Development of International Open science and Data Sharing in Big Data Era

Dong Ning

Shandong Technology & Business University, Yantai 264005, China

673077003@qq.com

Keywords: Open science, Data Sharing, Scientific data, Data librarian, Librar

Abstract: By tracking the development of policies and contents of international open scientific data, and the progress of library scientific data services, this paper summarizes the scientific data sharing practices in various international fields, so as to provide reference and guidance for the scientific data sharing practice in China.

1. Introduction

Nowadays, the world has entered the era of big data. The development and research of scientific activities and global academic exchanges are developing towards an increasingly open and open trend. Open science is universally recognized by the scientific community. The management and sharing of scientific research data has become the important task for the international organizations, scientific research institutions and research libraries. On November 10, 2015, the National Information Standards Organization (NISO) and the International Council for Scientific and Technical Information (ICSTI) convened a webinar on the theme "From Open Access and Data Sharing to Open Scientific Practice" to explore the development of data sharing to open science.

This paper intends to track the development of international open science and data sharing in recent years, and summarize its characteristics, in order to provide suggestions for the development of scientific data sharing practice in China.

2. The development of the policies of open science and scientific data sharing in the scientific community

2.1 Trends in open access, open data and open science

With the development of the peer review system, open access, open data, and open science have been unanimously recognized by the global scientific community. The EU, the United States, and Japan have introduced corresponding open science strategies. The Integrated Innovation Strategy released by the Cabinet Office of Japan on June 15, 2018 pointed out that the use of big data, artificial intelligence, cloud computing, etc., to achieve cross-border integration and implementation of data-driven development strategy based on data sharing platform. The strategy is to support the implementation of the Society 5.0 National Development Strategy [1]. On July 17, 2018, the National Academies released the "Open Science by Design: Realizing a Vision for 21st Century Research"[2], which organized academic experts to discuss policy recommendations to American society and government: The foundation of open science is the open access to scientific papers and the opening of scientific data. Sharing requires the full use of existing big data sharing platforms and measures to promote more scientific researchers to engage in scientific data management plans and scientific data sharing, and to combine data science and public science to create new knowledge and scientific value for the social economy.

The promulgation of policies from more and more countries and higher and higher levels fully demonstrates that open science has received strong support from the scientific community and even government policies. Open science has risen to the national strategic perspective

2.2 Data property rights policy guarantees the legality of scientific data sharing

Under the trend of open science development, the sharing of scientific data has been unanimously recognized by researchers and scientific research institutions. However, scientific data involves multiple stakeholders in the process of data transfer. The legality of data ownership and use needs to be clearly defined and standardized. This requires clear data rights. Data interest means that under the cooperation agreement, there is no disclosure, use or re-enforcement restriction between the data provider and the recipient to exchange data, unless otherwise specified in other terms [3].

On April 26, 2018, the European Union's official website published a report entitled "Study on Data Sharing between Companies in Europe"[4]. According to the report, the European Commission is committed to the development of the European data economy. China's network security Law (2017) and Data Management Measures (2018) also have provisions similar to GDPR, both of which are guided policies rather than specific areas. Our scientific research units, universities and other stakeholders should cooperate and explore legal provisions for the protection of their respective data rights.

3. Scientific data has become the focus of international attention in many fields

Scientific data is the achievement of scientific research, and the important foundation of scientific research as well. The interdisciplinary development of scientific research has become the focus of scientific data in various related fields. "The Open Data: The Engine of Innovation, Growth and Transparent Governance " issued by the European Commission's Horizon 2020 Strategy requires the EU and its members to establish relevant legal mechanisms and adopt corresponding fiscal measures to promote cooperation of open data[5].

The World Bank released a statement on data innovation for open data, public-private partnerships in data, and data literacy in January 2015[6]. In the field of cultural education, scientific data sharing has also become an important measure to promote the development of scientific research. In November 2014, the Ministry of Education and Culture of Finland issued "The Open Science and Research Roadmap 2014-2017", which clearly states that Finland will be the leader in the field of open scientific research worldwide after three years, and points out that the research results produced by publicly funded research projects, especially scientific data, should be open and shared [7].

4. Policy Supports of Scientific Research Funding Institutions for Scientific Data Sharing

The University of Cambridge Scientific Data Management Working Group summarizes 20 research data policies of major University research funders. They found that funders have detailed regulations on data sharing, data openness, and accessibility [8]. The Organization for Economic Co-operation and Development (OECD) has developed OECD principles and guidelines for access to research data from public funding [9]. The National Science Foundation (NSF), the National Institutes of Health (NIH), and the Medical Re-search Council (MRC) have all put forward corresponding policy requirements for the open sharing of scientific data for their funded research projects.

From the perspective of scientific data sharing, the research funding agencies' policies cover a wide range of disciplines, and provide a good environment and atmosphere for scientific data sharing and management. From the perspective of scientific research institutions and researchers, research funding agencies the introduction of scientific data related policies provides a guide to action for the implementation of scientific data sharing practices.

5. Scientific research institutions attach importance to scientific data

In December 2013, the League of European Research Universities (LERU) published the Roadmap for Research Data, which clarified the benefits and challenges of scientific data sharing and reuse, the cost of scientific data development, and the management of scientific data. Skills and knowledge required, as well as corresponding training and development opportunities, etc. The roadmap guides LERU member universities to implement scientific data management practices in the form of guidelines.

6. Research libraries have gradually become the main force for scientific data management and sharing

Since National Science Foundation (NSF) proposed mandatory policy requirements for scientific data management programs in research funding projects in 2011, research libraries, especially university libraries, have participated in scientific data. University libraries have taken scientific data management, sharing and data librarian training as their own development priorities.

6.1 The attention of the library community to scientific data management services

The Library Association actively promotes the research and project development of scientific data management. Take the American Library Association as an example. American Association of University and Research Libraries (ACRL), member of ALA, released "Environmental Scan 2015" in March 2015, stating that scientific data services have become the focus and hotspot of university library services [10].

6.2 University Libraries focus on scientific data management services

In 2013, Research and Markets, a market research company, surveyed research university libraries from the United States, the United Kingdom, Australia, India, South Korea and other countries. According to the survey results, 63.33% of the surveyed libraries provided scientific data management plan consulting services, and the proportion of this service in the United States reached 72.22%.

6.2.1 Progress in Library Scientific Data Service System

Take the scientific data service of the University of Illinois Urbana Champaign Library as an example to introduce the scientific data service content provided by the university library. The scientific data service (or, research data services, RDS) of the library is based on the data life cycle, mainly from the three aspects of data management, data organization, data storage and backup, providing professional knowledge and professional tools for all researchers in the university. The RDS provided by the library helps researchers to promote the planning, organization and management of scientific data and the smooth progress of the entire research activities. The scientific data service of the University of Illinois Urbana Champaign Library is jointly sponsored by the National Library of the United States, the National Supercomputer Application Center, the school library, and the Graduate School of Library and Information Science. It brings together scientific data management, data processing, and experts in various disciplines to form research data. Management plan, scientific research data organization, scientific research data preservation and sharing, scientific research data management training four parts of scientific research data service system [11].

6.2.2 Emerging role of library data librarians

With the development of digital environment, librarians assume an emerging role as data librarians to assist scientific researchers in scientific research. In the 2009 ARL (Association of Research Libraries) report, librarians were positioned as the last mile of scientific data infrastructure construction, which illustrated librarians should act as a bridge and link between systems, researchers, and data users in scientific research [12]. Based on the services provided by librarians

in the life cycle of scientific research, some scholars have suggested that data librarians can play the following roles: data manager, data engineer, data steward, data guardian, data scientists, etc.

American Library Association (ALA) is developing a core competence framework for this new profession of data librarians. During the 2018 International Library Congress (IFLA), there were reports suggesting that libraries and information science (LIS) should open courses for data librarians to quickly promote specialization in this field [13]. In recent years, Europe, the United States, Japan and other countries have achieved developments in open science and scientific data sharing. The progress is reflected in the government, international organization policy formulation, research funding agencies, scientific research institutions, research libraries in the scientific data sharing and service, management exploration and innovation.

7. Conclusion

At present, China's outline for promoting big data development clearly proposes to develop scientific big data, and actively promote the gradual open sharing of scientific data acquired and generated by public welfare research activities supported by the state public finance. The scientific community and the library community have begun to study the theory, policy, and norm-setting work in the fields of data openness, scientific data sharing and management. We can assimilate from the experience of international scientific data sharing and management, enhance the participation of different entities such as government, scientific research institutions, libraries, etc., and actively explore the development of open sharing of scientific data suitable for China.

References

[1] Cabinet House of Japan. (Cabinet Office of Government of Japan. Integrated innovation strategy.2018. http://www.cao.go.jp/cstp/tougosenryaku/index.html.

[2] National Academies. Open science by design: realizing a vision for 21st century research.2018. https://www.nap.edu/catalog/25116/open-science-by-design-realizing-a-vision-for-21st-century.

[3] Fu Wei, Yu Changyue. 2017. A Review of Research and Dynamic Analysis of Development on Data Rights at Home and Abroad, *Journal of Modern Information*, Vol. 7.(159-165).

[4] Study on Data Sharing Between Companies in Europe. 2018. https://publications.europa.eu/en/publication-detail/-/publication/004fcf02-49c7-11e8-be1d-01aa75ed71a1/language-en/format-PDF/source-69800191.

[5] Guidelines on open access to scientific publications and research data in horizon 2020.2015. http://ec.eu-ropa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-pilot-guide_en.pdf.

[6] The Next Step of the Data Revolution: Financing New Key Points. 2015. http://blogs.Worldbank.org/opendata/ch/next-step-data-revolution-financing-emerging-priorities

[7] The open science and research roadmap 2014-2017.2015.http://www.minedu.fi/export/sites/default/OPM/Julkaisut/2014/liitteet/okm21.pdf?lang=en.

[8] Research data policies of major University research funders. 2016. http://www.data.cam.ac.Uk/funders.

[9] OECD principles and guidelines for access to research data from public funding.2015. http://www.oecd.org/science/sci-tech/38500813.pdf.

[10] Environmental scan 2015. http://www.ala.org/acrl/sites/ala.org.acrl/files/content/publications/whitepaper/EnvironmentalScan15.pdf.

[11] Liu Ying, Wu Ming, Hu Hui, Chen Xiujuan. 2017. Research on the Research Data Service of Foreign University Libraries: Taking the University of Illinois at Urbana-Champaign Library as an Example. *New Century Library*, Vol. 6 (69-74).

[12] Gabridge T. 2009. The last mile: Liaison roles in curating science and engineering research data, *Research Library Issues: A Bimonthly Report from ARL, CNI, and SPARC*, Vol. 265 (15).

[13] Khan H R, DU Y F. 2018. What is a data librarian? A content analysis of job advertisements for data librarians in the United States academic libraries. http://library.ifla.org/id/eprint/2255.