
AI translation · View original & related papers at
chinaxiv.org/items/chinaxiv-202307.00413

A Review of Research on Open Sharing of Scientific Data at Home and Abroad (Postprint)

Authors: Sheng Xiaoping, Wu Tong

Date: 2023-07-26T00:00:00+00:00

Abstract

[Purpose/Significance] By reviewing research achievements on open sharing of scientific data both domestically and internationally, this study provides references for promoting the development of relevant theoretical research and practice in China.

[Method/Process] Using textual content analysis, this paper examines domestic and international research progress from seven dimensions: policies for open sharing of scientific data, platforms and projects, benefits and barriers, mechanisms and models, stakeholder roles and responsibilities, applications and surveys, and countermeasures and suggestions.

[Results/Conclusion] Currently, extensive exploration has been conducted on the topic of open sharing of scientific data both domestically and internationally, particularly yielding new insights into key issues such as policies and platforms, mechanisms and models, and benefits and barriers. However, the field as a whole remains in its initial stage, and it is still necessary to enhance the research level of open sharing of scientific data by effectively distinguishing the concepts and behaviors between open sharing of scientific data and scientific data sharing, establishing and improving mechanisms for open sharing of scientific data, and strengthening practical investigation and analysis of open sharing of scientific data.

Full Text

Preamble

Volume 63, Issue 17, September 2019
ChinaXiv Partner Journal

Review on Open Sharing of Scientific Data Across the World

Sheng Xiaoping¹, Wu Tong²

¹Department of Library, Information and Archives, Shanghai University, Shanghai 200444

²School of Economics and Management, South China Normal University, Guangzhou 510006

Abstract

[Purpose/Significance] By reviewing research results on open sharing of scientific data across the world, this paper provides a reference for promoting the development of relevant theoretical research and practice in China. **[Method/Process]** Using text content analysis method, this paper analyzes research progress across the world from seven aspects: policies, platforms and projects, benefits and obstacles, mechanisms and models, stakeholder roles and responsibilities, applications and investigations, and countermeasures and suggestions for open sharing of scientific data. **[Result/Conclusion]** Currently, the topic of open sharing of scientific data has been extensively explored across the world, particularly yielding new insights into key issues such as policies and platforms, mechanisms and models, and benefits and obstacles. However, it remains in its infancy overall, and it is necessary to improve the research level of open sharing of scientific data by effectively distinguishing between the concepts and behaviors of open sharing and traditional sharing of scientific data, establishing and perfecting the mechanism for open sharing, and strengthening investigation and analysis of open sharing practices.

Classification Number: G203

Keywords: scientific data, open sharing, research review

Introduction

Open sharing of scientific data can be traced back to the 1950s, when the concept of open access to scientific data emerged during the sharing of data collected by the International Council of Science in 1957-1958. However, subsequent research on open access to scientific data has been much less extensive than that on open access to scholarly publications [1]. Today, open sharing of scientific data has attracted significant attention. On the one hand, open access to scientific data is considered relevant to all types of data across all disciplinary fields and all researchers [2]; on the other hand, numerous policies on open sharing of scientific data have been promulgated, many initiatives or projects have been launched, and related research findings have been published both domestically and internationally.

On January 15, 2019, the authors first searched the Internet using Chinese terms “scientific data,” “research data,” or “study data” and English terms “open access to research data,” obtaining 15 relevant papers and research reports. Although some literature has separately reviewed foreign research on open scientific data [3-4] and scientific data sharing [5-6], the research theme of

open sharing of scientific data has not been effectively revealed. Therefore, this paper reviews research results across the world from seven aspects: policies, platforms and projects, benefits and obstacles, mechanisms and models, stakeholder roles and responsibilities, applications and investigations, and countermeasures and suggestions for open sharing of scientific data.

1. Research on Open Sharing Policies for Scientific Data

Research on open sharing policies for scientific data has become a focal point. Many international organizations—including the OECD, European Commission, European Research Council, League of European Research Universities, Science International, International Arctic Science Committee, Committee on Data for Science and Technology (CODATA), World Data System of the International Council for Science, Group on Earth Observations, and G8—as well as government agencies, research institutions, research funding bodies, and publishers in numerous countries such as the United States, United Kingdom, Canada, Germany, Australia, Sweden, Norway, Ireland, Finland, Denmark, Netherlands, Slovenia, and India have all formulated policies on open sharing of scientific data [7-11]. These policies can generally be categorized into three levels: macro, meso, and micro [12].

Although policies formulated by governments, research funding agencies, research and educational institutions, and academic publishers differ in policy elements, content, and implementation plans, these policy initiatives primarily focus on various stages of scientific data exchange channels [13], concentrating on data storage methods, data rights and stakeholder responsibilities, scope of open data, open access methods, data quality, data interoperability, data embargo periods, intellectual property protection, and data security. They exhibit several distinctive features: emphasizing the value of open sharing of scientific data, emphasizing intellectual property rights and their protection, emphasizing data interoperability, emphasizing high data quality, emphasizing data evaluation, and stressing stakeholder responsibility [14].

Although some literature conflates open sharing policies for scientific data with scientific data sharing policies [15] or institutional data policies [16], they are not the same concept. Open sharing policy for scientific data refers to the relevant regulations or institutional arrangements by which an organization or institution governs how to effectively implement open sharing of scientific data, sometimes also called scientific data open access policy. However, it differs from traditional scientific data sharing policies, which often require data sharing to be implemented within specific fields or limited scopes and under certain conditions, applicable only to users within a specific range and unavailable to others [10].

2. Research on Platforms and Projects for Open Sharing of Scientific Data

Today, numerous platforms for open sharing of scientific data have been established both domestically and internationally, mainly including: open registration platforms for scientific data, such as DataCite, re3data.org, Open Access Directory Data Repositories [7]91, Dataverse Network [17], and the Personal Genome Project Global Network [18]; open storage and service platforms for scientific data, such as Dryad, Figshare, PANGAEA, GBIF [7]95-100,192-193, Neurotycho.org [19], Open Spectral Database [20], Yale University Social and Policy Research Open Data Repository [21], Protein Data Bank [22], and Peking University Open Research Data Platform [23]; and open publishing platforms for scientific data, such as BioMed Central, PubMed Central, *Science*, *Nature*, and PLoS [7]102.

Existing research has extensively explored these platforms from various perspectives, including data repository structure [24], software stack design [20], technical infrastructure [7]95-100,192-193, data characteristics [19], data resources and services [22], platform construction content and application effects [23,25], and construction processes, challenges, and lessons learned [21]. Among them, open registration platforms for scientific data provide channels for individuals or institutions to submit and store scientific data in relevant fields; open storage and service platforms can implement open adoption software solutions based on open source [26], enabling both storage of scientific data and provision of browsing, retrieval, data linking, and even data correlation analysis for data papers; and open publishing platforms for scientific data build bridges for scientific data to move toward open publishing and also serve as “marketplaces” for open publishing of scientific data.

From the perspective of Indian institutional repositories, most repositories contain relatively few documents, with journal articles accounting for a high proportion, while conference papers and dissertations account for a relatively low proportion, and the numbers of preprints/working papers, teaching resources, and patents are very small [27].

Alongside platform development, numerous open sharing projects for scientific data have been implemented both domestically and internationally. Since 2007, the European Union has funded many scientific data projects, such as RECODE, ARIANDE, SERSCIDA, ODE, EURO-VO AIDA, CLARIN, CESSDA, ELIXIR [28], as well as ACUMEN, OAPEN, OpenAIRE, OpenAIREplus, PEER, and SOAP [29]. These projects mostly aim at building infrastructure for scientific data sharing and open access, seeking to promote storage, sharing, and reuse of scientific data from different perspectives. For example, the OpenAIRE project builds a collaborative open access infrastructure for European researchers, enabling them to access open access information and supporting the establishment of open access policies at the European level and providing usage statistics to policymakers [30]. The European project OpenUp achieves comparability, flex-

ibility, consistency, accessibility, and reproducibility of scientific data among member countries [31].

3. Research on Benefits and Obstacles to Open Sharing of Scientific Data

Open sharing of scientific data offers numerous substantial benefits, which also serve as motivations for stakeholders to participate in open sharing practices. However, it currently faces many obstacles.

3.1 Research on Benefits of Open Sharing of Scientific Data

Research on the benefits of open sharing of scientific data abroad began in the 1970s and 1980s [32]. Overall, open sharing of scientific data can generate numerous benefits: **Economic benefits**, including enabling researchers to obtain funding from research funding agencies when applying for certain research projects (as open sharing of research data is a prerequisite for funding some projects); achieving greater returns on public investment [33]; accelerating the transformation of research findings into products and services and bringing them to market more quickly [34]; stimulating downstream applications and commercial innovation [35]; gaining greater competitiveness [36]; and promoting economic growth while realizing the economic value of scientific data [37]. **Political and social benefits**, including enhancing government credibility, improving governance, and increasing transparency in government-funded scientific research [38]; facilitating public understanding of national scientific research [39]; providing publicly funded research results to the public [40]; promoting public participation in scientific research and improving public engagement and satisfaction [38]; and helping transfer factual data to developing countries and promote their capacity building [35]. **Technical and operational benefits**, including strengthening open scientific investigation; encouraging diversity in analysis and opinions; discovering new research topics and promoting novel research; reproducing and verifying research processes and results to improve research quality; avoiding duplication of work [41]; enabling broader access, use, and reuse of scientific data; creating more complete and transparent scientific records [42]; improving scientific transparency and accuracy to generate higher research impact [43]; accelerating the transformation and dissemination of research findings [44]; enhancing interdisciplinary, cross-sectoral, cross-institutional, and international research; helping maximize the research potential of new digital technologies and networks [35]; encouraging collaboration and improving crisis response [34]; increasing citation rates of papers and scientific data [45]; and improving the academic reputation and visibility of research institutions and researchers [46].

In summary, open sharing of scientific data can bring substantial benefits to stakeholders. However, the benefits obtained by different stakeholders may vary significantly, necessitating the establishment of benefit balancing mechanisms

[47].

3.2 Research on Obstacles to Open Sharing of Scientific Data

Obstacles to open sharing of scientific data mainly include the following aspects [36,38]: **Institutional obstacles**, such as lack of data sharing policies, risk-averse culture, absence of unified policies for data publication, only publishing non-valuable data, ineffective data promotion, revenue generation from data, and lack of incentives. **Task-related obstacles**, such as inability to discover valuable data, inability to access raw data, lack of information about open data quality, absence of indexes or other means to ensure easy discovery of correct data, users being unaware of potential data uses, data formats and datasets being too complex to process and use, lack of support tools or help desks, and particularly enormous challenges in open access to research data in digital dissertations [48]. **Usage-related obstacles**, such as discouragement of user data sharing, unclear understanding of data user needs, absence of processes for handling user input, neglect of open data opportunities, lack of time, mandatory registration before downloading data, fear of erroneous conclusions and potential negative consequences, lack of knowledge to utilize or understand data, lack of necessary capabilities to use information, lack of collaborative funding pathways between institutions, and restrictions from commercial contracts. **Legal obstacles**, such as lack of legal frameworks for data utilization, uncertainty brought by new laws, contradictory rules or laws regarding data sharing and usage rights, privacy concerns, lack of open licenses, threat of litigation, and prohibitions on commercial use. **Quality-related obstacles**, such as lack of information accuracy, incomplete information, outdated and invalid data, information redundancy, and data usability issues [49]. **Technical obstacles**, such as incompatible formats between institutions, lack of technical data infrastructure, inability to utilize crowdsourcing, absence of procedures for using user feedback, lack of metadata, lack of machine-readable formats, complexity in using open data portals, lack of programming and statistical skills, and language barriers.

The main obstacles to open sharing of scientific data in China include: lack of a clear national-level legal system, with existing policies lacking legal force; lack of effective management systems; lack of perfect platforms for data archiving, submission, and open sharing; low data quality and accessibility; existence of information blockades and barriers; many data resources held by the government remaining idle; small proportion of Chinese-language data online; market regulation of scientific data basically failing; lack of professional talent in the field of open sharing of data resources; and unwillingness or inability of data producers and managers to open and share data, necessitating further strengthening of institutional constraints and policy support in the research project funding system [50-51].

4. Research on Mechanisms and Models for Open Sharing of Scientific Data

Mechanisms and models for open sharing of scientific data are also focal issues in the field of open science research.

4.1 Research on Mechanisms for Open Sharing of Scientific Data

The mechanism for open sharing of scientific data refers to the interaction processes, methods, and operational patterns among elements of the open sharing system, including scientific data submission mechanisms, publishing mechanisms, dissemination mechanisms, alliance mechanisms, incentive mechanisms, evaluation mechanisms, and guarantee mechanisms for open sharing. Among them, the scientific data submission mechanism requires that data resources generated and managed by industry departments, research projects, and data centers supported by national fiscal funds be submitted or exchanged from the bottom up in accordance with the national “Management Measures for Scientific Data” and data submission management measures formulated by various departments. The scientific data publishing mechanism requires data to be made publicly available on the Internet, or authors to store scientific data in publicly accessible relevant databases before publishing papers, supporting researchers or organizations other than data providers to download, analyze, reuse, and cite the data [52]. The scientific data dissemination mechanism needs to increase search engine indexing, strengthen user utilization of data open platforms, use social networks to improve the dissemination speed and scope of openly shared data resources, and enhance the utilization rate of data resources to maximize the value of scientific data [52]. The scientific data alliance mechanism encourages data (platform) centers to follow jointly negotiated principles or declarations, using unified standards or interfaces (software systems) to achieve interoperability of metadata or data [52]. The scientific data incentive mechanism requires effective incentive measures to encourage data owners to open up their data and maximize the potential and energy of scientific big data. The evaluation mechanism for open sharing of scientific data requires strict peer review during the open sharing process to ensure data quality, usability, completeness, and standardization, and after open sharing, requires regulatory departments, funding agencies, and others to conduct regular or irregular investigations and supervision of open data conditions, as well as evaluate the quality, usability, value, and impact of open data [53]. The guarantee mechanism for open sharing of scientific data requires establishing and improving macro, meso, and micro-level policies, regulations, and systems, as well as technical platforms, while simultaneously building a cross-domain regional organization and service guarantee system [12]. These mechanisms can serve as strategic choices and action guidelines for domestic institutions implementing open sharing of scientific data.

4.2 Research on Models for Open Sharing of Scientific Data

Models for open sharing of scientific data involve both operational models and funding models. From the perspective of public welfare and commercial nature, there are typically two models for open sharing of scientific data within a country or region: the public welfare operation model led by government and the commercial operation model led by market forces. The former refers to a model where government departments (and their affiliated institutions) fund the production, collection, storage, and management of scientific data, and provide the data for free use and sharing to all members of society, allowing any organization or individual to copy, disseminate, and develop scientific data without restrictions from copyright agreements. This model is currently the most commonly used approach, but it also faces the dilemma of unsustainable follow-up funding. The latter refers to a scientific data sharing approach where funders and users of scientific data reach licensing agreements for paid development and utilization [12].

According to the source or scale of scientific data, models for open sharing of scientific data include [52]: **Single-source data sharing model for large scientific facilities**, which generates raw data according to data collection plans for relevant research tasks, stores it in data centers or service systems serving the large facility, prioritizes sharing among collaborators participating in the corresponding research tasks, and opens some or all data to the public after a certain period. The bottleneck of this model lies in the driving force for data open sharing and the data open policies of funding agencies. **Distributed registration model for wide-area cooperation**, which centrally registers loosely distributed data, metadata, products, and services according to unified standards and specifications, with a unified data portal responsible for connecting registered data resources to provide users with one-stop data retrieval and location services, and distributed data service systems providing data acquisition services. The bottleneck of this model is the controllability of open sharing standards and services across sub-centers. **Centralized deposit model based on data repositories**, which uses institutional repositories or relies on third-party repositories for centralized data storage and publication. The bottleneck of this model is the extremely high requirements for software and hardware environments including storage facilities, network facilities, security facilities, and management and publication services. **Scientific data publishing model**, which advocates achieving data openness through publishing forms and provides convenient channels for data reuse while implementing data quality control through publishing. The bottleneck of this model is the difficulty in determining academic evaluation systems and data intellectual property rights. **Data marketplace model**, which achieves interactive sharing for end users based on support from data centers or repositories, utilizing data trading platforms or single-entity data sources. The bottleneck of this model lies in issues such as sustainability of economic returns and data pricing.

The above models depict various operational approaches in existing open shar-

ing practices of research data, but they should not be equated with open access approaches for scientific data, which are often realized through open access data repositories, open access data journals, and data description files in open access scholarly journals [54]. However, the means for users to freely obtain scientific data [55] is not the same concept as open access approaches for research data. Furthermore, various funding models have now emerged, including government funding, consortium member funding, built-in cost funding, public-private partnership funding, charitable funding, research funding, user funding, service funding, volunteer funding, crowdfunding, social payment, and virtual currency payment [56-57], with government funding being the primary approach.

5. Research on Stakeholder Roles and Responsibilities in Open Sharing of Scientific Data

Stakeholders in open sharing of scientific data include governments, universities and research institutions, research funding agencies, academic societies and professional associations, publishers, libraries, researchers, and data professionals, who can serve as producers, funders, organizers, publishers, disseminators, managers, and users of scientific data. Each of these stakeholders is often associated with specific open access domains, open access processes, and specific data dissemination and preservation initiatives [41].

Different stakeholders can play different roles and have different responsibilities in open sharing of scientific data. Among them, governments are responsible for formulating legal and regulatory frameworks for research system operations and providing research funding; research funding agencies are responsible for formulating and implementing national research strategies and funding policies, allocating funds to various research activities; research institutions should, according to academic norms and funding guidelines, manage the implementation of publicly funded research projects by their employees, formulate policies for accessing and using research data, analyze researchers' data management needs, and ensure that each researcher receives appropriate research training; academic societies and professional associations are responsible for formulating and promoting scientific norms, values, and standards related to open sharing of scientific data; researchers should participate in formulating policies and standards for data annotation, preservation, and long-term access, comply with professional standards in their fields, establish data management plans at the beginning of each research project, promptly publish research data, methods, and other information contained in their publicly reported results to enable verification of published results, unless there are compelling reasons not to publish the data, in which case researchers should publicly explain why the data are being withheld; data professionals should be responsible for data analysis, archiving, preservation, and distribution, and participate in formulating data standards, formats, metadata, and data quality control mechanisms [35,58]. In addition, research institutions and research funders should collaborate with researchers and data professionals to develop, support, and implement plans that

meet researchers' data needs; research institutions, research funding agencies, professional associations, and journals should also promote sharing of research data through publication policies, public recognition, and funding [58]7-9.

6. Research on Applications and Investigations of Open Sharing of Scientific Data

The application domains of open sharing of scientific data are extensive, and practical investigations can further reflect the current state of open sharing.

6.1 Research on Applications of Open Sharing of Scientific Data

Open scientific data has been promoted and utilized in many fields, not only to promote scientific research and innovation [59], support verification of clinical trial data results [60], but also to implement biodiversity conservation actions [61] or conduct biodiversity conservation research [62] as well as planning of pelagic bioregionalization [63] and research on the Ross Sea [64], transform educational research models [65], enhance malaria research and control [66], and maximize the utility and public benefits of the Human Genome Project [67]. In these application practices, open sharing models do not violate specific provisions of intellectual property law; for data and related achievements that have obtained intellectual property protection, open sharing is often realized through signing licensing contracts. The open sharing model does not exclude intellectual property protection, and even when research institutions or researchers choose to open share the data and related achievements they have obtained, it does not mean they completely abandon seeking intellectual property protection [68]. However, merely implementing mandatory data disclosure policies may reduce welfare unless appropriate incentive measures are simultaneously provided to prevent delays [69].

6.2 Analysis of Investigations on Open Sharing of Scientific Data

From investigations on open sharing of scientific data, domestic open scientific data resources mainly originate from government departments, research universities, online scientific databases, and international data centers, each respectively forming characteristics of focusing on basic scientific data, focusing on disciplinary professional data with emphasis on natural sciences, focusing on Internet data, and focusing on proprietary data and partial opening of scientific data [70]. Universities, government agencies, research institutions, and municipal authorities in Finland retain tens of thousands of digital research datasets, but only a small portion is archived and reused [49], with similar situations existing in other countries. Scientists in low- and middle-income countries have growing interest in data sharing and have recognized the enormous potential of open science for the future [71].

Recent surveys by foreign scholars have found that scientists hold positive attitudes toward data sharing and reuse. Scientists acknowledge that sharing scien-

tific data can positively impact scientific progress by saving time and improving research efficiency, but when sharing their own research data, scientists express concerns, including fear of data misuse or misinterpretation. These concerns can be alleviated if proper acknowledgment and citation of used data can be ensured and original researchers are given some degree of control over data use. Additionally, helping researchers create metadata, add provenance information to datasets, and find adequate public data repositories can also help researchers adopt better data sharing practices [73].

7. Research on Countermeasures and Suggestions for Open Sharing of Scientific Data

To promote open sharing of scientific data, existing countermeasures and suggestions mainly include: Establishing and improving policies and regulations for open sharing of scientific data [74], and perfecting the top-level design of data sharing regulations [12]; Formulating universal scientific data management norms and technical standards, enhancing key technologies in all aspects of open sharing, and building a unified service platform portal [74]; Constructing an open sharing management system for research data [75], establishing public disclosure mechanisms, market regulation mechanisms, cooperation mechanisms, and guarantee mechanisms for scientific data [50,76], and exploring a path combining public welfare with marketization [74]; Establishing scientific data quality control and evaluation systems [12]; Building a culture of open sharing of scientific data [77], popularizing the social benefits of open science, not only providing citizens with the evidence needed to draw their own conclusions about scientific issues they care about, but also explaining the direct benefits of data sharing in specific work or life domains [78]; Strengthening education and training in scientific data management [12]; Strengthening intellectual property protection in the open sharing process [74]; Stakeholders (including governments, universities and research institutions, academic societies, academic and professional institutions, charitable organizations, researchers, etc.) should play their roles in open sharing of scientific data and actively promote its implementation [33]71-78; and Standardizing the operation and management of scientific data centers and accelerating the construction of a number of nationally influential scientific data centers [79].

The 2014 Belmont Forum open data survey results show that the main desires for open data stem from intrinsic research motivations, such as accelerating scientific research and applications, disseminating and recognizing research results, personal commitment to open data, and demands from data users. The most important obstacles to open scientific data are primarily the desire to publish results before releasing data, legal restrictions, loss of credit or recognition, and possible misunderstanding or misuse [72]. Vanden Eynden et al. also conducted an online survey on the benefits and obstacles of “open research” among Wellcome Trust grantees and ESRC funding recipients, finding that 80% of Wellcome Trust grantees and 66% of ESRC funding recipients had achieved open access

to datasets [46].

In summary, extensive research has been conducted on the topic of open sharing of scientific data both domestically and internationally, yielding a body of research results. However, current theories and practices of open sharing of scientific data both domestically and internationally are still in their infancy, with many deficiencies remaining.

First, from the perspective of theoretical research on open sharing of scientific data, some basic concepts and core issues have not been clearly explained, including: Failure to effectively distinguish between open sharing of scientific data and scientific data sharing, with some conflating open sharing policies for scientific data with scientific data sharing policies or government data open policies and institutional data policies, when in fact these concepts have essential differences. Future research needs to effectively define and distinguish these concepts. Lack of comprehensive and in-depth elaboration on the mechanisms of open sharing of scientific data; although some scholars have preliminarily explored this issue (see Section 4.1), there is a lack of innovative research on the dynamic mechanism, governance mechanism, and collaboration mechanism for open sharing of scientific data, and further research is still needed on guarantee and evaluation mechanisms to establish and improve the mechanism for open sharing of scientific data. Lack of investigation and analysis of practices and problems of open sharing of scientific data in China. Therefore, further strengthening of theoretical research on open sharing of scientific data is needed.

Second, from the perspective of domestic practice in open sharing of scientific data, although China's scientific data sharing services have achieved initial success, the macro-level policy system for domestic scientific data still needs expansion, policy implementation requires continued accumulation, many disciplines still need to improve data management awareness and capabilities [80], and a situation of open sharing for society has not yet been formed, with scientific data generated by various national science and technology plan projects not being effectively managed and utilized [81]. Compared with Western developed countries such as the United States and the United Kingdom, China has a clear gap in the level of open sharing of scientific data. However, China has recognized the strategic significance of scientific data sharing. The "Recommendations of the Central Committee of the Communist Party of China on Formulating the 13th Five-Year Plan for National Economic and Social Development" explicitly requires "implementing the national big data strategy and promoting open sharing of data resources." The "Management Measures for Scientific Data" promulgated by the General Office of the State Council on March 17, 2018 established the sharing principle of "openness as the norm, non-openness as the exception," and for the first time clarified the responsibilities of administrative departments of science and technology under the State Council, relevant departments under the State Council, relevant departments of provincial people's governments, research institutes, universities, enterprises as legal entities, and scientific data centers in scientific data management (including open sharing)

[82]. Therefore, China should gradually improve the systems or policies for open sharing of scientific data based on the “Management Measures for Scientific Data,” continuously enhancing the practical level of open sharing of scientific data in China.

References

- [1] Nicola, Caruso J, Archambault É. Open data access policies and strategies in the European research area and beyond [EB/OL]. [2019-01-15]. http://sciemetrix.com/pdf/SM_{{EC}}_{{OA}}_Data.pdf.
- [2] Mauthner NS, Parry O. Open access digital data sharing: principles, policies and practices [J]. *Social epistemology: a journal of knowledge, culture and policy*, 2013, 27(1): 47-67.
- [3] Hossain M A, Dwivedi Y K, Rana N P. State-of-the-art in open data research: insights from existing literature and a research agenda [J]. *Journal of organizational computing and electronic commerce*, 2016, 26(1-2): 14-40.
- [4] Huang Yongwen, Zhang Jianyong, Huang Jinxia, et al. Review of foreign open scientific data research [J]. *Modern Library and Information Technology*, 2013(5): 21-27.
- [5] Ruusalepp R. A comparative study of international approaches to enabling the sharing of research data [EB/OL]. [2019-01-15]. <https://www.era.lib.ed.ac.uk/handle/1842/3361>.
- [6] Qiu Chunyan, Huang Ruhua. New progress in international scientific data sharing in recent 3 years [J]. *Library and Information Service*, 2016, 60(3): 6-14.
- [7] Tie Mei, Gu Liping, Dong Jie, et al. Research on open access to foreign scientific data [M]. Beijing: China Financial and Economic Publishing House, 2017.
- [8] Zhang Yu'e, Wang Yongzhen. EU research data management and open access policy and its enlightenment—taking the “EU Horizon 2020” plan as an example [J]. *Library and Information Service*, 2017, 61(13): 70-76.
- [9] SPARC EUROPE. An analysis of open data and open science policies in Europe [EB/OL]. [2019-01-15]. <http://sparceurope.org/download/1654/>.
- [10] Zhang Xiaoqing, Sheng Xiaoping. Review of foreign scientific data open sharing policies [J]. *Library Tribune*, 2018, 38(8): 147-152.
- [11] Cui Yan. Research on stakeholder policies in scientific data open sharing [D]. Beijing: National Science Library, Chinese Academy of Sciences, 2016.
- [12] Wang Qing. On the operation model, guarantee mechanism and optimization strategy of scientific data open sharing [J]. *Journal of the National Library of China*, 2014, 23(1): 3-9.
- [13] Liu Xiwen, Xiong Rui. Analysis of characteristics of foreign scientific data open access policies [J]. *Information Studies: Theory & Application*, 2009, 32(9):

5-9, 18.

- [14] You Xianguang, Sheng Xiaoping. Comparison and characteristic analysis of scientific data open sharing policies of 8 international organizations [J]. *Information Studies: Theory & Application*, 2017, 40(12): 40-45.
- [15] Wen Fangfang. Research on foreign scientific data open sharing policies [J]. *Library Science Research*, 2017(9): 91-101.
- [16] Gu Liping. Policy research on open access to scientific data [M]. Beijing: Scientific and Technical Documentation Press, 2016.
- [17] Sweeney L, Crosas M. An open science platform for the next generation of data [EB/OL]. [2019-01-15]. <https://arxiv.org/ftp/arxiv/papers/1506/1506.05632.pdf>.
- [18] Haeussler M A, Greshake B, Blasimme A, et al. Open sharing of genomic data: who does it and why? [EB/OL]. [2018-10-08]. <https://doi.org/10.1371/journal.pone.0177158>.
- [19] Nagasaka Y, Shimoda K, Fujii N. Multidimensional recording (MDR) and data sharing: an ecological open research and educational platform for neuroscience [EB/OL]. [2019-01-14]. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3141074/>.
- [20] Chalk S J. The open spectral database: an open platform for sharing and searching spectral data [EB/OL]. [2019-01-15]. <https://jcheminf.springeropen.com/articles/10.1186/s13321-016-0170-2>.
- [21] Peer L, Green A. Building an open data repository for a specialized research community: process, challenges and lessons [J]. *The international journal of digital curation*, 2012, 7(1): 151-162.
- [22] Burley S, Berman H M, Christie C, et al. RCSB protein data bank: sustaining a living digital data resource that enables breakthroughs in scientific research and biomedical education [J]. *Protein science*, 2018, 27(1): 316-330.
- [23] Zhu Ling, Nie Hua, Cui Haiyuan, et al. Construction of Peking University open research data platform: exploration and practice [J]. *Library and Information Service*, 2016, 60(4): 44-51.
- [24] Abbà S, Birello G, Vallino M, et al. Shall we share? A repository for open research data in agriculture and environmental science [J]. *Bulletin OEPP/EPPO bulletin*, 2015, 45(2): 311-320.
- [25] Luo Pengcheng, Zhu Ling, Cui Haiyuan, et al. Construction of Peking University open research data platform based on Dataverse [J]. *Library and Information Service*, 2016, 60(3): 52-58.
- [26] Müller J, Heiss K, Oberhofer R. Implementation of an open adoption research data management system for clinical studies [EB/OL]. [2019-01-15]. <https://link.springer.com/article/10.1186/s13104-017-2566-0>.
- [27] Bhat M H. Exploring research data in Indian institutional repositories [J]. *Program: electronic library and information systems*, 2014, 48(2): 206-216.

- [28] Qiu Chunyan. EU scientific data open access practice and enlightenment [J]. *Information Theory and Practice*, 2016, 39(11): 138-144.
- [29] Tsimpogou F, Kouounidou V V, Prokopiou L. Open access to scientific results and data. European Union's efforts through OpenAIRE and OpenAIREplus FP7 projects: cyriot participation [EB/OL]. [2019-01-15]. <https://core.ac.uk/download/pdf/38296431.pdf>.
- [30] Rettberg N, Schmidt B. OpenAIRE—building a collaborative open access infrastructure for European researchers [J]. *Liber quarterly*, 2012, 22(3): 160-175.
- [31] Luzi D, Ruggieri R, Pisacane L. The OpenUp pilot on research data sharing, validation and dissemination in social sciences [C]//*Digital libraries: supporting open science*. Basel: Springer International Publishing, 2019: 248-258.
- [32] Stephen E F, Margaret E M, Miron L S. *Sharing research data* [M]. Washington, DC: National Academy Press, 1985.
- [33] The Royal Society. *Science as an open enterprise* [M]. London: The Royal Society, 2012.
- [34] Joseph H. The evolving U.S. policy environment for open research data [J]. *Information services & use*, 2016, 36(1-2): 45-52.
- [35] Uhlir P F, Schröder P. Open data for global science [EB/OL]. [2019-01-15]. <https://datascience.codata.org/articles/10.2481/dsj.6.OD36/gallery/367/download/>.
- [36] Stagars M. *Open data in Southeast Asia* [M]. Singapore: Palgrave Macmillan, 2016.
- [37] Houghton J, Sheehan P. Estimating the potential impacts of open access to research findings [J]. *Economic analysis and policy*, 2009, 29(1): 127-142.
- [38] Janssen M, Charalabidis Y, Zuiderwijk A. Benefits, adoption barriers and myths of open data and open government [J]. *Information systems management*, 2012, 29(4): 258-268.
- [39] Kowalczyk S, Shankar K. Data sharing in the sciences [J]. *Annual review of information science and technology*, 2010, 45(1): 247-294.
- [40] Borgman C L. The conundrum of sharing research data [J]. *Journal of the American Society for Information and Technology*, 2012, 63(6): 1059-1078.
- [41] Wessels B, Finn R L, Linde P, et al. Issues in the development of open access to research data [J]. *Prometheus*, 2014, 32(1): 49-66.
- [42] Fry J, Lockyer S, Oppenheim C, et al. Identifying benefits arising from the curation and open sharing of research data produced by UK Higher Education and Research Institutes [EB/OL]. [2019-01-15]. https://dspace.lboro.ac.uk/dspace-jspui/bitstream/2134/4600/1/JISC%20data%20sharing_{final}%20report.

- [43] Onyancha O B. Open research data in Sub-Saharan Africa: a bibliometric study using the data citation index [J]. *Publishing research quarterly*, 2016, 32(3): 227-246.
- [44] Yao Leye. Accelerating open sharing of social science data resources to promote the construction of philosophy and social sciences with Chinese characteristics [J]. *Journal of Xihua University (Philosophy & Social Sciences)*, 2018, 37(3): 68-72.
- [45] Piwowar H, Day R S, Fridsma D B. Sharing detailed research data is associated with increased citation rate [EB/OL]. [2019-01-15]. <http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0000308>.
- [46] Vanden Eynden V, Knight G, Vlad A, et al. Towards open research: practices, experiences, barriers and opportunities [EB/OL]. [2019-01-15]. <https://ndownloader.figshare.com/files/6739038>.
- [47] Guo Shilin. Research on benefit balancing mechanism of scientific data open sharing from policy perspective [D]. Harbin: Heilongjiang University, 2018.
- [48] Schöpfel J, Chaudiron S, Jacquemin B, et al. Open access to research data in electronic theses and dissertations: an overview [J]. *Library hi tech*, 2014, 32(4): 612-627.
- [49] Kuula A, Borg S. Open access to and reuse of research data—the state of the art in Finland [M]. Tampere: Finnish Social Science Data Archive, 2008.
- [50] Chen Chuanfu. Mechanism of public access to scientific data in China: characteristics, obstacles and optimization suggestions [J]. *China Soft Science*, 2004, 19(2): 8-13.
- [51] Wan Wanghui, Cui Chenzhou, Qiao Cuilan, et al. Analysis of open sharing policy and strategy for astronomical observation data [J]. *Astronomical Research & Technology*, 2015, 12(3): 364-373.
- [52] Zhu Yunqiang, Zhu Qi, Feng Zhuo, et al. Research on open sharing mechanism of scientific big data and its enlightenment to environmental information sharing [J]. *Chinese Journal of Environmental Management*, 2015, 7(6): 38-45.
- [53] Li Chengzan, Zhang Lili, Hou Yanfei, et al. Open sharing of scientific big data: models and mechanisms [J]. *Information Studies: Theory & Application*, 2017, 40(11): 45-51.
- [54] Liu Jingjing, Ma Jianhua. On three approaches to open sharing of research data [J]. *Journal of Intelligence*, 2015, 34(10): 146-150, 96.
- [55] Huang Ruhua, Liu Long, Chen Meng. Approaches to open access to scientific data [J]. *Library*, 2016(11): 26-33.
- [56] Kitchin R, Collins S, Frost D. Funding models for open access digital data repositories [J]. *Online information review*, 2015, 39(5): 664-681.

- [57] Eisfeldt-Reschke J, Herb U, Wenzlaff K. Research funding in open science [M]//Bartling S, Friesike S. Opening science: the evolving guide on how the Internet is changing research, collaboration and scholarly publishing. Heidelberg: Springer, 2014: 237-253.
- [58] Committee on Ensuring the Utility and Integrity of Research Data in a Digital Age. Ensuring the integrity, accessibility, and stewardship of research data in the digital age [M]. Washington, DC: The National Academies Press, 2009.
- [59] Xu G H. Open access to scientific data: promoting science and innovation [EB/OL]. [2019-01-15]. <https://datascience.codata.org/articles/10.2481/dsj.6.OD21/>.
- [60] Pencina M J, Louzao D M, McCourt B J, et al. Supporting open access to clinical trial data for researchers: the Duke Clinical Research Institute-Bristol-Myers Squibb supporting open access to clinical trial data initiative [J]. American heart journal, 2016, 172: 64-69.
- [61] Sullivan B L, Phillips T, Dayer A A, et al. Using open access observational data for conservation action: a case study for marine protected area network [J]. Ocean & coastal management, 2017, 148: 214-230.
- [62] Turner W, Rondinini C, Pettorelli N, et al. Free and open-access satellite data are key to biodiversity conservation [J]. Biological conservation, 2015, 182: 173-176.
- [63] Robertson L A, Lagabrielle E, Lombard A T, et al. Pelagic bioregionalisation using open-access data for better planning of birds [J]. Biological conservation, 2017, 208: 5-14.
- [64] Huettmann F, Schmidt M S, Humphries G R W. A first overview of open-access digital data for the Ross Sea: complexities, ethics, and management opportunities [J]. Hydrobiologia, 2015, 761(1): 97-119.
- [65] Gilmore R O, Adolph K E, Millman D S, et al. Transforming education research through open video data sharing [EB/OL]. [2019-01-15]. <https://files.eric.ed.gov/fulltext/EJ1106045.pdf>.
- [66] Moyes C L, Temperley W H, Henry A J, et al. Providing open access data online to advance malaria research and control [EB/OL]. [2019-01-15]. <https://malariajournal.biomedcentral.com/articles/10.1186/1475-2875-12-161>.
- [67] Pereira S, Gibbs R A, McGuire A L. Open access data sharing in genomic research [J]. Genes, 2014, 5(3): 739-747.
- [68] Wu Chunyan, Jiao Hongtao, Fan Jinde. Open sharing or intellectual property protection of human genetic data [J]. Intellectual Property, 2014(1): 55-60.
- [69] Mueller-Langer F, Andreoli-Versbach P. Open access to research data: strategic delay and the ambiguous welfare effects of mandatory data disclosure [J]. Information economics and policy, 2018, 42: 20-34.

[70] Li Zhifang, Deng Zhonghua. Analysis of distribution and characteristics of domestic open scientific data [J]. *Information Science*, 2015, 33(3): 45-49.

[71] Bezuidenhout L, Chakauya E. Hidden concerns of sharing research data by low/middle income country scientists [J]. *Journal of empirical research on human research ethics*, 2015, 10(3): 302-313.

[72] Schmidt B, Gemeinholzer B, Treloar A. Open data in global environmental research: the Belmont Forum's open data survey [EB/OL]. [2019-01-15]. <http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0146695>.

[73] Tenopir C, Christian L, Allard S, et al. Research data sharing: practices and attitudes of geophysicists [J]. *Earth and space science*, 2018, 5: 891-902.

[74] Ma Ning, Liu Zhao. Research on open sharing system of scientific data resources [J]. *China Science & Technology Resources Review*, 2017, 49(3): 1-7.

[75] Yu Fengcheng. Building an open sharing management system for research data [N]. *Study Times*, 2018-03-02(6).

[76] Shi Qin, Chu Jiewang. Research on guarantee mechanism of scientific data open sharing based on system dynamics [J]. *Journal of Intelligence*, 2018, 37(11): 143-149.

[77] Pampel H, Dallmeier-Tiessen S. Open research data: from vision to practice [M]//Bartling S, Friesike S. *Opening science: the evolving guide on how the Internet is changing research, collaboration and scholarly publishing*. Heidelberg: Springer, 2014: 213-224.

[78] Bisol G D, Anagnostou P, Capocasa M, et al. Perspectives on open science and scientific data sharing: an interdisciplinary workshop [J]. *Journal of anthropological sciences*, 2014, 92: 179-200.

[79] Yuan Yufei. Making open sharing of scientific data the norm [N]. *Guangming Daily*, 2018-04-05(2).

[80] Zhang Lili, Wen Liangming, Shi Lei, et al. Latest progress in scientific data management and open sharing at home and abroad [J]. *Bulletin of Chinese Academy of Sciences*, 2018, 33(8): 774-782.

[81] Yang Lan. Open scientific data for common development—the road of Chinese scientific data sharing [N]. *Science and Technology Daily*, 2014-04-24(12).

[82] General Office of the State Council. Notice of the General Office of the State Council on issuing the Management Measures for Scientific Data [EB/OL]. [2019-01-15]. http://www.most.gov.cn/mostinfo/xinxifenlei/fgzc/gfxwj/gfxwj2018/201804/t20180404_{139023}.htm

Author Contributions

Author Contributions: Sheng Xiaoping: paper writing and revision; Wu Tong: data collection and analysis.

Journal Information

Library and Information Service is a large, authoritative academic journal with over 60 years of history in library and information science and related fields, dedicated to publishing and exchanging theoretical, technical, methodological, and applied innovative achievements in library science, information science, and related interdisciplinary subjects. We welcome submissions of innovative research papers with theoretical contributions or application value that are thoughtful, creative, methodological, and evidence-based.

2019 Topic Guidelines (including but not limited to): 1. Research on the development of China's library and information cause in the 70 years since the founding of the People's Republic of China 2. Functions and influences of library societies (associations) in library development 3. Pre-research on China's library and information cause "14th Five-Year Plan" 4. Role positioning of libraries in the Open Science era 5. Functions and characteristics of library scientific communication in the new media era 6. Role of libraries in reconstructing academic communication systems 7. Artificial intelligence and smart libraries/smart services 8. Research on library laws and related laws in China and abroad 9. Theory and practice of library embedded services 10. From information literacy education to innovation literacy education 11. Resource organization and services across LAM (libraries, archives, museums) fields 12. Impact and effectiveness evaluation of new library construction and space recreation 13. Research and practice of libraries in promoting transformation of scientific and technological achievements 14. Key issues in next-generation institutional repository construction 15. Characteristics and requirements of library data resource construction 16. Construction of new-generation library systems driven by data 17. Innovation and application of information science theory and methods 18. Intelligence system reform under the overall national security concept 19. Innovation in intelligence analysis theory and methods 20. Intelligence service capabilities from a big data perspective 21. Library and information science in think tank construction and services 22. Think tank services and decision-making consulting service capability construction 23. Theory and method system of computational intelligence 24. Technology and methods of data management and services 25. Data governance and national intelligence security strategy 26. Intelligence sharing mechanism in military-civilian integration 27. Micro-mechanisms and macro-phenomena of information behavior 28. Regional and industrial intelligence service models and mechanisms 29. Utilization and value assessment of multi-source information resources 30. Altmetrics theory and practice research

Library and Information Service Magazine Office
December 2018

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

Source: ChinaXiv — Machine translation. Verify with original.