# Stakeholders' Responsibilities and Roles in Open Scientific Data Sharing: An Analysis Based on International Organization Policies - Postprint

**Authors:** Sheng Xiaoping, Wang Yi

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

## Abstract

[Purpose/Significance] By analyzing the responsibilities and roles of different stakeholders defined in international organizations' scientific data open sharing policies, this study provides references for improving China's "Administrative Measures for Scientific Data" and implementing scientific data open sharing. [Method/Process] Based on web surveys and text analysis, it discusses the responsibilities and roles of different stakeholders defined in international organizations' scientific data open sharing policies in scientific data open sharing activities. [Results/Conclusion] International organizations' scientific data open sharing policies define numerous stakeholders including government, researchers, research institutions, research funding agencies, libraries or archives, data centers, publishers, professional associations or societies, users, enterprises, etc., and their different responsibilities and roles in scientific data open sharing activities. China should learn from these experiences to formulate domestic scientific data open sharing policies and further enhance the level of scientific data open sharing.

## Full Text

## Preamble

**ChinaXiv Partner Journal**
Volume 63, Issue 17, September 2019

**Stakeholders' Responsibilities and Roles in the Open Sharing of Scientific Data: An Analysis Based on International Organization Policies**

Sheng Xiaoping, Wang Yi
Department of Library, Information and Archives, Shanghai University, Shanghai 200444

## Abstract

[**Purpose/Significance**] This paper analyzes the responsibilities and roles of different stakeholders defined in international scientific data open sharing policies to provide references for improving China's *Measures for the Management of Scientific Data* and implementing scientific data open sharing. [**Method/Process**] Based on web surveys and textual analysis, this paper discusses the responsibilities and roles of different stakeholders in scientific data open sharing as defined by international organization policies. [**Result/Conclusion**] International scientific data open sharing policies define numerous stakeholders—including governments, researchers, research institutions, research funding agencies, libraries or archives, data centers, publishers, professional associations or societies, users, and enterprises—and their distinct responsibilities and roles in scientific data open sharing activities. China should draw on these experiences to formulate domestic scientific data open sharing policies and further enhance the level of scientific data open sharing.

**Classification Number:** G203
**Keywords:** scientific data, open sharing, stakeholder, responsibility and role, international organization

The open access movement has promoted the global dissemination of knowledge and the sharing and reuse of open data, receiving increasing attention [1]. In recent years, scholars both domestically and internationally have begun to focus on stakeholders such as researchers, research institutions, funding agencies, publishers, and users regarding their responsibilities and roles in scientific data open sharing [2-4]. However, scientific data open sharing also involves other stakeholders like governments, libraries or archives, and data centers. On March 17, 2018, the General Office of the State Council promulgated the *Measures for the Management of Scientific Data*, which clarified the main responsibilities of government departments, research institutes, higher education institutions, enterprises, scientific data centers, and other legal entities, and required that scientific data formed by government budget funding should follow the principle of "open as the norm, non-open as the exception" to be shared with society and relevant departments.

## 1. Stakeholders Defined by International Organization Scientific Data Open Sharing Policies

Stakeholders in scientific data open sharing activities refer to individuals or organizations related to scientific data open sharing. As early as December 2006, the Organisation for Economic Co-operation and Development (OECD) issued the *OECD Principles and Guidelines for Access to Research Data from Public Funding*, setting a precedent for international organization scientific data open sharing policy formulation. Subsequently, other international organizations such as the International Arctic Science Committee (IASC), Group of Eight (G8), Group on Earth Observations (GEO), League of European Research Universi-

ties (LERU), European Research Council (ERC), and European Commission (EC) successively formulated scientific data open sharing policies (see ). These policies cover many aspects of scientific data open sharing, such as the scope of open data, data management plans, data formats and standards, data storage, data validation, data evaluation, data citation, data reuse, data interoperability, data security and curation, metadata, data repositories, embargo periods, fees or funding, intellectual property protection, and data access methods [5], as well as defining the responsibilities and roles of numerous stakeholders including governments, researchers, research institutions, research funding agencies, libraries or archives, data centers, publishers, professional associations or societies, users, enterprises, and society (see ).

## 2. Analysis of Stakeholders' Responsibilities and Roles

Although society can benefit from scientific data open sharing, since society is not a behavioral entity, it is difficult to define its responsibilities in scientific data open sharing. Therefore, the following analysis focuses on the responsibilities and roles of ten types of stakeholders in scientific data open sharing.

### 2.1 Government Responsibilities and Roles

Governments play an extremely important role in scientific data open sharing, serving as both leaders in open scientific data policy formulation and as funders, producers, publishers, disseminators, managers, and users of open scientific data. Their main responsibilities and roles are:

First, formulate national or regional government scientific data open sharing policies and standards, fulfilling the role of policy maker and leader. The *G8 Open Data Charter* fully demonstrates the government's leading role in formulating scientific data open sharing policies by clarifying five principles of open data, 14 priority open fields, and three common action plans (especially the action plans of G8 countries) [12]. In fact, following the release of the *G8 Open Data Charter* in June 2013, the UK government issued the *G8 Open Data Charter UK Action Plan 2013* in November 2013, making six commitments: (1) the UK will publish high-value datasets identified in the *G8 Open Data Charter*; (2) ensure all datasets are published through the national data portal data.gov.uk; (3) identify priority datasets for publication through communication with society, institutions, and the public; (4) support domestic and international open data innovators by sharing experiences and tools; (5) set a clear direction for the UK's open data work, with all government departments updating their departmental open data strategies by June 2014; and (6) establish a national-level information infrastructure for government data [18]. These six commitments essentially became the UK's action guide for implementing open data.

Second, fund scientific data open sharing activities by establishing national and local government funding. Governments can play a key role in funding scientific open sharing practices because scientific data that can be openly shared is often

publicly funded. The *OECD Principles and Guidelines for Access to Research Data from Public Funding* clearly stipulates that "publicly funded research data refers to research data obtained from research conducted by government agencies or departments, or research conducted using public funds provided by governments at all levels" [6]. Thus, governments are the main funders of scientific data open sharing, and government funding is one of the key success factors.

Third, create and publicly release various useful scientific data, fulfilling the roles of scientific data producer, publisher, and disseminator. Government departments are creators of scientific data in key fields. *Open Data in a Big Data World* clearly states that governments possess data of great value to the scientific enterprise [16]. The *G8 Open Data Charter* defines 14 priority open fields covering companies, crime and justice, earth observation, education, energy and environment, finance and contracts, geospatial, global development, government accountability and democracy, health, science and research, statistics, social mobility and welfare, and transportation and infrastructure [12]. Thus, governments are the main force in producing scientific data. Additionally, the *G8 Open Data Charter* requires governments to publish and freely provide high-quality open data in a timely, comprehensive, and accurate manner that facilitates access and reuse for all, and to share technical expertise and experience with other countries worldwide so that everyone can benefit from open data [12]. Therefore, governments also serve as publishers and disseminators of scientific data. In fact, governments at all levels have collected and stored millions of individual data records, from tax returns and unemployment claims to hospital reimbursements and energy usage, most of which is now electronically available and easily accessible through government websites or media. For example, the Canadian Open Government website (http://open.canada.ca/en) provides 46,472 datasets with 79,257 open data records [19], while the UK Open Government website (https://data.gov.uk/) provides 42,984 open data records, including 1,167 business and economic data, 12,941 environmental data, 7,696 mapping data, 721 crime and justice data, 4,977 government data, 3,181 social data, 189 defense data, 7,373 urban data, 1,378 education data, 2,175 health data, and 1,186 transportation data [20]. Publishing these data helps governments fulfill their roles as scientific data publishers and disseminators.

Fourth, improve the scientific validity and effectiveness of government decision-making by making full use of various open scientific data, enabling governments to benefit from open scientific data. The *G8 Open Data Charter* believes that publishing data in 14 priority fields not only helps improve national democratic systems but also facilitates the reuse of these data to generate significant social and economic benefits, such as improved governance and enhanced innovation [12]. Therefore, government agencies are also users of open scientific data. Government staff can use open data to improve decision-making, including using computer programs and visualizations to analyze large amounts of open data, make more logical and evidence-based decisions faster, and improve the scientific validity and effectiveness of government decision-making.

### 2.2 Researchers' Responsibilities and Roles

Researchers are the core force in scientific data open sharing. This is because researchers have sufficient knowledge of their data to determine what information must be made public so that others can verify their results and build upon their work, and because researchers can collaborate with research institutions, funders, and journals to provide data in ways that others can effectively understand and use [2]86. Researchers are producers, disseminators, managers, and users of scientific data, with the following main responsibilities and roles in scientific data open sharing activities:

First, comply with scientific data management laws and policies and respect intellectual property rights. The *OECD Principles and Guidelines for Access to Research Data from Public Funding* explicitly requires that researchers should negotiate and sign research data sharing agreements (i.e., data access agreements) with research funding agencies and research institutions as early as possible in the research project. These data access agreements should respect the legitimate rights and interests of all stakeholders, and access to and use of research data can only be restricted in special circumstances such as national security, privacy and confidentiality, business secrets and intellectual property rights, protection of rare, threatened, or endangered species, and data under consideration in legal proceedings. Meanwhile, data access agreements should consider the applicability of copyright or other intellectual property laws that may be related to publicly funded research databases [6]. The *Statement of Principles and Practices for Arctic Data Management* requires researchers to acknowledge the concerns, rights, and management practices of local knowledge holders and managers during scientific data open sharing, and to be familiar with and comply with local community or organizational data sharing practices and requirements [10]. The *WDS Data Sharing Principles* requires fully open sharing of data, metadata, products, and information, subject to compliance with national or international judicial and policy requirements, including respecting existing appropriate restrictions and conforming to international standards for ethical research conduct [15]. Therefore, researchers must comply with relevant scientific data management laws and policies and respect intellectual property rights in scientific data open sharing activities.

Second, create and provide scientific data, fulfilling the roles of scientific data producer and disseminator. Researchers are the main producers of scientific data. *Open Data in a Big Data World* clearly stipulates that scientists funded by the government have a responsibility to contribute to the public interest by creating and communicating new knowledge and data, and to make these data available to others as soon as possible after their generation so that they can be reused [16]. Researchers are also the main disseminators of scientific data. Today, many researchers informally exchange and collaborate on large amounts of data on the internet due to their personal and professional relationships and to support their respective research activities [21]. The *OECD Principles and Guidelines for Access to Research Data from Public Funding* explicitly requires

that the international research community should be able to access information at the lowest possible cost, preferably not exceeding the marginal cost of dissemination, on equal terms, and that this open access to publicly funded research data should be easy, timely, user-friendly, and preferably internet-based [6]. In 2016, a survey of over 500 Wellcome Trust grant recipients showed that about 50% of respondents made their data available to others [22].

Third, preserve and manage scientific data, fulfilling the role of scientific data manager. Scientific data open sharing cannot be separated from effective scientific data management support. The *LERU Open Research Data Statement* and the *Guidelines on Open Access to Scientific Publications and Research Data in Horizon 2020* require researchers to provide a data management plan when applying for funding, explaining how research data collected or generated during and after the project will be handled and preserved [8,13]. The *Guidelines on Implementation of Open Access to Scientific Publications and Research Data* requires beneficiaries to develop a data management plan at the start of the project to identify what data the project will generate or process and how these data will be utilized, curated, stored, and preserved [17]. The *ERC Guidelines for Open Access to Research Results* recommends that all funded researchers retain all research data produced and used in the course of their work and be prepared to share these data with other researchers, as long as the data are not subject to copyright constraints, rights restrictions, confidentiality requirements, or contractual terms [23]. Regarding data preservation methods, the *Guidelines on Open Access to Scientific Publications and Research Data in Horizon 2020* requires researchers to store research data (including data needed to verify results) in data repositories and take measures to enable third parties to freely access, mine, utilize, copy, and disseminate these data and metadata [13]. Meanwhile, *Open Data in a Big Data World* stipulates that data verifying scientific claims should be preserved in well-managed and trusted repositories with low access barriers as much as possible, and allow the validity of data to be tested through replication of experiments or observations [16]. Additionally, the *LERU Roadmap for Research Data* requires researchers to work with librarians and IT support staff to identify best practices for describing research data metadata [9]. Today, many researchers have become part-time or full-time data managers [21]. Therefore, researchers should effectively manage the core links or elements related to scientific data open sharing from the perspective of the scientific data lifecycle and fully play their key role in scientific data management.

Fourth, effectively utilize scientific data as core users. Researchers are both the main producers and the main users of scientific data. The common ways researchers use open scientific data are: (1) data citation, i.e., citing existing data to support their own research. The *WDS Data Sharing Principles* encourages researchers to appropriately cite data to respect data sources [15]. *Open Data in a Big Data World* requires researchers to reference data creators, sources, and permanent digital identifiers when using data created by others in academic publications, because citation is crucial for evidence-based reasoning practice and is one of the criteria for evaluating research contributions [16]. In 2016, Digital

Science's survey of over 2,000 researchers found that 68% of researchers valued data citation as much as paper citation, 10% considered data citation more important than paper citation, and only 2% did not value data citation [24]. A 2017 survey of over 2,300 researchers further found that researchers' willingness to reuse open datasets in their own research increased by 10% compared to the previous year, reaching 80% [22]. (2) data verification, i.e., verifying the authenticity of data or research results. The *Guidelines on Implementation of Open Access to Scientific Publications and Research Data* requires beneficiaries to provide information on available tools for verifying results through the repository when depositing research data and related metadata needed to verify results, to help other researchers better verify relevant data or viewpoints [17]. (3) data combination, i.e., creating new datasets through data combination [21].

### 2.3 Research Institutions' Responsibilities and Roles

Research institutions are a collective term for colleges, universities, research institutes, and other non-profit research organizations. Like researchers, research institutions are both major producers and managers, disseminators, and users of scientific data, and can fulfill the following important responsibilities and roles in scientific data open sharing activities:

First, create a favorable environment for scientific data open sharing. *Open Data in a Big Data World* clearly requires research institutions to be responsible for creating a favorable environment for open data, including providing training and related technical support in data management, preservation, and analysis, library and data management services, and establishing incentives and criteria for career advancement for those involved in open data [16]. The *OECD Principles and Guidelines for Access to Research Data from Public Funding* requires research institutions to consider adjusting existing reward systems and establishing new ones, including recognizing data management activities in retention and promotion reviews, to create a better scientific data open sharing environment [6].

Second, formulate institutional scientific data open sharing policies and standards. Research institutions need to follow national or industry scientific data open sharing policies, and can also formulate more detailed institutional scientific data open sharing guidelines and communicate these policies to researchers [2]118. The *OECD Principles and Guidelines for Access to Research Data from Public Funding* requires research institutions to cooperate with international organizations to develop new standards related to research data open sharing, including more general information and communication technology standards and scientific data open sharing evaluation standards. When developing evaluation standards, they should consider the overall public investment in research data production and management, the management performance of data collection and archiving institutions, the reuse degree of existing datasets, knowledge generated from reusing existing data, and data types most likely to be needed in the future [6]. The *LERU Roadmap for Research Data* requires each LERU

member to consider developing an institutional roadmap for research data to comply with the strategic goals, missions, and actions set by research funders, and on this basis formulate and promulgate an institutional data policy that clarifies the roles and responsibilities of all stakeholders in the scientific data management process [9].

Third, develop information and communication technology and infrastructure to support scientific data open sharing. Today, research institutions often have large amounts of individual scientific data or specialized scientific data centers and need to develop information and communication technology and establish infrastructure including institutional repositories, libraries, or data centers to support scientific data open sharing. The *LERU Open Research Data Statement* requires research institutions to shift to data-driven research, develop information and communication technology and infrastructure to support research, and cooperate with researchers and research funders to share infrastructure and best practices [8]. The *LERU Roadmap for Research Data* also requires research institutions to ensure that researchers have access to the infrastructure and support conditions needed for data utilization, such as external disciplinary data centers or internal institutional repositories [9].

Fourth, provide long-term preservation and access to institutional open scientific data. Research institutions can and should play a leading role in their academic and knowledge resource management and have a special responsibility to proactively provide preservation and access to research data [2]119. The *OECD Principles and Guidelines for Access to Research Data from Public Funding* requires research institutions to assume formal responsibility for long-term data preservation at the start of each new project, particularly identifying the most appropriate data storage facilities to ensure effective preservation, management, and utilization of research data [6].

Fifth, manage publicly funded scientific data projects and ensure the security of institutional scientific data. Research institutions should manage the implementation of publicly funded research plans and projects by their employees according to academic norms and funding sources (public and private, internal and external) [21]. The *OECD Principles and Guidelines for Access to Research Data from Public Funding* emphasizes that research institutions and other stakeholders should particularly support the use of technologies and tools to ensure the integrity and security of research data from intentional or unintentional loss, destruction, modification, and unauthorized access, while protecting datasets and the equipment storing them from environmental hazards such as high temperature, dust, power surges, magnetism, and static electricity discharge [6].

Sixth, evaluate scientific data open sharing practices. The *OECD Principles and Guidelines for Access to Research Data from Public Funding* requires research institutions to regularly evaluate cost-effectiveness and data access agreements. Data access agreements should consider the applicability of copyright or other intellectual property laws that may be related to publicly funded research databases, promote institutional practices such as formulating rules and

regulations and explaining the responsibilities of parties involved in data-related activities, pay appropriate attention to relevant international data documentation standards, explain good practices in methods, techniques, and means used for collecting, disseminating, and accessing archived data, and further improve cost-effectiveness within the global scientific system by describing good practices in data management and specialized support services [6]. Therefore, evaluating scientific data open sharing activities and data access agreements is a basic responsibility of research institutions.

Seventh, support or be responsible for training in data management. The *LERU Roadmap for Research Data* requires research institutions to incorporate data management into training courses, introduce knowledge about data management contracts, regulations, and laws in training classes, and involve a wide range of stakeholders (such as graduate school leaders responsible for training programs, human resources leaders, research librarians, IT directors, certification bodies, and decision-makers) in training [9]. *Open Data in a Big Data World* also requires research institutions to provide training in data management, preservation, and analysis, and train researchers in big data and linked data analysis and open data management [16]. Therefore, research institutions need to support or be responsible for researchers' data management training, provide funding for such training, and facilitate the development of data professionals [2]118-119.

### 2.4 Research Funding Agencies' Responsibilities and Roles

Research funding agencies are responsible for providing and allocating public funds or other funding to support various research activities. When research funding agencies determine funding priorities, they have strong capacity to influence the overall data policies and management systems of the research institutions they support [21]. *Open Data in a Big Data World* believes that research funding agencies should include open data costs in research projects as an inherent part of research costs, and provide necessary resources and policies for the long-term sustainable development of infrastructure and repositories; research funding agencies have a responsibility to promote the open data process by funding relevant hardware and software infrastructure, stimulating research on the fundamental principles of data science, establishing incentives for research institutions, and helping them fulfill their responsibilities. Moreover, research funders can also reward data sharing by improving research and impact assessment analysis, and support the reuse of open data by providing dedicated funding [16]. However, the *LERU Open Research Data Statement* points out that as a condition of funding, research funders may have certain requirements for research data generated by their funding, such as copyright issues, but these requirements are usually listed in funding agreements, and researchers and research groups should ensure that any consortium or partnership agreements reflect the requirements of funding agreements [8].

### 2.5 Libraries or Archives and Data Centers' Responsibilities and Roles

In scientific data open sharing activities, libraries or archives and data centers have unique and important roles, serving as organizers, managers, and open sharing service providers of scientific data. Existing international scientific data open sharing policies stipulate the following responsibilities and roles for libraries or archives and data centers:

First, formulate policies and service standards related to scientific data open sharing. *Open Data in a Big Data World* clearly stipulates that libraries and archives are responsible for formulating and providing data services and technical standards to ensure others can access relevant data [16]. In fact, libraries or archives and data centers can influence the development of national data policies. For example, in Canada, the research library community has strong representation in the government research data strategy working group; similarly, in Australia, senior library community members are appointed to the Australian National Data Service Steering Committee. Additionally, library directors and other senior library staff can play a key role in helping local government officials understand the challenges of scientific data management and formulate reasonable scientific data management policies [25]. Particularly in the process of formulating institutional scientific data open sharing policies or service standards, libraries or archives and data centers can play a leading role.

Second, collect, organize, and store scientific data. Libraries or archives and data centers can serve as organizers of scientific data. *Open Data in a Big Data World* requires libraries to continue playing the role of collecting, organizing, preserving, and utilizing knowledge (including scientific data), adapting to technological changes from print to digital formats, and addressing open data management issues [16]. Over the past 30 years, the traditional concept of library collections has been expanded to cover tangible and intangible digital objects, which can be stored and accessed locally or remotely [25]. Libraries or archives and data centers can acquire more digital resources through both "ownership" and "access" approaches, including establishing local datasets and providing access to publicly available remote data products and services.

Third, support scientific data management. Libraries or archives and data centers can fulfill the role of scientific data managers and effectively support scientific data open sharing activities. In this regard, the *LERU Open Research Data Statement* proposes ten recommendations for research libraries [26]: (1) provide research data management support, including data management plans for funding applications, intellectual property advice, and information materials; assist staff in developing data management plans and incorporate data management into training courses; (2) participate in developing metadata and data standards and provide metadata services for research data; (3) create data librarian positions and develop data librarian professional skills; (4) actively participate in formulating institutional research data policies or resource plans, appropriately encourage and adopt open data strategies in the research data lifecycle; (5) net-

work and cooperate with researchers, research groups, data archives, and data centers to provide interoperable infrastructure for data access, discovery, and sharing; (6) support the research data lifecycle by providing storage, discovery, and permanent access services; (7) promote research data citation by applying permanent identifiers for research data; (8) provide institutional data catalogs or data repositories based on available infrastructure; (9) participate in specific thematic data management practices; (10) provide or facilitate cooperation with institutional IT departments on secure storage of dynamic and static research data, and seek to utilize appropriate cloud services. These recommendations also apply to other libraries, archives, and data centers.

Fourth, provide data training and curation services. Libraries or archives and data centers can provide data training and curation services for other stakeholders. The *LERU Roadmap for Research Data* requires stakeholders (including libraries or archives and data centers) to participate in and provide scientific data training, incorporate data curation into training courses, and provide scientific data curation services. These data curation services include: (1) identifying documentation and metadata needs from the beginning of the project and considering them throughout the data lifecycle; (2) identifying best practices for metadata describing scientific data; (3) selecting interoperable metadata formats; (4) establishing sound scientific data infrastructure to provide a common framework for creating, processing, and sharing scientific data [9]. Today, some US university libraries, such as Purdue University Library and Georgia Tech Library, attach great importance to developing local data curation capabilities and have used institutional repositories as the basis for providing scientific data curation services; some libraries jointly provide data curation services with supercomputing or research computing centers; a group of librarians, IT staff, and graduate students at the University of Wisconsin-Madison Library and Information Studies jointly provide research data digital curation services. Additionally, some US librarians can provide data consulting services to researchers on key issues in data management, such as copyright and intellectual property rights, metadata and technical standards, data archiving and preservation [25].

### 2.6 Publishers' Responsibilities and Roles

With the development of the open access movement, publishers' position in scientific data sharing has somewhat weakened but remains important. Publishers' responsibilities and roles in scientific data open sharing activities mainly include the following two aspects:

First, publish and share scientific data and metadata. *Open Data in a Big Data World* clearly stipulates that publishers have a responsibility to provide data to reviewers during the review process, require data to be openly accessible simultaneously with publications, and master the situation of referencing and citing these data. Publishers also have a responsibility to enable scientific records to be used for subsequent analysis through open metadata provision and open access to text and data mining [16].

Second, provide data for verifying scientific claims. *Open Data in a Big Data World* requires that data providing evidence for published scientific claims must be published simultaneously to enable rigorous examination of the logic between data and claims and to allow the validity of data to be tested through replication of experiments or observations. To this end, publishers must make open data a necessary condition for publication, store data verifying scientific claims in trusted repositories, and make data verifying scientific claims freely and openly available according to principles of transparency and reproducibility [16].

Therefore, publishers can take the following practical actions to promote scientific data open sharing: (1) formulate and improve open data publishing policies; (2) advocate good scientific data open practices among different research communities; (3) improve journal publishing policies to encourage good scientific data open practices; (4) provide scientific data open sharing solutions to help researchers share their own data and discover and use scientific data; (5) cooperate with research communities to establish scientific data open sharing solutions, such as establishing the Research Data Alliance to improve research data policies, standards, data linking, and citation [22].

### 2.7 Professional Associations or Societies' Responsibilities and Roles

Professional associations or societies are major participants in formulating scientific norms, values, and standards and scientific responsibilities. They can provide a liaison center for interaction and communication among specific disciplinary groups, especially national academic groups, and continuously understand the data generated by their research community members [21]. They are an important force participating in scientific data open sharing, with main responsibilities and roles including the following two aspects:

First, formulate open scientific data policies or guidelines and coordinate scientific data open sharing activities. *Open Data in a Big Data World* clearly stipulates that professional associations or societies should formulate open data guidelines and policies, provide various opportunities to reflect or enhance members' cognitive norms and practices, and influence "bottom-up" open data initiatives by expressing their specific field's research principles and priorities [16]. Professional associations or societies can also coordinate scientific data open sharing behaviors among members and institutions in their professional fields by formulating formal and informal open data codes of conduct [21]. In some fields, professional associations or societies can take the lead in establishing scientific data oversight committees to better perform their coordination functions, such as making recommendations for industry data repository construction, evaluating the future value and practical utility of scientific data, and organizing interactions with researchers or institutions [2]111.

Second, promote scientific data open sharing activities. Professional associations or societies can vigorously promote scientific data open sharing activities by holding academic conferences, creating academic journals, interacting with

policymakers and researchers, and promoting open data concepts within professional communities [22]. In this regard, the *OECD Principles and Guidelines for Access to Research Data from Public Funding* requires professional associations to fully communicate with research communities about the development of research data quality standards when they do not exist, and actively promote data citation and record citation practices in indexes [6]. Therefore, professional associations or societies can fulfill the roles of promoters and coordinators of scientific data open sharing activities.

### 2.8 Users' and Enterprises' Responsibilities and Roles

In a broad sense, all individuals or institutions that utilize scientific data are scientific data users, including researchers and various institutions. Since the responsibilities and roles of researchers and related institutions have been discussed above, this section only explains the responsibilities and roles of individuals or institutions (including enterprises) that use scientific data but do not produce or provide scientific data:

First, promote the free sharing of scientific data. Users (including enterprises) are direct beneficiaries of scientific data open sharing, and enabling users to freely share scientific data is the main purpose of the open scientific data movement. The *Implementation Guidelines for the GEOSS Data Sharing Principles* strongly encourages users to receive data and metadata at zero cost, enabling users to discover data and information sources without restriction [7]. The *Data Sharing Principles for Developing Countries* also clearly stipulates that data should be free to end users [14]. The *Guidelines on Implementation of Open Access to Scientific Publications and Research Data* [17] and the *Guidelines on Open Access to Scientific Publications and Research Data in Horizon 2020* [13] both require that users can freely obtain, mine, exploit, copy, and disseminate research data.

Second, achieve efficient utilization of scientific data. Maximizing the potential or future value of scientific data mainly depends on users' efficient utilization of scientific data. The *Implementation Guidelines for the GEOSS Data Sharing Principles* requires promoting the reuse and redistribution of shared data, metadata, and products by GEOSS users to maximize GEOSS social benefits [7]. The *Statement of Principles and Practices for Arctic Data Management* requires users to adhere to the ethical norms and data sharing norms of the Polar Information Commons when utilizing scientific data [10]. The *Data Sharing Principles for Developing Countries* requires that any subsequent user of data must mark the data source when reusing data (i.e., citing data) and must not misuse data in any way [14]. The *G8 Open Data Charter* allows users to provide feedback on data quality and quantity and its revision to obtain open data of the highest quality standards [12]. Additionally, the *OECD Principles and Guidelines for Access to Research Data from Public Funding* requires users to participate in regular evaluations of data access agreement implementation [6]. All these provisions help users efficiently utilize open access scientific data.

## 3. Conclusion and Implications

International scientific data open sharing policies define the responsibilities and roles of different stakeholders (including governments, researchers, research institutions, research funding agencies, libraries or archives, data centers, publishers, professional associations or societies, users, enterprises, etc.) in scientific data open sharing. China's newly promulgated *Measures for the Management of Scientific Data* aims to "further strengthen and standardize scientific data management, ensure scientific data security, improve open sharing levels, and better support national scientific and technological innovation, economic and social development, and national security." Regarding scientific data open sharing, it requires [27]: (1) the State Council's science and technology administrative department to coordinate and promote standardized management, open sharing, and evaluation of scientific data; (2) relevant State Council departments and provincial government departments to coordinate and plan the construction of scientific data centers in their departments (regions) and promote scientific data open sharing; (3) legal entities such as research institutes, higher education institutions, and enterprises to establish scientific data management systems, publish scientific data open catalogs, and open them to society; (4) scientific data centers to undertake integration, submission, classification, processing, analysis, and mining of scientific data in relevant fields, ensure scientific data security, and promote scientific data open sharing according to laws and regulations; (5) competent departments and legal entities to strengthen scientific data lifecycle security management, formulate scientific data security protection measures, and prevent malicious use of data; (6) competent departments and legal entities to establish and improve evaluation systems for scientific data management and open sharing work. Although the *Measures for the Management of Scientific Data* makes preliminary provisions for the responsibilities of governments, research institutes, higher education institutions, enterprises, and scientific data centers in scientific data open sharing, it does not define in detail the responsibilities and roles of different stakeholders including researchers, research funding agencies, libraries or archives, publishers, professional associations or societies, and users in scientific data open sharing.

To vigorously promote the implementation of the national big data strategy and advance scientific data open sharing, we should: First, based on the *Measures for the Management of Scientific Data*, China's government competent departments should draw on foreign successful experiences to formulate government scientific data open sharing regulations, establishing the rights and responsibilities of relevant government competent departments and institutions in government scientific data open sharing. Second, China's various research funding agencies should revise their funding management measures to mandatorily require various funded projects to implement open sharing of research data during and after completion, and make this requirement one of the necessary conditions for obtaining funding. Third, based on the *Measures for the Management of Scientific Data*, China's research institutions (including research

institutes and higher education institutions), libraries or archives, data centers, publishers, professional associations or societies, and enterprises should draw on foreign successful experiences to formulate their own institutional scientific data open sharing guidelines, clarifying the rights and responsibilities of institutions and individuals in scientific data open sharing, providing detailed implementation rules for scientific data open sharing for institutions and individuals, such as the *Measures for the Management and Open Sharing of Scientific Data of the Chinese Academy of Sciences (Trial)* [28], thereby effectively improving the level of scientific data open sharing.

## References

[1] THE ROYAL SOCIETY. Science as an open enterprise[M]. London: The Royal Society, 2012: 18.
[2] 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.
[3] LYON L. Dealing with data: roles, rights, responsibilities and relationships[R/OL]. [2019-01-27]. http://opus.bath.ac.uk/412/1/dealing_{{{with}}{{data}}}{{{report}}{{final}}}.
*[4] Xiang Liling, Wang Qing. On the legal attributes, responsibilities and obligations of open data and related mechanisms[J]. Journal of the National Library of China, 2013(5): 38-44.*
*[5] You Xiaguang, 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.*
*[6] OECD. OECD principles and guidelines for access to research data from public funding[EB/OL]. [2019-01-27]. https://www.oecd.org/sti/sci-tech/38500813.pdf.*
*[7] Group on Earth Observations. Implementation guidelines for the GEOSS data sharing principles[EB/OL]. [2019-01-27]. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.684.*
*[8] LERU. Open research data[EB/OL]. [2019-01-27]. http://www.leru.org/files/publications/Open*{{{Access}}
[9] LERU. LERU roadmap for research data[EB/OL]. [2019-01-27]. https://www.fosteropenscience.eu/sites/default/files/pdf/598.pdf.
[10] IASC. Statement of principles and practices for arctic data management[EB/OL]. [2019-01-27]. https://iasc.info/images/data/IASC_{{{data}}{{statement}}}.pdf.
*[11] G8 Science Ministers Statement London UK, 12 June 2013[EB/OL]. [2019-01-27]. https://www.gov.uk/government/uploads/system/uploads/attachment*{data}/file/206801/G8_{{{Scien
*[12] G8. G8 open data charter[EB/OL]. [2019-01-27]. http://opendatacharter.net/wp-content/uploads/2015/10/opendatacharter*{charter}F.pdf.
[13] EUROPEAN COMMISSION. Guidelines on open access to scientific publications and research data in Horizon 2020 version 1.0[EB/OL]. [2019-01-27]. http://www.gsrt.gr/EOX/files/h2020-hi-oa-pilot-guide_{en}.pdf.
[14] CODATA. Data sharing principles in developing countries (The Nairobi data sharing principles)[EB/OL]. [2018-10-09]. https://www.rd-alliance.org/sites/default/files/attachment/NairobiDataSharingPrinciples.pdf.

[15] ICSU-WDS. WDS data sharing principles[EB/OL]. [2018-10-09]. https://www.icsu-wds.org/services/data-sharing-principles.
[16] Science International. Open data in a big data world[EB/OL]. [2019-01-27]. http://www.science-international.org/sites/default/files/reports/open-data-in-big-data-world_{{{long}}}*{{en}}}.pdf.*
*[17] EUROPEAN RESEARCH COUNCIL (ERC). Guidelines on implementation of open access to scientific publications and research data[EB/OL]. [2019-01-27]. http://ec.europa.eu/research/participants/data/ref/h2020/other/hi/oa-pilot/h2020-hi-erc-oa-guide*{en}.pdf.
[18] CABINET OFFICE. G8 open data charter: UK action plan 2013[EB/OL]. [2019-01-27]. https://www.gov.uk/government/publications/g8-open-data-charter-uk-action-plan-2013.
[19] Open government portal[EB/OL]. [2018-06-27]. https://open.canada.ca/data/en/dataset?portal_{type}=
[20] Find open data[EB/OL]. [2018-06-27]. https://data.gov.uk/.
[21] UHLIR P F, SCHRODER P. Open data for global science[EB/OL]. [2019-01-27]. https://datascience.codata.org/articles/10.2481/dsj.6.OD36/gallery/367/download/.
[22] FIGSHARE. The state of open data 2017: a selection of analyses and articles about open data[R/OL]. [2019-01-27]. https://ndownloader.figshare.com/files/9558481.
[23] EUROPEAN RESEARCH COUNCIL. Open access guidelines for research results funded by the ERC[EB/OL]. [2019-01-27]. https://erc.europa.eu/sites/default/files/document/file/ERC revised_{{{feb}}}*{{2016}}}.pdf.*
*[24] FIGSHARE. The state of open data: a selection of analyses and articles about open data[R/OL]. [2019-01-27]. https://ndownloader.figshare.com/files/6558051.*
*[25] CORRALL S. Roles and responsibilities: libraries, librarians and data[M]//PRYOR G. Managing research data. London: Facet Publishing, 2012: 105-133.*
*[26] THE LIBER WORKING GROUP ON E-SCIENCE. Ten recommendations for libraries to get started with research data management[EB/OL]. [2019-01-27]. http://libereurope.eu/wp-content/uploads/The%20research%20data%20group%202012%20v7%20final.pdf.*
*[27] General Office of the State Council. Notice of the General Office of the State Council on issuing the Measures for the Management of Scientific Data[EB/OL]. [2019-01-27]. http://www.most.gov.cn/mostinfo/xinxifenlei/fgzc/gfxwj/gfxwj2018/201804/t20180404*{139023}
[28] Chinese Academy of Sciences. Measures for the management and open sharing of scientific data of the Chinese Academy of Sciences (Trial)[EB/OL]. [2019-02-21]. http://m.cas.cn/tzgg1/201902/P020190220358041915907.pdf.

## Author Contributions

Sheng Xiaoping: Responsible for paper writing and revision;
Wang Yi: Responsible for data collection and policy content analysis.

---

**The Responsibilities and Roles of Stakeholders in the Open Sharing of Scientific Data: An Analysis Based on Policies of Open Sharing of Scientific Data in International Organizations**

Sheng Xiaoping, Wang Yi
School of Library, Information and Archives, Shanghai University, Shanghai 200444

**Abstract:** [Purpose/significance] By analyzing the responsibilities and roles of different stakeholders defined by the open sharing policies of scientific data in international organizations, this paper provides a reference for perfecting the *Measures for the Management of Scientific Data* and implementing the open sharing of scientific data in China. [Method/process] Based on network survey and text analysis, this paper discusses the responsibilities and roles of different stakeholders in the open sharing of scientific data defined by the open sharing policies of scientific data in international organizations. [Result/conclusion] The open sharing policies of scientific data in international organizations define a wide range of stakeholders including governments, researchers, research institutions, research funding institutions, libraries or archives, data centers, publishers, professional associations or institutes, users and enterprises, who have different responsibilities and roles in the open sharing activities of scientific data. We should draw lessons from these experiences to formulate domestic open sharing policies of scientific data, and further improve the level of open sharing of scientific data in China.

**Keywords:** scientific data, open sharing, stakeholder, responsibility and role, international organization

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