out" right allows users to disable personalized recommendations at any time, with the guarantee that related behavioral data collection ceases afterward. Additionally, algorithms should be designed to prevent excessive tracking by establishing "data collection cooling-off periods," avoiding frequent, purposeless data gathering. Conducting regular security audits of algorithms is necessary to identify and mitigate vulnerabilities that could lead to privacy breaches

4.3. Management and Service Dimension

At the management and service level, standardizing processes and enhancing capabilities are essential. This involves strengthening librarian training in privacy protection to elevate professional competence. Privacy protection should be integrated into routine training programs, covering relevant laws and regulations, library-specific privacy policies, and secure data operation protocols. Utilizing methods such as case studies and scenario simulations will improve librarians' ability to identify privacy risks and handle incidents effectively. Furthermore, establishing an assessment mechanism for privacy protection linking evaluation results to performance metrics—will reinforce a sense of responsibility among staff. Concurrently, a robust emergency response mechanism for privacy breaches must be implemented. This requires the development of a "Smart Library Privacy Breach Emergency Response Plan," which clearly defines reporting procedures, responsible parties, containment measures, and remediation steps. In the event of a data leak, an immediate emergency response should be activated to investigate the cause and scope, promptly notify affected users, and execute data containment and system rectification. A dedicated channel for handling user privacy complaints should also be established to ensure timely responses and resolutions within specified timeframes, thereby minimizing potential harm.

4.4. User Participation

Enhancing privacy protection awareness among users requires proactive educational outreach. Libraries should utilize multiple channels such as official social media accounts, on-site posters, and informational brochures in service areas to disseminate knowledge on privacy risks specific to smart libraries, personal information protection techniques, and available recourse channels. Outreach efforts should be tailored to different user groups, such as elderly patrons or students, to increase relevance and effectiveness. Furthermore, libraries must ensure users' control over their personal data by establishing clear mechanisms for data access, rectification, and erasure. Users should be able to review what information is collected and how it is used, request corrections to inaccuracies, and apply for the deletion of data no longer necessary, through library apps, official websites, or service desks. Dedicated feedback channels, such as designated email addresses or hotlines, should be set up for users to submit privacy-related inquiries and suggestions, with libraries committed to providing timely responses and implementing service improvements based on this feedback.

5. Conclusion

5.1. Research Findings

This study focuses on the issue of user privacy protection within the context of smart libraries. By clarifying core concepts and analyzing current risks, it proposes multi-dimensional optimization strategies. The research finds that user privacy in smart libraries encompasses core categories such as basic personal information and behavioral data. In technological applications, it faces multiple risks including excessive data collection, insufficient security in storage and transmission, misuse of algorithmic tracking, and inadequate regulatory systems. Addressing these issues requires constructing a four-pronged protection framework integrating "legal systems, technical safeguards, management services, and user participation." This involves defining privacy protection boundaries and rules at the legal and regulatory level; strengthening security measures across the entire data lifecycle at the technical level; standardizing service procedures and enhancing staff capabilities at the management level; and promoting user education and rights protection at the user level. This framework can serve as a reference for libraries to balance service quality with privacy protection during their intelligent transformation.

5.2. Limitations

This study primarily employed literature review and case analysis as its main research methods. It did not conduct field research to obtain primary data, leading to insufficient analysis of the differentiated privacy protection needs across libraries of different types and scales. The proposed optimization strategies are general in nature, and their specificity and operational feasibility require further validation. Furthermore, the study did not deeply explore the innovative applications of emerging technologies such as AI and blockchain in privacy protection, indicating that the research depth could be enhanced.

5.3. Future Research Directions

As digital technologies continue to evolve, privacy protection in smart libraries will face increasingly complex challenges. Future research should prioritize empirical investigations to develop nuanced solutions that address the distinct needs of different library contexts. For instance, field-based studies could help design differentiated privacy frameworks tailored to the specific operational and user-profile characteristics of academic, public, and specialized libraries. Moreover, researchers should explore the integration of emerging technologies such as blockchain to enable decentralized data storage, enhance transparency in data usage, and ensure tamper-proof traceability of user information. Such technical advancements can significantly strengthen privacy safeguards while maintaining service efficiency. Finally, international comparative studies are essential to examine privacy regulations, technical standards, and ethical practices across different regions. By learning from advanced international experiences, libraries can develop adaptable and culturally sensitive privacy protection frameworks, foster global collaboration and promote the sustainable development of smart libraries. [7]

References

[1] An, L. (2024). A study on the conceptual definition of library users' personal information: From the perspective of China's Personal Information Protection Law. Journal of Library Science, 46(2), 1–6, 14.
[2] Cai, Y. J., & Chen, B. W. (2025). Research on the integrated path of data certification and data compliance in China: Insights from the EU General Data Protection Regulation. Journal of Qingdao University of Science and Technology

(Social Sciences Edition), 41(4), 69–82.
[3] Liu, X. C. (2017). IFLA's statement on the right to be forgotten: New trends in reader privacy protection. Library Journal, 36(12), 91–95.
[4] Ma, C. H. (2021). Research on library user privacy protection strategies in the modern information technology environment. Henan Library Journal, 41(12), 112–114.
[5] National People's Congress Standing Committee of the People's Republic of China. (2021). Data Security Law of the People's Republic of China. Gazette of the Standing Committee of the National People's Congress of the People's Republic of China, 352(5), 951–956.
[6] Tian, Y. (2015). Analysis of strategies adopted by American libraries to address reader privacy issues involving RFID technology. Hebei Sci-Tech Library Journal, 28(1), 9–12.
[7] Yang, L. L. (2022). From personalized service to privacy intrusion: A review of research on reader privacy protection in the context of smart library development. Library World, (2), 47–52.