Influencing Factors and Optimization Strategies for the Quality of Library Services for Scientific Data Intellectual Property in the Algorithm Era: Postprint

Authors: Xu Yue, Xi Zijie, Pan Chao

Date: 2024-04-03T00:00:00+00:00

Abstract

[Purpose/Significance] In the era of big data recommendations, library information presentation channels have shifted from offline physical entities to online digital platforms, during which the usage frequency of scientific data has exhibited exponential growth. This study aims to analyze the intrinsic mechanism through which the service guarantee of intellectual property rights for scientific data influences the development of digital libraries. [Method/Process] Based on the questionnaire survey method, a questionnaire was administered to highly educated groups. The survey yielded 252 valid samples, and confirmatory factor analysis was employed for testing. The results revealed that the reliability of each questionnaire factor ranged between 0.724~0.913, and the questionnaire demonstrated good validity through confirmatory factor analysis testing. [Results/Conclusion] Five categories of factors—policy and regulation implementation efficiency, talent team construction, data management technology, service model diversification, and the number of data property rights sharing agreements—significantly and positively influence scientific data intellectual property services. Based on these findings, five optimization strategies are proposed for libraries in the field of scientific data intellectual property services.

Full Text

Factors Influencing and Optimization Strategies for the Quality of Library Services for Scientific Data Intellectual Property Rights in the Algorithm Era

XU Yue[1], XI Zijie[2], PAN Chao[3] [1]Harbin Medical University, Harbin 150076 [2]Harbin National Railway Technology Group Co., Ltd., Harbin 150001 [3]Fourth Affiliated Hospital of Harbin Medical University, Harbin 150001

Abstract: [Purpose/Significance] In the era of big data recommendations, libraries' information delivery channels have shifted from offline physical entities to online digital platforms, during which the frequency of scientific data usage has exhibited exponential growth. This study aims to analyze the internal mechanisms through which intellectual property rights (IPR) service guarantees for scientific data influence the development of digital libraries. [Method/Process] Based on questionnaire surveys, we conducted a survey of highly knowledgeable groups, obtaining 252 valid samples. The reliability of questionnaire factors ranged between 0.724 and 0.913, and confirmatory factor analysis verified that the questionnaire demonstrated good validity. [Results/Conclusions] Five factors—policy and regulation implementation efficiency, talent team building, data management technology, service model diversification, and the number of data property rights sharing agreements—all significantly and positively influence scientific data IPR services. Based on these findings, we propose five optimization strategies for libraries in the field of scientific data IPR services.

Keywords: algorithm era; scientific data; intellectual property rights; talent team construction; service quality

1 Research Background

In the information age, particularly in the algorithm era, libraries must provide services such as information retrieval, data source provision, and data analysis [2]. Simultaneously, libraries face challenges related to data security and privacy protection. The high-frequency characteristics of data usage have made information collection and utilization easier, but have also introduced numerous privacy and security risks [3]. With the development of the internet and social media, access to information has become more diversified and convenient, requiring libraries to meet increasingly diverse and personalized user demands. As the scale of data in library information services continues to expand, the difficulty of providing broader information resources and more intelligent services has increased [1].

Scientific research data, as an important information product, constitutes a crucial knowledge resource for promoting scientific research and technological development. As the scale and importance of scientific data continue to grow, IPR protection issues in library information services have become increasingly prominent. Effective IPR protection can safeguard the innovation motivation and research enthusiasm of knowledge creators, thereby promoting scientific research and technological development [19]. Scientific data is typically held and managed by data creators, providers, and other stakeholders who have the right to protect their data resources and legitimate interests [20]. In the digital era, library scientific data IPR services have concentrated on issues such as weak service awareness, unclear management standards, and low service quality [29]. Current research on scientific data IPR protection primarily focuses

on insufficient legal constraints, inconsistent implementation standards, inadequate regulatory and protection measures, and insufficient awareness among data owners and managers [21-23].

2 Literature Review

2.1 Algorithm Applications and Definitions in Library Services

2.1.2 Research on Library Service Development in the Algorithm Era In the field of library information services, research on algorithm applications is abundant both domestically and internationally, primarily focusing on three aspects. First, machine learning algorithms applied to library recommendation services, represented by automatic navigation algorithms for library book-finding robots [8] and book recommendation systems based on library loan records and bibliographic information [9]. Second, big data mining algorithms applied to library personalized services, represented by intelligent library big data system design [10] and user clustering analysis for targeted services [11]. Third, cloud computing algorithms applied to library security construction and data preservation services, represented by improved encryption algorithms for cloud computing user privacy protection [13] and multi-model data preservation systems based on algorithmic cloud computing [14].

However, existing research has primarily focused on the positive impacts of algorithms on library services, with less attention paid to potential risks and preventive measures. Few studies have examined IPR issues related to scientific data in libraries during the algorithm era. This study employs quantitative methods to explore the influencing factors of library scientific data IPR service quality and propose targeted optimization strategies.

2.1.1 Algorithm Definition and Connotation In contemporary society, algorithm definitions span multiple disciplines, primarily involving computer science and information communication. With technological advancements, algorithms have evolved from computer-related terms to 演绎 data involving artificial intelligence and human society, acquiring richer connotations. Algorithms are generally defined as instructions for solving specific problems, manifested as logical code sequences that produce controlled results. Two main interpretations exist: tool-based algorithms referring to computer science concepts, and power-based algorithms referring to algorithms utilized by social power to achieve value orientations, shaping society, culture, and even acquiring status as truth [7]. In library-related fields, algorithms primarily refer to the tool-based connotation.

2.2 Scientific Data and Intellectual Property Protection in the Algorithm Era

2.2.1 Scientific Data and IPR Protection Scientific data refers to data related to scientific research. Domestic scholars define scientific data as research data in digital form, including data stored on computers and other types of data convertible to digital form [15], encompassing various data generated during scientific research activities [16]. International definitions consider scientific data to encompass broad categories, including raw experimental results, graphics, digital drafts, result data, and derivative data obtained during research [18].

The importance of scientific data IPR protection has garnered academic attention. Effective IPR protection safeguards data resources from unauthorized use, preventing legal confusion and significant economic losses [20]. Specific protection measures include national regulations such as China's "Measures for Scientific Data Management" (2018), "Digital China Construction Development Strategy" (2019), and "Several Policy Measures for High-Quality Economic Development" (2019), which positioned digital data as a production factor [24]. In library practice, scientific data IPR services include patent database customization, patent novelty searches, IPR training, and consultation services [26]. However, current services primarily focus on result data and research data patent applications, with insufficient specialized services for process data and derivative data [33].

2.2.2 Scientific Data IPR Service Content and Quality Influencing Factors Current research indicates that scientific data IPR protection faces issues of legal constraints, inconsistent implementation standards, inadequate regulatory measures, and insufficient awareness. These deficiencies affect service quality. IPR services for process data and derivative data are particularly limited due to their heterogeneity and complexity. The quality of scientific data IPR services is highly sensitive to new technologies and environments, significantly influenced by algorithmic technology and policy implementation efficiency [32]. Service quality is affected by factors including policy implementation efficiency, talent team building, data management technology, service model diversification, and data property rights sharing agreements.

---

3 Research Process and Results

3.1 Research Hypotheses

Based on the above analysis of scientific data IPR service content and library service development in the algorithm era, this study proposes the following hypotheses regarding factors influencing service quality:

H1: Policy and regulation implementation efficiency significantly positively influences service quality.

H2: Talent team building significantly positively influences service quality.
H3: Data management technology significantly positively influences service quality.
H4: Service model diversification significantly positively influences service quality.
H5: The number of data property rights sharing agreements significantly positively influences service quality.

3.2 Questionnaire Design

The questionnaire comprises five dimensions:

3.2.1 Policy and Regulation Implementation Efficiency: Adapted from Ye Lan's [40] maturity questionnaire for university library scientific data management capabilities.

3.2.2 Talent Team Building: Based on Zhang Hao et al.'s [41] evaluation indicators for university librarian team construction, assessing technical talent development, management systems, and personnel training.

3.2.3 Data Management Technology: Selected from Liu Xuhui's [42] questionnaire on technology usefulness for university library services.

3.2.4 Service Model Diversification: Adapted from Chen Lili's [43] exploration of diversified service models, including specialized channel services.

3.2.5 Data Property Rights Sharing Agreements: Based on Li Jian et al.'s [44] research, assessed through paid agreements and sharing restrictions.

3.2.6 Scientific Data IPR Service Quality: Evaluated from three dimensions—convenience, usefulness, and security—referencing Chen Huiqi et al.'s [45] research.

3.3 Research Subjects and Sampling

The study targeted high-level scientific research practitioners, including doctoral and postdoctoral researchers, while excluding master's students. To ensure sample diversity and heterogeneity, we also included researchers from research institutes and enterprises. Given the dispersed geographical distribution of samples, online distribution was adopted. From March 11 to May 22, 2023, we distributed 300 questionnaires, recovered 273, and obtained 252 valid questionnaires after excluding responses with completion times under 300 seconds or over 3,000 seconds. The sample included 176 males and 97 females, with 153 teachers, 58 researchers and enterprise personnel, and 62 doctoral and postdoctoral researchers, reflecting a balanced demographic distribution.

3.4 Statistical Analysis Results

3.4.1 Reliability Analysis SPSS reliability analysis was conducted on the five influencing factors and service quality variable. As shown in , reliability

coefficients ranged from 0.853 to 0.860, all exceeding 0.7, indicating acceptable reliability. Corrected item-total correlations exceeded 0.5 within each factor, demonstrating high association between items and factors. Cronbach's alpha if item deleted was lower than factor reliability, confirming that removing items would not improve reliability.

3.4.2 Exploratory Factor Analysis The overall questionnaire KMO value was 0.815, exceeding 0.7 and approaching 0.9, indicating suitability for factor analysis. Bartlett's sphericity test yielded ^{2} = 2,189.556 ($p < 0.001$). Six components with eigenvalues > 1 were extracted, explaining 81.28% of total variance. As shown in and , rotated eigenvalues ranged from 2.696 to 2.056, with all item loadings > 0.5 and no cross-loadings, confirming the preliminary variable structure.

3.4.3 Confirmatory Factor Analysis A first-order confirmatory factor analysis model was constructed with 18 observed items as indicators and 6 latent factors. Using Amos 26.0, all standardized factor loadings exceeded 0.5, indicating strong item-dimension relationships. Model fit indices () showed CMIN/DF = 1.778 (< 3), RMSEA = 0.024 (< 0.05), TLI > 0.9, and CFI > 0.9, demonstrating good model fit ().

3.4.4 Correlation Analysis Correlation analysis () revealed significant positive correlations between all independent variables and service quality: policy implementation efficiency ($r = 0.511$, $p < 0.01$), talent team building ($r = 0.432$, $p < 0.01$), data management technology ($r = 0.399$, $p < 0.01$), service model diversification ($r = 0.528$, $p < 0.01$), and data property rights sharing agreements ($r = 0.554$, $p < 0.01$).

3.4.5 Multiple Linear Regression Analysis Multiple linear regression was conducted with service quality as the dependent variable. The model summary () shows adjusted $R^2 = 0.816$, indicating that independent variables explain 81.6% of variance in service quality. The standard error was 2.54456, and Durbin-Watson = 1.778, suggesting good independence of residuals. Normal P-P plots ([Figure 1: see original paper], [Figure 2: see original paper]) confirmed approximate normality.

Regression coefficients () show all five hypotheses were supported: - Policy implementation efficiency: B = 0.275, p = 0.024 - Talent team building: B = 0.264, p = 0.018 - Data management technology: B = 0.169, p < 0.001 - Service model diversification: B = 0.358, p = 0.045 - Data property rights sharing agreements: B = 0.329, p = 0.048

All VIF values were below 10, indicating no multicollinearity issues. These results validate H1-H5, confirming that all five factors significantly and positively influence service quality.

4 Conclusions and Discussion

The study identifies five key factors influencing library scientific data IPR service quality in the algorithm era: policy implementation efficiency, talent team building, data management technology, service model diversification, and data property rights sharing agreements.

4.1 Policy Implementation Efficiency Mechanism

Policy implementation efficiency positively influences service quality. Policy formulation and implementation strengthen IPR protection, increase service provider accountability, and enhance market competition and transparency, thereby motivating providers to improve service quality [46, 48].

4.2 Talent Team Building Mechanism

Talent team building positively influences service quality. High-quality talent constitutes the primary productive force for scientific data IPR services. Professionals with strong research and innovation capabilities can deliver more scientific achievements and IPR outcomes, providing continuous momentum for service quality improvement [47].

4.3 Data Management Technology Mechanism

Data management technology positively influences service quality. Advanced data management facilitates data sharing, enhances reproducibility, ensures data accuracy and reliability, and helps scientists control data usage, thereby protecting privacy and security [38].

4.4 Service Model Diversification Mechanism

Service model diversification positively influences service quality through four mechanisms: (1) enhancing service provider capabilities, (2) increasing service adaptability and flexibility, (3) promoting knowledge and experience sharing, and (4) fostering healthy competition that drives innovation [31].

4.5 Data Property Rights Sharing Agreements Mechanism

The number of data property rights sharing agreements positively influences service quality. Increased agreements raise researchers' awareness of data protection, promote the development of comprehensive data management systems, and enhance data security, reliability, and availability.

---

5 Optimization Strategies

5.1 Strengthen Policy Implementation

Establish robust policy implementation mechanisms, particularly for commercial scientific data protection referencing patent laws. Enhance policy publicity to improve public awareness and participation. Ensure coordination among policies to form integrated implementation frameworks.

5.2 Enhance Talent Team Building

Strengthen talent development through training programs that improve professional competence and innovation capacity. Enhance management-level talent reserves to provide strategic leadership. Establish incentive mechanisms with preferential policies and rewards to encourage innovation in scientific data IPR services.

5.3 Advance Data Management Technology

Invest in R&D of data management technologies to improve applicability and effectiveness. Promote technology application across all stages of scientific data IPR services. Establish standardized technical specifications to ensure service quality and consistency.

5.4 Promote Service Model Diversification

Drive service model innovation by integrating multiple service approaches to create synergistic effects. Develop customized and personalized services tailored to user needs and service characteristics to enhance satisfaction and user engagement.

5.5 Strengthen Inter-Library Collaboration

Improve the number of data property rights sharing agreements through enhanced library collaboration. Establish legal frameworks defining data property rights. Develop standardized agreement protocols and strengthen implementation supervision to ensure effective execution and improve service quality.

References

[1] DU H, LI C L, ZHAO P P, et al. The reconstruction of traditional media news production mode by “big data”[J]. 2016(1): 66-78.

[2] GE X L. Research on ethical issues of information dissemination in algorithm era[D]. Nanjing: Nanjing Normal University, 2019.

[3] YANG N, ZHANG Z Q. The new position and challenge of digital library in scientific big data era - Postscript on the 16th advanced digital library seminar[J]. Library & information, 2020(1): 127-135.

[4] CHAKRABORTY T, BHATIA S, CARAGEA C. Special section on mining knowledge from scientific data[J]. Expert systems, 2021, 38(4): e12710.

[5] LIN H C. Vertical innovation, foreign direct investment, and asymmetric imitation: A welfare analysis of intellectual property protection[J]. Southern economic journal, 2021, 88(2): 789-827.

[7] WANG J, SONG Y F, LI X C. Research on intelligent technology fusion mechanism in cooperative governance of smart library in algorithmic era[J]. Library, 2020(1): 53-58, 87.

[8] RASHID M R, UZZAMAN N, HOSSAIN S, et al. Development of a self-navigating algorithm for library book finder robot[C]//2017 3rd International Conference on Electrical Information and Communication Technology (EICT). Piscataway, New Jersey: IEEE, 2017: 1-6.

[9] TSUJI K, YOSHIKANE F, SATO S, et al. Book recommendation using machine learning methods based on library loan records and bibliographic information[C]//2014 IIAI 3rd International Conference on Advanced Applied Informatics. Piscataway, New Jersey: IEEE, 2014.

[10] SIMOVIĆ A. A big data smart library recommender system for an e-educational institution[J]. Library hi tech, 2018, 36(3): 498-523.

[11] XU B J, RECKER M. Teaching analytics: A clustering and triangulation study of digital library user data[J]. Educational technology & society, 2012, 15(3): 103-115.

[12] HE S, FENG X L, WU Q H, et al. Research on personalized services of library based on user behavior modeling and big data mining[J]. Library and information service, 2017, 61(1): 40-46.

[13] LU K, LIU H, ZHANG X X, et al. Research on data risk governance of smart libraries in China in the era of "society 5.0"[J]. Library, 2023(4): 22-27.

[14] XUE Y, ZHU X F. Research on improved encryption algorithm for privacy-preserving problems under cloud computing environment[J]. Information science, 2016, 34(9): 145-149.

[15] XIONG Y. Empirical resources and utilization of international open science data[J]. Library theory and practice, 2016(1): 12-14.

[16] LI X H. Research data management and service pattern in libraries[J]. Journal of library science in China, 2011, 37(5): 46-52.

[17] WEN Z J. Model of OCLC worldcat cloud computing digital library[J]. Library and information service, 2012, 56(15): 54-60.

[18] WU D, XU H. Human-centered data science education: New development of library and information science[J]. Documentation, information & knowledge, 2021, 38(6): 12-20.

[19] CHEN S X, LIU G F, LIU Q. Research progress and implementation of FAIR principles for scientific data management[J]. Journal of library and information science in agriculture, 2022, 34(8): 30-41.

[20] SUN S. The concept restoration and system integration of intellectual property law from the perspective of legal interest protection theory[J]. Zhejiang academic journal, 2021(4): 85-94.

[21] WANG L P, YANG B, QIN X, et al. The contents, models and trends of patent information service in university libraries[J]. Library and information service, 2015, 59(6): 113-119.

[22] DU P P, LI Y K, ZHANG X Y. Research on conceptual domain and management boundary of data set in China's measures for scientific data management[J]. Library journal, 2022, 41(4): 96-105, 114.

[23] YANG L C. Research on the ability framework and development strategy of librarian scientific data literacy[J]. Library theory and practice, 2018(9): 90-95.

[24] GUO J. The legal protection of database by the anti-improper competition law[J]. Journal of the second northwest institute for ethnic minorities, 2005(3): 92-97.

[25] LUO W X, LIU X W, WANG Y M. Exploration of online services in university libraries for the whole process of intellectual property rights - A case study of the intellectual property information service platform of Peking university[J]. Research on library science, 2023(2): 67-75.

[26] ZHANG G P, CHEN H Y, CHEN J, et al. A research of the factors of patent information service capability of academic libraries[J]. Research on library science, 2022(3): 41-51.

[27] ZHANG S J, CHEN W J, YUAN Q, et al. Construction of patent information service system of university libraries for industrial technology innovation[J]. Information science, 2021, 39(4): 75-84.

[28] LU Y, HU J Q, ZHANG Y T, et al. Research on library data service system of different stakeholders based on the dissemination of scientific data[J]. Journal of academic libraries, 2021, 39(4): 63-72.

[29] ZENG J J, LIU T, ZHANG R. Patent information service strategies of academic libraries oriented to patent supply chain[J]. Journal of library and information science in agriculture, 2021, 33(5): 40-50.

[30] LIU C, WANG Z X. The experience of u.s. global change data and information sharing and its indication to China[J]. Advance in earth sciences, 2002, 17(1): 151-157.

[31] BAO Z Y, ZHAO N X. Construction and application of patent information service mode in university from the perspective of big data[J]. Information studies: Theory & application, 2020, 43(3): 111-115.

[32] DUAN M Z, CHU J L, ZHANG D R, et al. On the current situations of smart library development in Chinese "double first-class" universities[J]. Library tribune, 2022, 42(1): 91-101.

[33] WU F J, GU H H. The accurate service mode and countermeasures of patent information in university library driven by data[J]. Information science, 2022, 40(4): 49-55.

[34] ZENG P, GAO F, NING L. Analysis and evaluation of library management system based on RFID[J]. Library and information service, 2013, 57(9): 75-79.

[35] LIU J, SHAO S Y, PENG P. Health assessment system construction and empirical research on the science and technology information service ecosystem[J]. Library and information service, 2018, 62(11): 88-96.

[36] GUO Y J, YUAN Y M, ZHANG T F. Ecological mechanism of user information interaction in the metaverse environment[J]. Journal of library and information science in agriculture, 2022, 34(6): 4-13.

[37] WANG X F, WANG X L, SHI Q, et al. Thoughts on the construction of intellectual property service quality evaluation system[J]. Journal of university of electronic science and technology of China (social sciences edition), 2019, 21(3): 24-30.

[38] YE L. Research on the construction and application of the service capability maturity model for research data management in university libraries[J]. Library and information service, 2022, 66(6): 3-14.

[39] YAN L G, SHEN J L, WEI L X. Research on the cooperation model of intellectual property services in university libraries and patent agencies[J]. Library science research & work, 2021(10): 32-35, 45.

[40] ZHANG H, ZHANG J, XU H, et al. The practice of innovative librarian team construction in academic libraries - Taking Shanghai Jiao Tong university library as an example[J]. Journal of academic libraries, 2023, 41(2): 5-12.

[41] FENG Z Y, CHEN R, YAO M, et al. Challenges and opportunities for co-development and sharing of literature resources in library of our age[J]. Chinese journal of medical library and information science, 2020, 29(8): 62-65.

[42] LIU X H. Research on the current mobile reading behavior of college students and the service of college library - An empirical analysis based on the technology acceptance model[J]. Research on library science, 2021(16): 82-95.

[43] CHEN L L. Research on diversified service model of information literacy education in public library[J]. Library work and study, 2021(11): 5-10.

[44] SU R Z, ZHANG K. Research on personalized service mode of university library based on situational awareness[J]. Library work and study, 2022(7): 5-12.

[45] LI J, ZHAO L L, LIN J, et al. Exploration of university library's intellectual property information services oriented to scientific research[J]. Information science, 2021, 39(10): 56-62.

[46] CHAO R, HUANG X L, ZHOU Y. Research on the supply mode of intellectual property information service in university libraries - Based on the perspective of technological innovation demand[J]. Library work and study, 2021(5): 55-65.

[47] CHEN H Q, LÜ Y, LIU M R. Research on online training quality of intellectual property information services for subject librarians in universities - Based on data from FULink alliance librarians[J]. Library and information service, 2021, 65(21): 76-83.

[48] ZUO T T. Analysis on the external service mode of university digital library from the perspective of network diversification[J]. Information science, 2022, 40(1): 66-72.

Note: Figure translations are in progress. See original paper for figures.