



Research on the Ecosystem Construction of Scientific Data Sharing in the Open Science Environment

LI Mengru*

Beijing Academy of Science and Technology, Beijing, China

limr2021@163.com

Abstract. With the advent of big data era, open science has gradually become the development trend of global scientific research. Under the background of open science, the construction of scientific data ecology has become an important means to promote scientific research and technological innovation. Based on open science, with scientific data sharing as the core, this paper aims to explore the connotation, development status, challenges and solutions of scientific data ecological construction. Through combing and analyzing the status quo of scientific data ecology construction at home and abroad from four aspects: management mechanism, legal guarantee, technical support and social cognition, it is found that there are still many problems and challenges in its development. Finally, it puts forward countermeasures and suggestions for the construction of scientific data sharing ecology from many aspects, and provides references for further promoting the implementation of open science in our country and improving the social participation and management level of scientific data sharing.

Keywords: open science; open data; data sharing; data sharing ecosystem

1 Introduction

With the acceleration of the construction of national innovation system, the rapid growth of scientific research input, the vigorous development of scientific and technological personnel, the development and popularization of open science, scientific data sharing has been paid more and more attention. The development of network communication, cloud storage and other technologies in open science environment opens up new opportunities for the research of trusted data storage, data sharing and data reuse. The public demand for open science, open sharing of scientific research data and information, and access to credible research results is more urgent. At present, promoting the development of scientific data sharing has more realistic needs. It is of great significance to promote the transparency and quality of scientific research, improve the efficiency and credibility of scientific research, and accelerate the opening of results and social application.

In order to build a perfect ecological system of scientific data sharing, it is necessary to explore the sharing mode of scientific data, formulate the norms of scientific data sharing, establish the sharing platform and service institutions of scientific data, and

improve the security and credibility of scientific data sharing. The purpose of these research works is to break the "Information Island" of scientific data, promote the exchange and cooperation of scientific research, and achieve a deeper and more extensive development of scientific research^[1].

2 The concept and practice of scientific data sharing

Under the guidance of the open scientific value system, in order to cope with the problems and challenges in the process of scientific research and promote the development process of scientific data sharing, a series of reform measures for scientific data sharing have been carried out in succession, starting from Europe and expanding, in order to improve the accessibility of science, improve the scientific research process and value realization chain, and improve the scientific evaluation system^[2].

2.1 Scientific data sharing concept

Scientific data refers to the activities of science and technology or obtained by other means to reflect the essence of the objective world, characteristics and change rules of basic data and the original system according to different activities of science and technology require processing all kinds of data sets.

The strategic position of scientific data in science and technology innovation has been continuously enhanced, and it has become an important part of the global science system infrastructure, and is also considered to be an important strategic resource for economic and social development and an important commanding point for national science and technology innovation and competition. Scientific data sharing has reached an unprecedented level in both scale and degree^[3].

2.2 Trends of scientific data sharing at home and abroad

Scientific data sharing has become a trend, and more and more international organizations begin to issue relevant policy guidance, as shown in Table 1. Many governments and international organizations, such as UNESCO, OECD and WHO, have also issued policy guidance on scientific data sharing to encourage countries to strengthen data management and sharing.

Table 1. Implementation trends of representative scientific data sharing in foreign countries

Year	Country/Area/Organization	Policy/Action
2002	EU	Budapest Open Access Initiative
2003	OECD	Policy Guide for the Development of Digital Science and Technology
2006	US	Federal Research Public Access Act
2007	OECD	Policy Principles for Global Science and Technology Development

2010	Germany	Grundsätze zum Umgang mit Forschungsdaten
2011	EU	EU Digital Agenda 2011-2020 Strategy
2012	ALLEA	Open Science for the 21st Century
2012	US	Big Data Research and Development Initiative
2013	EU	Policy recommendations for open access to research data in Europe
2013	UK	Seizing the data opportunity: a strategy for UK data capability
2013	US	Fair Access to Science and Technology Research Act
2013	US	Open Data Policy-Managing Information as an Asset
2014	Finland	The Open Science and Research Roadmap2014—2017
2015	EU	Horizon 2020 plan
2016	EU	Open Science Policy Platform, OSPP
2016	US	The Federal Big Data Research and Development Strategic Plan
2017	Netherlands	National Plan Open Science
2018	LIBER	LIBER Open Science Roadmap
2018	EU	EU Open Science Strategy
2018	ERC	Plan S
2018	US	Open science by design: Realizing a vision for 21st century research
2019	US	Federal Data Strategy Action Plan
2020	EU	A European strategy for data
2021	UNESCO	UNESCO Recommendation on Open Science
2021	SE	Practical Guide to Sustainable Research Data

With the rapid development of science and technology, the significance of scientific data management and open sharing is more and more important. We also starts to gradually establish and improve the policies and regulations of scientific data management and open sharing, as shown in Table 2.

In June 2018, Measures for the Management of Scientific Data were officially released, which for the first time set out comprehensive regulations on the management, preservation, utilization and open sharing of scientific data^[4].

In September 2019, Ministry of Science and Technology issued the Opinions on Strengthening the Management of Science and Technology Innovation Data, which clearly promotes the pilot of the Open Science Program (OSP), establishes the working system of science and technology innovation data management, and improves the system of policies and regulations on science and technology innovation data management.

Table 2. Implementation trends of representative scientific data sharing in China

Year	Organization	Policy/Action
2002	China's Ministry of Science and Technology	National Scientific Data Sharing Project
2004	Chinese Academy of Sciences/ National Natural Science Foundation of China	Berlin Declaration on Open Access
2006	General Office of the State Council	Outline of the National Program for Medium - and Long-Term Scientific and Technological Development (2006-2020)
2006	China Seismological Bureau	Measures for the Administration of seismic Science Data Sharing

2008	National Earth System Science Data Sharing Platform	Charter of Earth System Science Data Sharing Platform
2014	Chinese Academy of Sciences/ National Natural Science Foundation of China	Chinese Academy of Sciences policy Statement on Open access to papers published in publicly funded research projects
2015	General Office of the State Council	Action Plan to Promote Big Data Development
2016	National Earthquake Science Data Sharing Center	Data sharing Standard
2016	General Office of the State Council	Interim Measures for the Administration of Government Information Resources Sharing
2018	General Office of the State Council	Scientific Data Management Measures
2018	Ministry of Science and Technology/Ministry of Finance	Measures for the Administration of the National Science and Technology Resource Sharing Service Platform
2018	World Congress for the Promotion of Public Science Literacy	Beijing Declaration on the Promotion of Scientific Literacy among the World Public
2018	National Natural Science Foundation of China/National Science and Technology Library	Support for Plan S
2019	International Council for Science Data Committee	Beijing Declaration on Scientific Research Data
2019	Ministry of Science and Technology/Ministry of Finance	Notice on the implementation of the Data Sharing Support Plan for Major Scientific and Technological Innovation Projects
2019	China's Ministry of Science and Technology	Opinions on Strengthening the Management of Scientific and Technological Innovation Data
2019	Chinese Academy of Sciences	Measures for the Management and Open Sharing of Scientific Data of the Chinese Academy of Sciences (Trial)
2021	National People's Congress	Outline of the 14th Five-Year Plan for National Economic and Social Development of the People's Republic of China and the Long-range Goals for 2035
2021	Ministry of Science and Technology	Technical and Management Standards of Scientific Data Exchange for the Formation of Science and Technology Plans
2021	General Office of the State Council	Science and Technology Progress Law

3 Current situation and dilemma of scientific data sharing development

With the wide application of big data in scientific research, the orderly organization and effective utilization of scientific data will no longer be determined by a single data management principle and collection and storage methods. Data practice is affected by multiple individual, institutional, disciplinary and infrastructure factors. Therefore, there is still a gap between the openness and development level of scientific data sharing in China and that in the world. Influenced by the inherent development paradigm of science and the conflicts among multi-stakeholder bodies, we are faced with many challenges. It is not easy to realize the mission of scientific data sharing^[5].

3.1 The absence of management system

At present, developed countries have basically formed a clear and cohesive data division management organization structure among governments, scientific research institutions and universities. In China, the overall scientific data management system is not perfect, and the responsibilities, authority and business boundaries between different data management subjects are not clear. To some extent, there are phenomena such as the unsmooth organizational system of data management, the absence of data supervision by the government and project sponsors, and the weak data management and support services.

The specific manifestations are as follows: First, the lack of specialized institutions or organizations to be responsible for the management and coordination of scientific data sharing. Second, the lack of unified scientific data sharing standards and technical specifications is another problem. Third, the lack of scientific data trust mechanism and reward and punishment system. Fourth, the management of business organizations is not standardized.

3.2 The inadequacy of legal safeguards

Perfect government information laws and regulations provide sufficient system basis for scientific data sharing management. Compared with western countries, where government departments, research project funding institutions and universities have formulated their own scientific data management policies, China has not formed data management policies at the level of the state, research institutions and universities.

The specific manifestations are as follows: First, lacking of systematic laws and regulations system. Second, laws and regulations have no clear provisions on the content and authority of scientific data sharing. Third, there is no clear definition of personal privacy protection in scientific data sharing^[6].

3.3 The complexity of technical support

Scientific data sharing requires the establishment of efficient data storage, management and exchange platform, big data processing and analysis capabilities, as well as safe and reliable data transmission and storage technology.

First of all, there is a great variety of scientific data, involving multiple subject areas and various data types. Second, the scale and complexity of scientific data is also increasing, requiring the adoption of technologies such as big data and artificial intelligence to process and manage. In addition, the security and privacy of scientific data is also an important issue, involving data encryption, access control, data sharing protocol and other aspects. Although the national level has built a number of scientific data sharing platforms, such as the National population and Health science data sharing platform and the National Earth system science data sharing platform, these platforms have a single type of data resources and a low degree of integration and sharing.

3.4 The imbalance of social cognition

Since the policy of open and sharing scientific data has not been widely agreed, and the sharing mechanism still has many incomplete links, it will take time for the public to accept open science and fully share its data and results.

On the one hand, restricted by privacy, security, ethics, politics and other factors, many researchers not only have a low awareness of data sharing, but also have a low perception of the scientific research data management system and its data services. On the other hand, the lack of unified management and standards leads to the confusion of scientific data management, unpurified data classification and arrangement, information loss and data tampering in the process of data collection and processing. In addition, there is a lack of motivation and incentive mechanisms for scientific data sharing, leading to underutilization of its potential^[7].

4 Countermeasures and suggestions for constructing scientific data sharing ecosystem

After years of efforts, China has made great progress in the open sharing and management of scientific data, gradually deepening the understanding of scientific data sharing, and has achieved initial results in promoting the open sharing of scientific data, the construction of scientific data centers, identification and publication of scientific data, but it should be noted that there are still a lot of work to be carried out in the sharing and management of scientific data. In order to further promote the implementation of open science and improve the level of social participation and management of scientific data sharing, relevant countermeasures and suggestions are put forward from the following four aspects.

4.1 Establish a scientific data sharing and management service mechanism

Sharing and management service mechanism is a very important part in constructing a perfect scientific data sharing ecosystem. On the one hand, it can clarify the status and function of scientific data sharing; On the other hand, it can promote the standardization and controllability of scientific data sharing. We can learn from the sharing experience and tactics of international organizations, clarify sharing mechanism, formulate perfect sharing policies and regulations guarantee system, form scientific data management and sharing service system which is "reasonably structured, oriented to the whole society, network and intelligent", promote the sharing work for a long and stable development.

First, establish and improve scientific data management norms and standards, formulate common data standards, data classification, and data quality evaluation standards. Secondly, the rules and procedures of data sharing should be formulated. Finally, a national scientific data sharing platform should be established to provide a unified platform for data storage, sharing and inquiry for researchers and institutions, so as to promote the convenience and efficiency of scientific data sharing.

4.2 Strengthen policy guidance and legal guarantee

In terms of policies and regulations, only by improving relevant policies, laws and regulations, clarifying the legal status and responsibilities of data sharing, and establishing the legitimacy and security of data sharing, can the legitimate rights and interests of data providers and data users be protected and a good atmosphere of scientific data sharing be created for the whole society.

Firstly, Improve legislative guarantee. The government should strengthen legislation, establish a legal system, policies and regulations related to scientific data sharing. Secondly, strengthen protection of intellectual property rights. The ownership and protection of intellectual property rights should be clearly defined, and an intellectual property protection mechanism should be established to prevent unnecessary losses and infringements of scientific researchers and research institutions in the process of data sharing. Thirdly, attaches great importance to privacy protection. In the process of data sharing, it is necessary to clarify the protection measures of personal privacy information to ensure that personal privacy information is safe and not abused. Fourthly, clarify reward and punishment systems. The crackdown on illegal infringement and misuse of scientific data will be strengthened and active participation in data sharing and contributors will be encouraged and rewarded.

4.3 Improve scientific data security protection

In the face of international scientific data governance situation, China urgently needs to further improve data security protection to ensure data security and privacy.

When it comes to technical support and security protection measures for scientific data sharing, there are several key aspects to consider: Firstly, data format standardization; Secondly, data storage and processing; Thirdly, construction of data access and sharing platform; Fourthly, data protection and security. Only by solving the problems of autonomy and security in data sharing through technological means can we ensure the effectiveness and sustainability of scientific data sharing, carry out more reciprocal international data exchange and sharing, and promote the convergence and communication with international standards.

4.4 Improve public's scientific literacy and awareness

At present, the scientific quality of Chinese citizens needs to be improved. The public has strong demand for the acquisition and participation of scientific and technological information. Chairman Xi pointed out that "Scientific and technological innovation and popularization of science are the two wings of innovation and development. We should place the popularization of science in an equally important position as scientific and technological innovation."

The following measures can be taken: First, strengthen publicity and education; Second, enhance social trust in scientific data sharing; Third, strengthen audience participation; Fourth, encourage collaboration on scientific data sharing: including across dis-

ciplines, institutions and countries. Through the above measures, more people can understand scientific data sharing, establish the values and concepts of data sharing, and lay a solid social cognitive foundation for the construction of the ecological system of scientific data sharing.

5 Conclusions

In the context of open science, scientific data has become an important basic strategic resource for scientific research and innovative discovery. It has become a common consensus of the international community to strengthen the open sharing of scientific data and ensure the security of scientific data. Compared with the developed countries in Europe and the United States, our scientific data sharing and management practice and policy-making start late, and there are many space to improve and improve. We should make policy and institutional efforts to properly handle prominent problems and relationships in scientific data sharing, strengthen the practice of scientific data sharing, improve the system of data standards, raise the level of scientific data sharing and utilization, explore the ecosystem construction of scientific data sharing, and promote scientific data to support scientific and technological innovation and economic and social development.

Foundation

BJAST Program “Evaluation and Spillover Effect of Science and Technology Innovation Efficiency Based on Innovation Value Chain (23CB074).

References

1. Joseph H., Building momentum to realign incentives to support open science. *Data Intell*, 2021, 3: 71–78.
2. United Nations Educational, Scientific and Cultural Organization (UNESCO). UNESCO Recommendation on Open Science. 2021, <https://unesdoc.unesco.org/ark:/48223/pf0000379949?3=null&queryId=65f84ab3-e6ec-4dd7-abe0-cc839b1a559c>, 2023-04-25.
3. European Commission. European Open Science Cloud (EOSC). 2022, <https://ec.europa.eu/info/research-and-innovation/strategy/strategy-2020>, 2023-04-25.
4. 2024/our-digital-future/open-science/european-open-science-cloud-eosc_en.
5. European Commission. European Open Science Cloud. 2022, <https://digital-strategy.ec.europa.eu/en/policies/open-science-cloud>, 2023-04-25.
6. UK Research and Innovation. Common Principles on Data Policy [EB/OL]. <http://fff8ed76bbfb4194ad1adad59591f163e1asokn5qnoqvfnu6x5n.fgy.hlju.cwkeji.cn/funding/information-foraward-holders/data-policy/common-principleson-data-policy/,2023-04-25>.
7. Koundouri P, Chatzistamoulou N, Dávila O G, et al. Open access in scientific information: Sustainability model and business plan for the infrastructure and organization of Open AIRE. *J Benefit Cost Anal*, 2021, 12: 170–198.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits any noncommercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

