

University of Nebraska - Lincoln

DigitalCommons@University of Nebraska - Lincoln

---

Copyright, Fair Use, Scholarly Communication,  
etc.

Libraries at University of Nebraska-Lincoln

---

1-29-2021

## A Review of Open Research Data Policies and Practices in China

LILI ZHANG

ROBERT R. DOWNS

JIANHUI LI

LIANGMING WEN

CHENGZAN LI

Follow this and additional works at: <https://digitalcommons.unl.edu/scholcom>



Part of the [Intellectual Property Law Commons](#), [Scholarly Communication Commons](#), and the [Scholarly Publishing Commons](#)

---

This Article is brought to you for free and open access by the Libraries at University of Nebraska-Lincoln at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in Copyright, Fair Use, Scholarly Communication, etc. by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.



# A Review of Open Research Data Policies and Practices in China

REVIEW

LILI ZHANG 

ROBERT R. DOWNS 

JIANHUI LI

LIANGMING WEN 

CHENGZAN LI

*\*Author affiliations can be found in the back matter of this article*

 ubiquity press

## ABSTRACT

This paper initially conducts a literature review and content analysis of the open research data policies in China. Next, a series of exemplars describe data practices to promote and enable the use of open research data, including open data practices in research programs, data repositories, data journals, and citizen science. Moreover, the top four driving forces are identified and analyzed along with their responsible guiding work. In addition, the “landscape of open research data ecology in China” is derived from the literature review and from observations of actual cases, where the interaction and mutual development of data policies, data programs, and data practices are recognized. Finally, future trends of research data practices within China and internationally are discussed. We hope the analysis provides perspective on current open data practices in China along with insight into the need for additional research on scientific data sharing and management.

## CORRESPONDING AUTHORS:

**Lili Zhang**

Computer Network  
Information Center, Chinese  
Academy of Sciences, Beijing,  
China

[zhll@cnic.cn](mailto:zhll@cnic.cn)

**Jianhui Li**

Computer Network  
Information Center, Chinese  
Academy of Sciences, Beijing,  
China

[lijh@cnic.cn](mailto:lijh@cnic.cn)

---

## KEYWORDS:

Open data; data policy; data practices; China; research data sharing

## TO CITE THIS ARTICLE:

Zhang, L, Downs, RR, Li, J, Wen, L and Li, C. 2021. A Review of Open Research Data Policies and Practices in China. *Data Science Journal*, 20: 3, pp. 1–17. DOI: <https://doi.org/10.5334/dsj-2021-003>

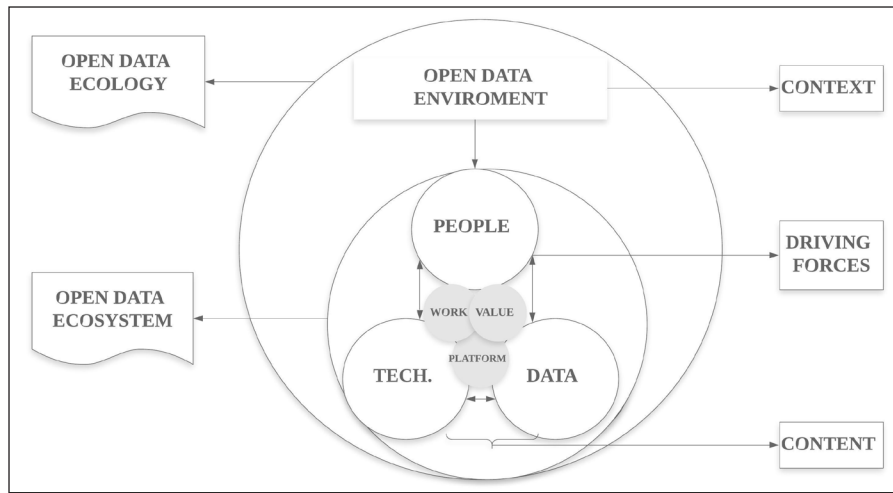
Data has become the new currency of both the global economy and the scholarly community (OECD, 2017). Thus, scientific, research, and scholarly communities around the globe are endeavoring for sound research data management and sharing practices. Many funders, such as the Arts and Humanities Research Council (AHRC), the Biotechnology and Biological Sciences Research Council (BBSRC), the Engineering and Physical Sciences Research Council (EPSRC), the European Commission (EC), the National Science Foundation (NSF), the United States Geological Survey (USGS), and the Australian Research Council (ARC), among others, are requiring data management plans for proposals to conduct sponsored research (Nature Editorial, 2018). Large Data centers, like the Finnish Social Science Data Archives, the Australian National Data Service (ANDS), and the United Kingdom (UK) Data service have accepted responsibilities for national long-term data preservation. Universities and libraries offer training opportunities to improve the skills, knowledge, and capabilities of those who are accepting responsibilities for data stewardship (Adamick, Reznik-Zellen, & Sheridan 2012; Fong and Wang, 2015; Thielen, et al 2017). The Consultative Committee for Space Data Systems (CCSDS, 2012), the Digital Curation Centre (DCC) (Whyte and Wilson, 2010), DataOne (2018), OpenAIRE (2015), and many other international organizations have established guidelines for managing research data and the systems for hosting the data. These and many other government agencies, research institutions, and other stakeholder entities of the research enterprise have developed holistic data policies to address the recognized need to improve capabilities for responsible open data access and stewardship (Gerunov, 2017; Mauthner and Parry, 2013; Wyatt, 2017).

Recognizing that, in addition to published research articles, data represent first-class research deliverables, progress is materializing. Publishers of scientific journals, within various disciplines, are establishing data policies to facilitate the reproducibility of data that have been used for studies reported in their scholarly publications (Charbonneau and Beaudoin, 2015; Naughton and Kernohan, 2016). Furthermore, several instruments have been developed for evaluating the maturity of research data and for assessing the trustworthiness of the repositories where research data are housed (Peng, 2018). The Findable, Accessible, Interoperable, and Reusable (FAIR) principles for data, initially proposed in 2014, have been widely adopted (Wilkinson, et al, 2016; Vita, et al, 2018; Corpas, et al, 2018; Boeckhout, Zielhuis and Bredenoord, 2018; Holub, et al, 2018). And, the importance of Transparency, Responsibility, User focus, Sustainability, and Technology (TRUST) for the stewardship of research data also has been recognized (Lin, et al, 2020). These efforts to improve capabilities for open data access and stewardship contribute to responsible open research data policies and practices, internationally. Taken together, this progress reflects an international movement for ensuring continuing accessibility and usability of open data products, services, and research-related information that is produced by research and scientific endeavors.

Like many of the countries involved in the international efforts to improve capabilities for open research data sharing and stewardship, China, one of the most productive scientific countries, plays a key role in improving data policies and practices (UNESCO, 2015; Shen, et al, 2015; Tollefson, 2018). According to the National Science and Technology Infrastructure Center (NSTI) of the China Ministry of Science and Technology (MOST) (2019), original research data generation in China had reached 83.72PB by the end of 2017. Information about the evolution of data policies and corresponding practices in China can improve understanding about current progress and additional opportunities for research data management and sharing, internationally.

Here, the ecology perspective is leveraged to provide a systematic review of the general progress of open research data in China. As is frequently used in information ecology (Nardi and O'Day, 2000; Wang et al, 2015), data ecology borrows the concept from biological science and studies the environment, and the relationships among organisms within and across ecosystems constituted by data, people, technologies, and their interactions, as well as other intersectional aspects, like platforms, work and value (Nardi and O'Day, 2000; Parsons, et al, 2011; Pritchard and Martel, 2019). These potential components form a complexity of dynamic equilibrium and could be organized into three general components, including context, content, and driving forces (see [Figure 1](#)). Such an open data ecology also emphasizes open service trends within different ecosystems.

Furthermore, we envision that improvements in data policies and data practices can serve as a threshold from context and content aspects into the current state of data ecology, as well as its prospects for the future. Taken together, the general data ecology analysis encompasses data policy (content), data practice (context) and people (driving forces), which includes policy-



**Figure 1** Potential components within an open data ecology and its ecosystem.

makers and research sponsors, as well as the data producers, end-users, data stewards, and others whose work is supported by the sponsors. Among the different elements in the ecology of data, the needs of the people are the driving forces, especially when an ideal and harmonious data ecology is still emerging.

The design and methodology of this paper includes the following elements:

- The Literature review is conducted within the China National Knowledge Infrastructure (CNKI) ([www.cnki.net](http://www.cnki.net)), which is one of the largest online platforms for Chinese journals, dissertations, standards, and patents, along with reference databases for Chinese policies. Further governmental portals are also explored for particular content-level policy studies.
- Field observation has been carried out by performing analysis and participating in data program planning, development, and operations activities and related efforts, as part of the authors' work in the Computer Network Information Center (CNIC), Chinese Academy of Sciences (CAS). CNIC is the CAS-level leading institution for general data infrastructure and various data services and has been managing data programs sponsored by major funding agencies in China for years. Thus, members of this community have gained in-depth understanding and insight into the development of data policies and data practices in China. Typical open research data practice examples are selected from research programs, data repositories, data journals, and citizen science. Furthermore, the top four exemplars, representing salient initiatives in China, from both policy and funding perspectives, are identified and their endeavors to foster open data have been introduced.
- The conceptual model for open research data ecology has been framed and future work for a better world of open data also has been discussed.

## 2. EVOLVING DATA POLICY ENVIRONMENT

### 2.1 NATIONAL LEVEL LEGISLATIONS

Data policy should not only simplify the path towards “effective research data stewardship and infrastructure development” (Parsons, 2013), but also help maximize the data benefits through data sharing (Federer, et al, 2018). **Table 1** provides an overview of the current state-level rules governing open research data in China.

According to **Table 1**, generally, the “Law of the People’s Republic of China on Science and Technology Progress” (2008) established the fundamental rules for research data stewardship, stating that, “The Science and Technology Administrative Department of the State Council shall, in conjunction with the relevant competent departments of the State Council, establish information systems for scientific and technological resources, such as S&T research bases, scientific instruments and scholarly literature, S&T data and natural resources...and should release the distribution and usages of all the research sources as well”.<sup>1</sup> That law also addresses the disclosure of governmental information since it is one of the largest contributors of open research data in China.

<sup>1</sup> Law of the People’s Republic of China on Science and Technology Progress, Standing Committee of the National People’s Congress, P.R.C., Art.65 (2008). Available at: [http://www.gov.cn/ziliao/flfg/2007-12/29/content\\_847331.htm](http://www.gov.cn/ziliao/flfg/2007-12/29/content_847331.htm) [Accessed: 8 Sep 2020].

LEGISLATIONS	ISSUED BY
<b>Law of the People's Republic of China on Science and Technology Progress (2008 amended)</b>	Standing Committee of the National People's Congress (SCNPC, P.R.C.)
<b>Copyright Law of the People's Republic of China (2010 amended)</b>	
<b>Law of the People's Republic of China on Promoting the Transformation of Scientific and Technological Achievements (2015 amended)</b>	General Office of the State Council, P.R.C.
<b>Cybersecurity Law of the People's Republic of China (2017)</b>	
<b>Measures for Managing Scientific Data (2018)</b>	

**Table 1** Key legislations guiding open research data in China.

Other rules also affect research data stewardship by providing provisions on cybersecurity and intellectual property, as well as those governing particular elements, such as research outcomes. In particular, the “Data Security Law of the People's Republic of China (Draft Version)”<sup>2</sup> has been released during June 2020 for public comment. Like the practice of General Data Protection Regulation (GDPR) in Europe, this law aims to ensure the flow of data for the protection of data rights within a safe environment, nationwide. Key measures include clarity and implementation of data security protection obligations for different stakeholders, promotion of various data flow in sound manners, and guidance for institutional measures to ensure the safety of governmental data sharing.

By contrast, “Measures for Managing Scientific Data”,<sup>3</sup> brings research data management and sharing to a new phase in which general rules covering comprehensive data aspects have been set up for implementation, including enhanced entire lifecycle data management, strengthened capability for data assets protection, sustained open data ecology supported by continuing funding for specific data programs, clear data property rights, and responsibility for long-term data stewardship. In effect, many institutions are revising their implementation guidelines and plans accordingly. For example, CAS plans to initiate several steps to complete this implementation, including the development of CAS level data policies following national laws, construction of strengthened CAS data centers and mature data infrastructure, exploration of innovative open data models, and improvements in fair incentive mechanisms for multiple stakeholders, etc. (AOCAS, 2018). By the end of 2019, fifteen provincial-level administrative regions, including Anhui, Chongqing, Gansu, Guangxi Autonomous Region, Hainan, Heilongjiang, Hubei, Inner Mongolia Autonomous Region, Jiangsu, Jilin, Shaanxi, Shandong, Sichuan, Tianjin, Xinjiang, and Yunnan, have published regional rules to govern scientific data in line with the national-level law. And more steps, in accordance with the “Measures for Managing Scientific Data” (2018), also are being planned for completion.

## 2.2 INSTITUTIONAL DATA POLICIES

In addition to these national laws, administrative institutions have produced more guidelines for sound data stewardship and these call for open data and sharing of other research resources generated from public funded research programs.

Government information, serving as one of the major sources for data generation and sharing, is governed by the “Regulations of the People's Republic of China on Disclosure of Government Information”.<sup>4</sup> Also, the document, “Interim Measures for the Management and Sharing of Government Information resources”,<sup>5</sup> provides guidance for electronic records, which helps to facilitate the governance of scientific data from an administrative perspective.

<sup>2</sup> Data Security Law of the People's Republic of China (Draft Version), solicit comments through 3 Jul 2020 to 16 Aug 2020. Available at: <http://www.npc.gov.cn/flcaw/> [Accessed: 9 Sep 2020].

<sup>3</sup> Measures for Managing Scientific Data, General Office of the State Council, P.R.C., Art.21 (2018). Available at: [http://www.gov.cn/zhengce/content/2018-04/02/content\\_5279272.htm](http://www.gov.cn/zhengce/content/2018-04/02/content_5279272.htm) [Accessed: 9 Sep 2020].

<sup>4</sup> Regulations of the People's Republic of China on Disclosure of Government Information of 2007, State Council, P.R.C. (2019). Available at: <http://www.mofcom.gov.cn/article/swfg/swfgbg/201905/20190502864821.shtml> [Accessed: 9 Sep 2020].

<sup>5</sup> Interim Measures for the Management and Sharing of Government Information resources of 2016, State Council, P.R.C. (2016). Available at: [http://www.gov.cn/zhengce/content/2016-09/19/content\\_5109486.htm](http://www.gov.cn/zhengce/content/2016-09/19/content_5109486.htm) [Accessed: 9 Sep 2020].

The Ministry of Science and Technology in China (MOST) provides another research data governance example. Early in 2001, the “Interim Provisions on the Administration of National S&T Plans”<sup>6</sup> states the responsibilities of the MOST to establish databases, archiving systems, and rules for the preservation, usage, and sharing of data and related information. Following up with rules for program management, the MOST also points out that “All national science and technology plans should establish compatible databases to share information and data resources”.<sup>7</sup> In particular, as a key component of research outcomes, scientific data is highlighted for deposit into archives and research facilities that provide guidance to ensure their integrity, completeness, and accuracy (MOST, 2003).<sup>8</sup> In effect, the MOST becomes the motivator for initializing and implementing first state-level guidance on “Measures for Managing Scientific Data” (2018). In line with these efforts, the Chinese Academy of Sciences (2019),<sup>9</sup> The Chinese Academy of Agricultural Sciences (2019)<sup>10</sup> and others also published their institutional open data rules based on a decade of data work experience, in which, mechanisms, responsibilities, workflow, and plan for data center development are underlined accordingly.

### 2.3 DISCIPLINARY DATA POLICIES

Data has not been emphasized enough in every field to be regulated by specific data rules, but there are some data-intensive research fields taking the lead in developing data policies. For example, searching the China Legal Knowledge Database (CLKD)<sup>11</sup> retrieved 254 excerpts of disciplinary policies that were entitled with “data”. Shown in [Table 2](#) are several selected disciplinary research data policy examples covering geoscience, medical and health sciences, meteorology, ocean science, seismology and space science.

SUBJECTS	POLICYMAKER(S) (EFFECTIVE SINCE): POLICY TITLE
Geoscience	State Oceanic Administration, PRC (2018): Measures for Managing Polar Expedition Data of China
Geoscience	Ministry of Natural Resources, PRC (Former Ministry of Land and Resources, PRC) (2010): Interim Measures for Managing Land and Resources Data
Medical and health sciences	National Health Commission, PRC (2018): Interim Provisions on National Health Care Big Data Standards, Safety and Service Management
Meteorology	China Meteorological Administration (2018): Fengyun Meteorological Satellite Data Management Measures (for Trial Implementation)
Meteorology	China Meteorological Administration (2017): Measures for Meteorological Data Exchange and Management
Ocean science	China Oceanic Information Network (2015): Data Sharing and Service Procedures in Marine Ecological Environment Monitoring (for Trial Implementation)
Seismology	China Earthquake Administration (2006): Measures for Managing Seismological Science Data
Space science	State Administration of Science, Technology and Industry for National Defense, PRC; National Development and Reform Commission, PRC; Ministry of finance, PRC (2018): Interim Measures for the Management of National Civil Satellite Remote Sensing Data
Space Science	State Administration of Science, Technology and Industry for National Defense, PRC; China National Space Administration (2016): Measures for Managing Scientific Data in Lunar and Deep Space Exploration

**Table 2** Examples of disciplinary rules governing research data sharing in China.

6. Interim Provisions on the Administration of National S&T Plans of 2000, Ministry of Science and Technology, P.R.C., Chap. 5, Art.21 (2001). Available at: [http://www.most.gov.cn/fggw/bmgz/200811/t20081105\\_64825.htm](http://www.most.gov.cn/fggw/bmgz/200811/t20081105_64825.htm) [Accessed: 10 Sep 2020].

7. Interim Measures for the Management of National S&T Plan Projects of 2000, Ministry of Science and Technology, P.R.C., Chap. 3, Art.35 (2001). Available at: [http://www.most.gov.cn/fggw/bmgz/200811/t20081105\\_64826.htm](http://www.most.gov.cn/fggw/bmgz/200811/t20081105_64826.htm) [Accessed: 10 Sep 2020].

8. Interim Provisions on Strengthening the Management of the National S&T Plan Achievements of 2003, Ministry of Science and Technology, P.R.C., Art.6 (2003). Available at: [http://www.gov.cn/gongbao/content/2003/content\\_62434.htm](http://www.gov.cn/gongbao/content/2003/content_62434.htm) [Accessed: 10 Sep 2020].

9. Measures for the Management and Sharing of Scientific Data of the Chinese Academy of Sciences (for Trial Implementation), Chinese Academy of Sciences (2019).

10. Measures for the Management of Open Access and Scientific Data of the Chinese Academy of Agricultural Sciences, Chinese Academy of Agricultural Sciences (2019).

11. China Legal Knowledge Database, available at: <http://law.cnki.net/> [Accessed: 31 Jul 2020].



### 3. PROACTIVE OPEN DATA PRACTICES

Data practices often reflect policies and also demonstrate how such policies have been implemented. Based on their importance and popularity, exemplars for four different types of open data practices are described, below, and include open data in research programs, repositories, data journals, and citizen science.

#### 3.1 IN RESEARCH PROGRAMS

According to the 13th Five-year science plan,<sup>12</sup> there are five types of major projects, and among those, three major research projects have policies governing data and information. The other two are talents and enterprise innovation-related grants, in which we suppose more flexibility is left due to the complexity of such work as well as priorities for Intellectual Property protection.

Shown in [Table 3](#), data responsibilities, data infrastructures (i.e. data platforms and databases), data curation and sharing, are stressed within these three program provisions. Moreover, rules for sharing of national major scientific research infrastructures and large-scale scientific research instruments also strengthen the capability of data production and broaden the scope of data sharing by reducing investment costs and increasing the efficiency of facility usage ([Table 3](#)).

POLICYMAKER(S) (ENACTED SINCE): POLICY TITLE	CONTENTS
<b>NSFC (2015):</b> <b>Measures for the Management of Research Outputs in Projects Funded by the National Natural Science Foundation of China</b> <sup>13</sup>	<ul style="list-style-type: none"> <li>PI should collect and reserve the original records with sound manners and submit them to the supporting unit as required, whilst also ensuring the accuracy, completeness and consistency of the scientific data within the report for research outputs.... (Art.9)</li> <li>NSFC shall set up platform(s) for the exchanges of research outputs and ensure the persistency, completeness and openness of research resources. Relying units should share valuable research outputs (i.e. databases, samples, research facilities) gratis or no gratis. (Art.13)</li> </ul>
<b>MOST&amp;MOF (2017):</b> <b>Interim Measures for the Administration of National Key R&amp;D Programs</b> <sup>14</sup>	<ul style="list-style-type: none"> <li>Research information disclosure (Art.4).</li> <li>Establishment and implementation of Public Service Platform for National Science and Technology Information System (Art.5)</li> <li>Assign responsibilities of project outcomes application and information sharing to MOST, Program management institutions, PI affiliated institutions et al. (Art. 8/11/12/39)</li> </ul>
<b>MOST, NDRC, MOF (2017):</b> <b>Measures for Managing National S &amp;T Major Projects</b> <sup>15</sup>	<ul style="list-style-type: none"> <li>Establishment of an information management platform and sub-branches for major special programs and merged into the National Science and Technology Management Information System (Art.53).</li> <li>Detail definition of “Information contents” in which all research workflow is included (Art.54) together with information security rules (Art.59).</li> </ul>
<b>MOST, NDRC, MOF (2017):</b> <b>Administrative Measures for the Opening and Sharing of National Major Scientific Research Infrastructures and Large-scale Scientific Research Instruments</b> <sup>16</sup>	<ul style="list-style-type: none"> <li>Scientific research facilities and equipment should be open to the public ...except as otherwise provided by law (Chap.1, Art.5).</li> <li>Responsibilities for managing units to take (i.e. MOST, MOF) to share research facilities as well as to provide guidance for assessment, awards &amp; punishments are addressed in detail (Chap.2 &amp; 4).</li> </ul>

**Table 3** Typical cases of program-level rules governing research data sharing in China.

Furthermore, especially after the release of the national-level rules, “Measures for Managing Scientific Data” (2018), the “data submission agreement” is compulsory for MOST grant

<sup>12</sup> “13th Five-Year” National Science and Technology Innovation Plan, State Council, P.R.C., Chap.18, Art.2(2016). Available at: [http://www.gov.cn/zhengce/content/2016-08/08/content\\_5098072.htm](http://www.gov.cn/zhengce/content/2016-08/08/content_5098072.htm) [Accessed: 10 Sep 2020].

<sup>13</sup> Measures for the Management of Research outputs in Projects Funded by the National Natural Science Foundation of China, Natural Science Foundation of China, Art.9, 13 (2015). Available at: <http://www.nsfc.gov.cn/publish/portal0/tab475/info73668.htm> [Accessed: 10 Sep 2020].

<sup>14</sup> Interim Measures for the Administration of National Key R&D programs, Ministry of Science and Technology, P.R.C. and Ministry of Finance, P.R.C., Art.4–5,8,11,12,39 (2017): Available at: [http://www.most.gov.cn/xxgk/xinxifenlei/fdzdgnr/fgzcfjxwj/gfxwj2017/201706/t20170628\\_133796.html](http://www.most.gov.cn/xxgk/xinxifenlei/fdzdgnr/fgzcfjxwj/gfxwj2017/201706/t20170628_133796.html) [Accessed: 10 Sep 2020].

<sup>15</sup> Measures for Managing National S &T Major Projects, Ministry of Science and Technology, P.R.C., National Development and Reform Committee, P.R.C., Ministry of Finance, P.R.C., Art.53–54, 59(2017). Available at: [http://www.most.gov.cn/xxgk/xinxifenlei/fdzdgnr/fgzcfjxwj/gfxwj2017/201706/t20170627\\_133757.html](http://www.most.gov.cn/xxgk/xinxifenlei/fdzdgnr/fgzcfjxwj/gfxwj2017/201706/t20170627_133757.html) [Accessed: 10 Sep 2020].

<sup>16</sup> Administrative Measures for the Opening and Sharing of National Major Scientific Research Infrastructures and Large-scale Scientific Research Instruments, Ministry of Science and Technology, P.R.C., National Development and Reform Committee, P.R.C., Ministry of Finance, P.R.C., Chap. 1,2,4 (2017): Available at: [http://www.gov.cn/gongbao/content/2018/content\\_5257406.htm](http://www.gov.cn/gongbao/content/2018/content_5257406.htm) [Accessed: 10 Sep 2020].

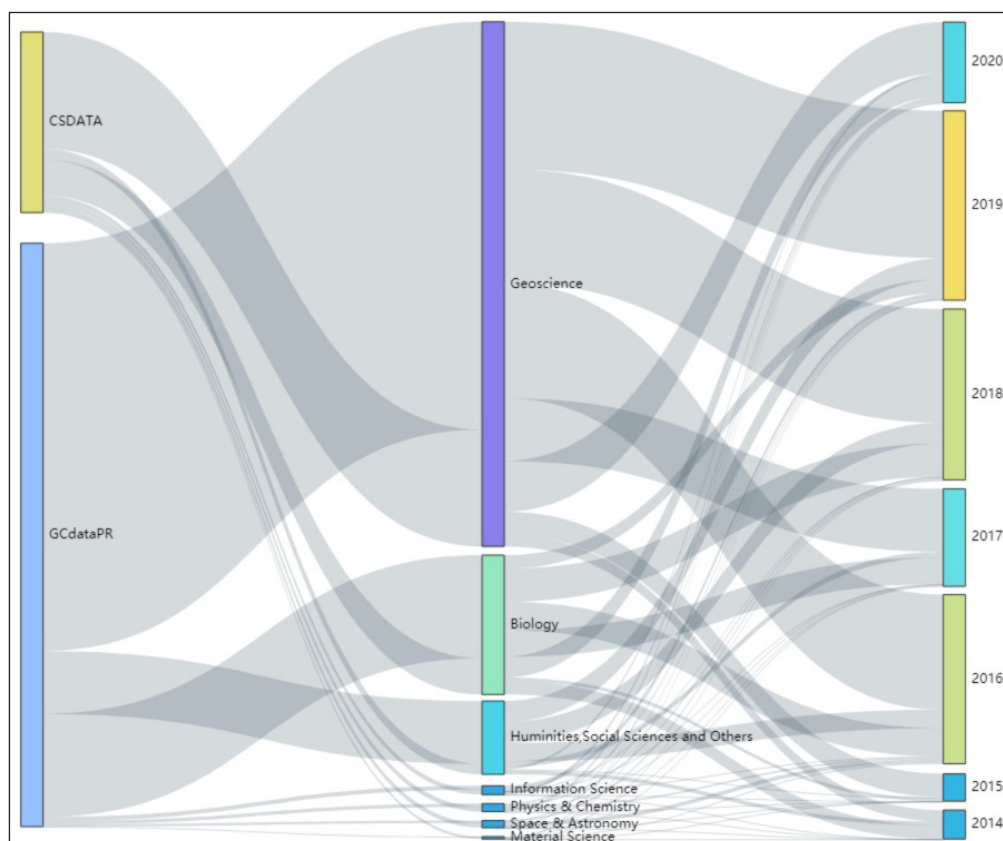
programs (Fan, 2018). This requirement can be traced back to early 2016 (MOST, 2016) and the main purpose is to reach legal agreements between the funding agency (MOST), individual researchers, and their affiliated institutions to guarantee complete and on-time data capture and submission to the MOST platforms during the research and data sharing that commences after an embargo period. Such conditions are considered mandatory requirements during proposal review and award processes.

### 3.2 IN DATA REPOSITORIES

Data repositories, such as data centers and archives, often provide users with integrated data platforms that offer data curation capabilities to enable efficient data publication. Data services can include in-depth quality control as well as data sharing. Some data repositories also provide linkage between datasets, data papers, and publications (Karcher, Kirilova and Weber, 2016). Many data repositories specialize in providing services to specific disciplines, such as the geoscience data repository, GSCloud ([www.gscloud.cn](http://www.gscloud.cn)), omics data center, GSA ([bigd.big.ac.cn/gsa/](http://bigd.big.ac.cn/gsa/)), and institutional repositories (i.e., Peking University Open Research Data Platform, [opendata.pku.edu.cn](http://opendata.pku.edu.cn)), as well as facilities serving the general public (i.e. ScienceDB, [www.sciencedb.cn](http://www.sciencedb.cn)).

### 3.3 IN DATA JOURNALS

In addition to data published as supplements to scholarly journals, data journals, which jointly publish data papers with datasets, have become popular. *China Scientific Data* ([www.csdata.org](http://www.csdata.org)), which was established in 2015 and began releasing data in 2016, and *Global Change Research Data Publishing & Repository* ([www.geodoi.ac.cn](http://www.geodoi.ac.cn)), which established online services in 2014 and began publishing in 2017, take the lead in promoting FAIR data by publishing data papers and datasets. Until June 2020, *China Scientific Data* has<sup>17</sup> shared over 224 data papers and datasets, from across sundry disciplines, with around 300,000 page views yearly. Among the represented disciplines, geoscience and biology data sharing rank highly. In addition, *Global Change Research Data Publishing & Repository* (GCRDP) has published 267 GB of datasets online with over 245,000 data files downloaded cumulatively.<sup>18</sup> Distribution of disciplinary data articles in the two journals is listed in [Figure 2](#). Furthermore, *GigaScience* also publishes datasets



**Figure 2** Disciplinary distribution of data articles in CSDATA & GCdataPR.<sup>17,18</sup>

<sup>17</sup> Statistical data available at: [http://www.csdata.org/en/p/paper\\_search/?q=](http://www.csdata.org/en/p/paper_search/?q=) [Accessed: 31 Jul 2020].

<sup>18</sup> Statistical data available at: <http://www.geodoi.ac.cn/WebEn/> [Accessed: 31 Jul 2020].



with joint efforts from GigaDB (Sneddon, Li and Edmunds, 2012). Also, *Big Earth Data* includes geoscience big data publishing as part of their scope (Guo, 2017).

In addition, the implementation of norms for data also guarantees and facilitates the reuse of data. According to the CNKI standard database,<sup>19</sup> there are over 700 data-related works of national-level and disciplinary standards, which contribute to the quality of data throughout the data life cycle. Such data norms include quality control for metadata and data (i.e. “Information Technology – Big Data – Terminology GB/T 35295-2017”), methods for data security (i.e. “Information Security Technology – Personal Information Security Specification GB/T 35273-2017”), data processing, exchange & communication (i.e. “Technical Specification for Environmental Thematic Spatial Data Processing HJ 927-2017”, “Specification for Drafting Basic Dataset of Ecology and Environment Information HJ 966-2018”), data sharing (i.e. “Information Technology Big Data Governmental Data Sharing Part 1: General Provisions GB/T 38664.1-2020”), data metrics and evaluation (Zhang and Yuan, 2017; Ou and Zhou, 2016), as well as information systems (Briney, Goben and Zilinski, 2017; Bao, Fan and Li, 2015) and information technology (i.e. “Information Technology – Big Data – Technical Reference Model GB/T 35589-2017”) for data.

### 3.4 IN CITIZEN SCIENCE

Citizen science techniques offer another way of collecting open data, leveraging multiple contributions to capture and analyze data. This approach is becoming popular for researchers who study daily life and have access to data collection communities, such as Birdnet ([www.birdnet.cn](http://www.birdnet.cn)) and the Chinese Field Herbarium (CFH, [www.cfh.ac.cn](http://www.cfh.ac.cn)). Records in the “Database of Cetacean Stranding Records around Hainan Island” are also partly contributed by volunteers (Zhang, Li, Lin, et al, 2016). However, data exchanges are not enough, as some citizen scientists go even further. For example, some contribute to the exploration of new species or participate in a non-governmental organization (NGO) for more citizen science data collection and analysis opportunities (Science China, 2014).

## 4. RESPONSIBLE DRIVING FORCES

Understanding the driving forces behind data policy efforts offers insight into the intentions and objectives of data policy initiatives (Braun, 1998). It appears that key national data policymakers in China include, but may not be limited to the MOST, the Chinese Academy of Sciences (CAS), the China Association for Science and Technology, and the National Natural Science Foundation of China (Qiu, 2014; ChinaAccess4EU, n.d.). These agencies have promoted improvements in data practices to facilitate the implementation of data policies (Noorden, 2014).

### 4.1 MINISTRY OF SCIENCE AND TECHNOLOGY (MOST)

As is illustrated in the discussion of institutional data policies, being one of the major funding agencies for research, the MOST<sup>20</sup> takes the lead in pushing open data across domains. Since its initial efforts in 2001, the MOST began supporting the NSTI program with the initial establishment of 13 scientific data centers covering agriculture, forestry, seismicity, meteorology, marine science, Earth systems, population and health, biology, chemistry, materials, and energy, as well as others.<sup>21</sup> In recent years, the funding mechanism has evolved from pre-funding awards to subsidies for further development of selected data portals to foster data sustainability (MOF, 2013). Design, establishment, and evaluation of national scientific data centers is another major contribution enhancing the generic service capability of national data infrastructures facing open data and open science currently (MOST and MOF, 2019).

### 4.2 CHINESE ACADEMY OF SCIENCES (CAS)

As one of the most important and largest research institutions in China, CAS takes the lead in promoting the production of research data, contributing substantially to the efforts of the science community in different disciplines (i.e. Chen, 2018; Zhang, 2018). Among all of the

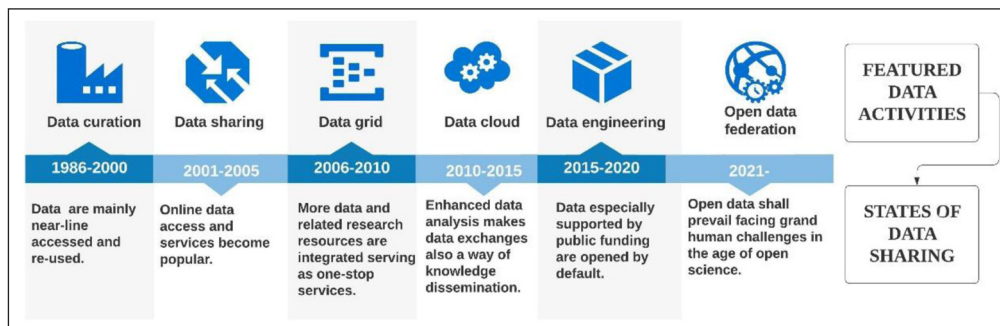
<sup>19</sup> CNKI standard database, available at: <https://kns.cnki.net/kns8?dbcode=CISD> [Accessed: 31 Jul 2020].

<sup>20</sup> Ministry of Science and Technology, P.R.C. Available at: <http://en.most.gov.cn/eng/index.htm> [Accessed: 11 Sep 2020].

<sup>21</sup> NIST programs. Available at: <https://escience.org.cn/> [Accessed: 28 Aug 2018].

CAS programs, the Scientific Database Program (SDP) has been focusing on research data generation, curation, and sharing as its primary goals since 2006 (CESDCAS, 2009).

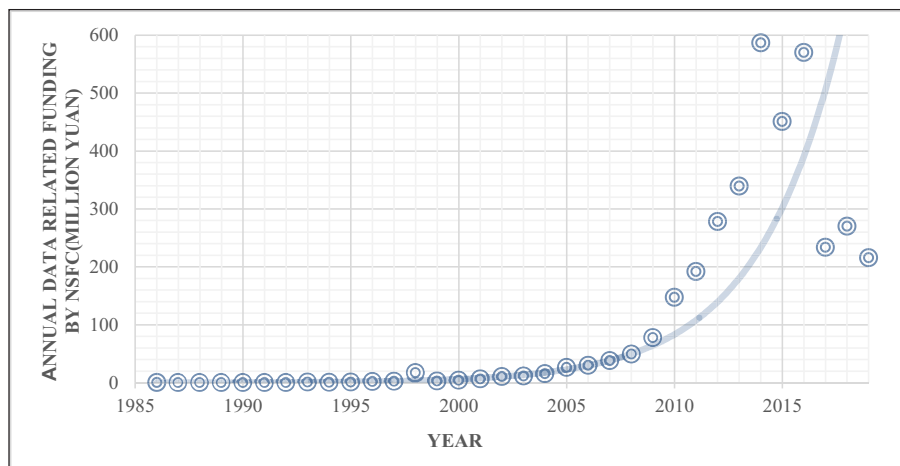
Before 2000, data was disseminated offline, near-line, and, to some extent, online. The SDP program mainly focused on expanding the scale of research data and curation to improve data management and data sharing capabilities (Zhang and Li, 1997). Subsequently, during the next five years, data sharing capabilities were emphasized as online data services were developed further (Gui, 2007). Then, data-sharing efforts expanded considerably through 2010, as the volume of research data increased and the data grid was employed to integrate data from different sources to foster new scientific discoveries (CIOCAS, 2011; Hu, et al, 2012). From 2011 to 2015, the adoption of the data cloud has provided an opportunity for attaining flexible, but more robust data infrastructures, and also for supporting value-added data analysis (Liu, et al, 2016). By the end of 2015, multidisciplinary scientific data volume had reached 655 TB with over 96,290,000 unique visits and 456 TB downloads in total (Li, et al, 2016). The following “Big Data Engineering” program continues the trend of open data based on consolidating engineering construction (CAS, 2017). And the “Measures of Managing Scientific Data” further clarify the general duty to provide open, public-funded data by default. These actions leave no doubt that open data shall prevail across domains and regions in accordance with the open science paradigm. The evolving history of SDP is depicted in [Figure 3](#).



**Figure 3** Data management and sharing in CAS Scientific Database Program.

### 4.3 NATIONAL NATURAL SCIENCE FOUNDATION OF CHINA (NSFC)

As one of the vital funding agencies in China, the NSFC supports the sharing of research outcomes, including data. In 2014, NSFC jointly announced, with CAS, open-access rules for all their programs, providing direction to release articles as open access, with an embargo of no more than 12 months after publishing (OpenAIRE, 2014). Moreover, during the last 30 years, data-intensive research has gained notable support from the NSFC through approximately 6,000 different programs. With up to 3.58 billion yuan in total, the annual tendency for data-related grants is depicted in [Figure 4](#).<sup>22</sup> The NSFC also supports data stewardship activities as a necessary part of other types of research. Incidentally, the NSFC has been merged into the MOST in 2018 (Cyranoski, 2018; NSFC, 2018), but will continue to be one of the major funding agencies, especially for fundamental research.



**Figure 4** Tendency of Yearly NSFC Funding on Data Topics.<sup>22</sup>

<sup>22</sup> NSFC, available at <http://kd.nsfc.gov.cn/baseQuery/supportQuery> [accessed 15 Jul 2020], programs entitled with “data” are selected and their grants are totaled by year. Program types on this website have been re-categorized since the initial article submission, thus data from 2017–2019 have been manually totaled for the revised paper.

#### 4.4 CHINA ASSOCIATION FOR SCIENCE AND TECHNOLOGY (CAST)

CAST<sup>23</sup> is the largest society of science and technology (S&T) professionals that operates as a non-governmental organization in China. The focus of CAST includes various topics, like database sharing and data exchanges. In the typical project, titled “Discipline development in CAST member societies”, data have been recognized as the main source of research outcomes, and sharing of databases for scientific research, provided by over 200 national-level academic societies, has been highlighted in particular (CAST, n.d.).

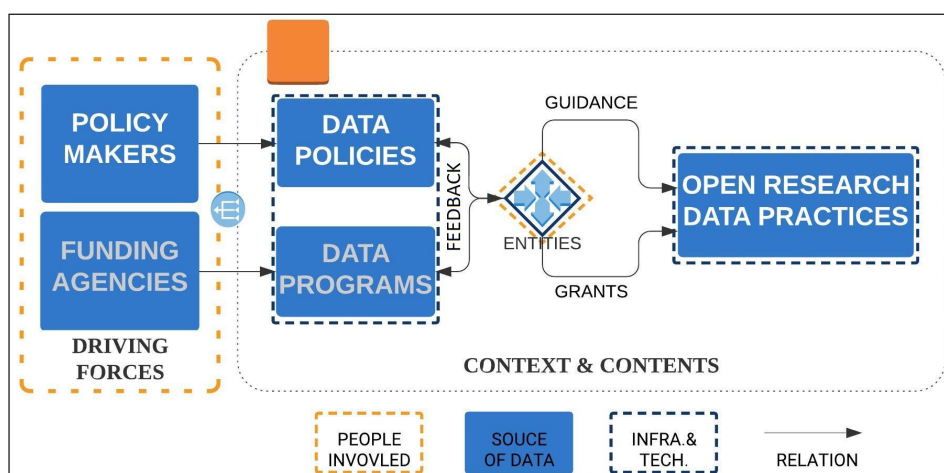
Moreover, other driving forces include international organizations in China, such as CODATA China, and the World Data System (WDS). These international organizations and their members promote regional data exchanges through comprehensive rules, showcases, trainings, and workshops, as well as through other innovative ways to facilitate lifecycle data sharing and communication. Furthermore, local administrative departments, research associations and their sub-branches also serve as stakeholders, contributing to the driving forces within the data ecosystem. While the importance of people and the research community is implied in the discussion of such organizational stakeholders, it is vital to emphasize that data producers, data stewards, and data users are the stakeholders whose efforts and needs are of paramount importance and necessary for the development and evolution of an ecosystem for curating and sharing open data.

### 5 LANDSCAPE OF OPEN RESEARCH DATA ECOLOGY IN CHINA

In China, policymakers within national, regional, and local levels develop and disseminate data policies and promote recommended implementation practices, while funding agencies provide support to improve open data practices via various data programs, which often will request for additional funding. Along with the research community, China policymakers and funding agencies have served as catalysts for much of the evolution of open data practices in China and their decisions guide the mainstream of the data ecology. Therefore, we pinpoint them as the driving forces who initialize the open data cycle in data ecology.

In practice, data policies serve as guidance for developing and operating data programs; the latter provides data curation and sharing experiences and exposes problematic issues to be considered for revising data policies (Ragusa, 2010; World Bank’s open government data working group, 2013). Furthermore, open research data practices offer insight into the adoption of open data policies by providing feedback about enforcement of those policies. Data programs support data practices with direct funding and strict program guidance, including data requirements. Moreover, we can take policy, program, and practice together as a compound of context and contents for open data ecology, since they are usually reflected through entity actions simultaneously. In effect, certain open data practices simultaneously decode the policy constraints and program support to afford progress.

Essentially, driven by the initiatives of policymakers and funding agencies, the three primary elements that contribute to the open data ecology in China are data policies, data programs, and open research data practices (Li and Zhang, 2017). These interactions are depicted in [Figure 5](#).



**Figure 5** Landscape of open research data ecology in China.

<sup>23</sup> China Association for Science and Technology profile. Available at: <http://english.cast.org.cn/col/col471/index.html> [Accessed: 11 Sep 2020].

Comprehensive data policies employ a bottom-up approach driven by the rapid growth of research data scales and the recognition of community needs. Some provinces (i.e. Shandong, Henan, Hainan, Guizhou, etc.) have taken the lead to set up independent bureaus for data-driven city governance. Moreover, data programs serve as effective data engines to support data production with grants. Such programs also emphasize issues about data sharing and data management, encouraging clarification to differentiate between restricted and open data, and stimulating the development of strategies for managing sensitive data to address issues of privacy and public safety. Data policies are being developed to guide such data practices and are expected to be implemented fully as data practices evolve over time. Although the ecosystem for data policy is just beginning to flourish, the flexible approach of many data programs in China is fostering the acquisition of research resources and raises awareness of the need to continually improve scientific data management and sharing practices.

## 6 DISCUSSION AND CONCLUSION

In essence, the formation of an ideal and harmonious data ecology is under construction in China. Data policies have evolved from constitutions of general legislation to independent rules focusing mainly on publicly-funded data governance. Also, like the evolution of data sharing practices internationally (Fecher, Friesike and Hebing, 2015), the focus in China has shifted from primarily data management to both data governance and data sharing throughout the entire scientific data lifecycle. Furthermore, like the increasing acceptance of data sharing practices among researchers internationally (Tenopir, et al, 2015), we can observe that the gap between positive attitudes of sharing data and not-so-active data sharing behaviors in China is shrinking (Li, et al, 2016; NSTI, MOST, 2019). Relatedly, as improvements to data management and data sharing policies are occurring in the United States to face the complexities of confidentiality, privacy, and intellectual property concerns (Doshi, et al, 2016), China also spares no efforts in the governance of those areas, as analyzed earlier in this paper. These improvements in open data sharing policies and practices in China also appear to be consistent with observed trends in open knowledge practices in China (Montgomery and Ren, 2018). Stronger participation in open access practices among research communities (Zhang, 2014), as well as regular and close collaboration (Joseph, 2016), should be encouraged across disciplines, including the natural sciences and social sciences. While challenges remain for scientific data management and data sharing in China and internationally, cultural norms for data sharing appear to be improving and these may work together with data policies and data practices to create a friendly data culture that also nourishes the whole society (Zuidervijk and Janssen, 2014).

Moreover, as global public goods (CODATA, 2019), the use of research data is embracing the grand human and societal challenges and the sharing of data has become a vital part of research collaboration. Based on the review of evolving data policies and data practices in China, we see the future of open data progressing across the international realm in the following manner:

- Open data is being adopted into mainstream practices, but faces many challenges in various sectors. The outbreak of COVID-19 presents a compelling case for the adoption of data sharing, globally: the leading role of the World Health Organization (WHO); the re-affirmation to share data during COVID-19 by almost 160 key stakeholders, internationally, including government agencies, research institutions, publishers, universities, and companies (Wellcome, 2020); the emergence of ongoing international efforts like the “RDA COVID-19 Recommendations and Guidelines on Data Sharing” (RDA COVID-19 Working Group, 2020), the 2019-nCoV related genomic databases released from Chinese Academy of Sciences (CAS, 2020), and the cases and testing data collection shared by the Johns Hopkins Coronavirus Resource Center and many other institutions around the world. These and other endeavors that have been initiated and the data that is shared demonstrate how we have united for the sake of the global village. However, there are still obstacles that must be addressed, such as privacy, security, intellectual property, and other ethical concerns. Therefore, more innovation is needed to maintain a balance between closed and open data.
- Stable and robust open data ecology calls for diversity. As we know, diversity can contribute to the stability and resilience of an ecosystem. And, when considering an

open data ecology, diversity may refer to: (1) Involvement of diversified stakeholders. As illustrated partly in [Figure 4](#), stakeholders within an open data ecology may include driving forces (such as policymakers, funding agencies), affiliated institution(s), individual researchers, as well as publishers, and research associations, etc. Participation of different stakeholders jointly contributes to the ecological balance through optimal allocation of resources throughout the entire data lifecycle. (2) Multi-factor open data governance. Impelling data programs and data policies, together with various data practices, all serve as compulsory factors in an open data ecology. And effective operation of the open data ecosystem may require an intermediate approach in which there's top-down policies of regional and national-level guidance, but also bottom-up exemplars to follow. The routine of following general data practices when developing data policy might be cost-effective, but some data policy development could take the lead, if necessary (Timney, 2002). (3) Various open data models. Diversity also could mean that there's no fixed-up way to facilitate the sharing and use of open data, but better exploration of information technology may help. For instance, the application of blockchain technology (Li, et al, 2017; Naz, et al, 2019) bypasses access to sensitive datasets, and shares models, algorithms, and results online without the need to migrate local data, which can make sharing sensitive data possible without violating complex data protection rules.

- Under the umbrella of open science, open data also needs to be curated through efficient data facilities. Such facilities should be networked across international research efforts, interoperable between platforms, and integrated to leverage different research resources. The European Open Science Cloud, the United States National Research Cloud, the Chinese Science & Technology Cloud, and the African Open Science Platform, as well as others, are current examples. Such data infrastructures serve as integrated platforms that foster the implementation of data policy and technical support of data programs, while also serving as consolidated bases for various data practices. But that's not enough. If we envision a future with open data at larger scales and better linkage across regions and domains, then we need improved data infrastructure to guarantee capabilities for connecting sundry research resources together. The idea of building up a global open science cloud is still at an early stage with active participants from the CNIC, CAS, EGI Foundation, and their African partners as well as other partners around the world.
- According to our field observations, fairness, trust, and sustainability are essential ingredients that involve many interdependent components within an open data ecology. Fairness is necessary for data policy-making and program design. Likewise, for data practices, fairness provides general principles for facilitating the use of open data, as previously mentioned in terms of FAIR capabilities. TRUST also is necessary for implementing capabilities to adopt data policies and programs while ensuring the continuing quality of data for reuse and protecting the legal rights of contributors. Such efforts also can improve trust between data producers and data re-users. Often, data programs survive on trust, especially when data policies and programs are under development. With trust, successful data transactions are possible. Providing such systems sustainably also requires efficiency, while offering incentives and metrics for assessing data policies and programs, along with the business models that have been adopted to foster enduring data practices.

## APPENDIX: LIST OF OPEN RESEARCH DATA RELATED POLICIES IN CHINA

Law of the People's Republic of China on Science and Technology Progress (《中华人民共和国科学技术进步法》)

Copyright Law of the People's Republic of China (《中华人民共和国著作权法》)

Law of the People's Republic of China on Promoting the Transformation of Scientific and Technological Achievements (《中华人民共和国促进科技成果转化法》)

Cybersecurity Law of the People's Republic of China (《中华人民共和国网络安全法》)

Measures for Managing Scientific Data (《科学数据管理办法》)



Data Security Law of the People's Republic of China (Draft Version) (《中华人民共和国数据安全法(草案)》)

Regulations of the People's Republic of China on Disclosure of Government Information (《中华人民共和国政府信息公开条例》)

Interim Measures for the Management and Sharing of Government Information resources (《政务信息资源共享管理暂行办法》)

Interim Provisions on the Administration of National S&T Plans (《国家科技计划管理暂行规定》)

Interim Measures for the Management of National S&T Plan Projects (《国家科技计划项目管理暂行办法》)

Interim Provisions on Strengthening the Management of Achievements of the National Science and Technology Plan (《关于加强国家科技计划成果管理的暂行规定》)

Measures for the Management and Sharing of Scientific Data of the Chinese Academy of Sciences (for Trial Implementation) (《中国科学院科学数据管理与开放共享办法(试行)》)

Measures for the Management of Open Access and Scientific Data of the Chinese Academy of Agricultural Sciences (《中国农业科学院开放获取与科学数据管理办法》)

Measures for Managing Polar Expedition Data of China (《中国极地考察数据管理办法》)

Interim Measures for managing Land and Resources Data (《国土资源数据管理暂行办法》)

Interim Provisions on National Health Care Big Data Standards, Safety and Service Management (《国家健康医疗大数据标准、安全和服务管理办法(试行)》)

Fengyun Meteorological Satellite Data Management Measures (for Trial Implementation) (《风云气象卫星数据管理办法(试行)》)

Measures for Meteorological Information Exchange and Management (《气象探测资料汇交管理办法》)

Data Sharing and Service Procedures in Marine Ecological Environment Monitoring (for Trial Implementation) (《海洋生态环境监测数据共享服务程序(试行)》)

Measures for Managing Seismological Science Data (《地震科学数据共享管理办法》)

Interim Measures for the Management of National Civil Satellite Remote Sensing Data (《国家民用卫星遥感数据管理暂行办法》)

Measures for Managing Scientific Data in Lunar and Deep Space Exploration (《月球与深空探测工程科学数据管理办法》)

"13th Five-Year" National Science and Technology Innovation Plan (《“十三五”国家科技创新规划》)

Measures for the Management of Research outputs in Projects Funded by the National Natural Science Foundation of China (《国家自然科学基金资助项目研究成果管理办法》)

Interim Measures for the Administration of National Key R&D programs (《国家重点研发计划管理暂行办法》)

Measures for Managing National S & T Major Projects (《国家科技重大专项(民口)管理规定》)

Administrative Measures for the Opening and Sharing of National Major Scientific Research Infrastructures and Large-scale Scientific Research Instruments (《国家重大科研基础设施和大型科研仪器开放共享管理办法》)

Measures for the Administration of the National S & T Resources Sharing and Services Platform (国家科技资源共享服务平台管理办法)

## ACKNOWLEDGEMENTS

This work has been supported by the Chinese Academy of Sciences (No. XDA 19020104/ No. 241711KYSB20200023), the National Science and Technology Infrastructure Center, MOST (No. 2018DDJ1ZZ14), and the Columbia University Center for International Earth Science Information Network (CIESIN).



Portions of this work are based on “Open Research Data Policies and Practices in China”, by LI Jianhui and ZHANG Lili, which was presented to the Board on Research Data and Information/ US CODATA during the symposium, International Coordination for Science Data Infrastructure Board Meeting, at the National Academy of Sciences in Washington, DC on November 1, 2017.

## COMPETING INTERESTS

The authors have no competing interests to declare.

## AUTHOR CONTRIBUTIONS

L. Z. drafted the paper, R. D. revised it, J. L. designed the framework and provided facts, L. W. and C. L. cleaned data and prepared some figures.

## AUTHOR AFFILIATIONS

**Lili Zhang**  [orcid.org/0000-0003-1847-6683](https://orcid.org/0000-0003-1847-6683)

Computer Network Information Center, Chinese Academy of Sciences, Beijing, China

**Robert R. Downs**  [orcid.org/0000-0002-8595-5134](https://orcid.org/0000-0002-8595-5134)

Center for International Earth Science Information Network, Columbia University, Palisades, NY, USA

**Jianhui Li**

Computer Network Information Center, Chinese Academy of Sciences, Beijing, China

**Liangming Wen**  [orcid.org/0000-0003-0284-1527](https://orcid.org/0000-0003-0284-1527)

Computer Network Information Center, Chinese Academy of Sciences, Beijing, China

University of Chinese Academy of Sciences, Beijing, China

**Chengzan Li**

Computer Network Information Center, Chinese Academy of Sciences, Beijing, China

## REFERENCES

- Adamick, J, Reznik-Zellen, RC and Sheridan, M.** 2012. Data Management Training for Graduate Students at A Large Research University. *Journal of eScience Librarianship*, 1(3): e1022. DOI: <https://doi.org/10.7191/jeslib.2012.1022>
- Administrative Office, Chinese Academy of Sciences (AOCAS).** 2018. The Chinese Academy of Sciences Held A Video Conference on the “Scientific Data Management Measures”. Available at: [http://www.cas.cn/sygz/201807/t20180719\\_4658696.shtml](http://www.cas.cn/sygz/201807/t20180719_4658696.shtml) [Accessed: 13 Sep 2018] (In Chinese).
- Bao, D, Fan, Y and Li, M.** 2015. Data Governance and Framework of University Libraries. *Library and Information Service*, 18: 134–141. DOI: <https://doi.org/10.13266/j.issn.0252-3116.2015.18.020> (In Chinese).
- Boeckhout, M, Zielhuis, GA and Bredenoord, AL.** 2018. The FAIR Guiding Principles for Data Stewardship: Fair Enough? *European Journal of Human Genetics*, 26: 931–936. DOI: <https://doi.org/10.1038/s41431-018-0160-0>
- Braun, D.** 1998. The Role of Funding Agencies in The Cognitive Development of Science. *Research Policy*, 27(8): 807–821. DOI: [https://doi.org/10.1016/S0048-7333\(98\)00092-4](https://doi.org/10.1016/S0048-7333(98)00092-4)
- Briney, K, Goben, A and Zilinski, L.** 2017. Institutional, Funder, and Journal Data Policies. In: Johnston, LR. (ed.) *Curating Research Data: Practical Strategies for Your Digital Repository*. Chicago: Association of College and Research Libraries A division of the American Library, pp. 61–78.
- CAS.** 2017. The 13th Five-Year Informatization Special Science Big Data Project of the Chinese Academy of Sciences: Guideline for “Big Data Driven Discipline Innovation Demonstration Platform” proposal. Available at: <http://m.cas.cn/tzgg/201709/W020170906345809917086.pdf> [Accessed: 28 Aug 2018] (In Chinese).
- CAS.** 2020. CAS Research Institutions Develop Three Databases for 2019–nCoV Research. Available at: <http://international-talent.cas.cn/front/mobi/newInfo.html?id=f7452efccb11d7242900fa2bd7f895ce> [Accessed: 17 Aug 2020]
- Charbonneau, DH and Beaudoin, JE.** 2015. The State of Data Guidance in Journal Policies: A Case Study in Oncology. *International Journal of Digital Curation*, 10(2): 136–156. DOI: <https://doi.org/10.2218/ijdc.v10i2.375>
- Chen, N.** 2018. Global Network to Use Big Data to Protect Earth. Available at: [http://english.cas.cn/newsroom/news/201802/t20180213\\_190117.shtml](http://english.cas.cn/newsroom/news/201802/t20180213_190117.shtml) [Accessed: 16 Jul 2018].
- China Association for Science and Technology (CAST, n.d.).** Disciplinary development. Available at: <http://english.cast.org.cn/n214236/c346026/content.html> [Accessed: 11 Jun 2018].

- ChinaAccess4EU.** n.d. National Funding Programs in Mainland China. Available at: <http://www.access4.eu/China/274.php> [Accessed: 14 Jul 2018].
- CODATA.** 2019. The Beijing Declaration on Research Data. Available at: [www.codata.org/uploads/BeijingDeclaration-19-11-07-FINAL.pdf](http://www.codata.org/uploads/BeijingDeclaration-19-11-07-FINAL.pdf) [Accessed: 17 Aug 2020].
- Committee of Experts on Scientific Database, Chinese Academy of Sciences (CESDCAS).** 2009. Resource Integration and Sustainable Development of Scientific Databases, Chinese Academy of Sciences. *eScience*, 1: 1–5.
- Consultative Committee for Space Data Systems (CCSDS).** 2012. Reference Model for An Open Archival Information System (OAIS), Recommended Practice CCSDS 650.0-M-2. Also published as ISO 14721. Available at: <https://public.ccsds.org/Pubs/650x0m2.pdf> [Accessed: 28 Aug 2018].
- Corpas, M, Kovalevskaya, NV, McMurray, A,** et al. 2018. A FAIR Guide for Data Providers to Maximise Sharing of Human Genomic Data. *PLoS Computational Biology*, 14(3). DOI: <https://doi.org/10.1371/journal.pcbi.1005873>
- Cyber Infrastructure Office, Chinese Academy of Sciences (CIOCAS).** 2011. Chinese Academy of Sciences Informatization development report 2011. Available at: <http://www.ecas.cas.cn/xxkw/fzbg/201505/P020150526358132242288.pdf> [Accessed: 14 Jul 2018].
- Cyranoski, D.** 2018. China Creates Grand Science Ministry. *Nature*, 2018, 555: 425–426. DOI: <https://doi.org/10.1038/d41586-018-03246-w>
- DataONE.** 2018. Data Management Skill Building Hub. Available at: <https://dataoneorg.github.io/Education/> [Accessed: 17 Jul 2018].
- Doshi, J, Hendrick, F, Graff, J,** et al. 2016. Data, Data Everywhere, but Access Remains a Big Issue for Researchers: A Review of Access Policies for Publicly Funded Patient-Level Health Care Data in the United States. Washington, DC: EGEMS, 4(2). DOI: <https://doi.org/10.13063/2327-9214.1204>
- Fan, Q.** 2018. Changes in 2018 Key Special Program Proposal. Available at: [http://www.stdaily.com/kjzc/jiedu/2018-04/18/content\\_660751.shtml](http://www.stdaily.com/kjzc/jiedu/2018-04/18/content_660751.shtml), 29 Apr 2018 [Accessed: 27 Aug 2018].
- Fecher, B, Friesike, S and Hebing, M.** 2015. What Drives Academic Data Sharing? *PLoS ONE*, 10(2): e0118053. DOI: <https://doi.org/10.1371/journal.pone.0118053>
- Federer, LM, Belter, CW, Joubert, DJ,** et al. 2018. Data Sharing in PLOS ONE: An Analysis of Data Availability Statements. *PLoS ONE*, 13(5): e0194768. DOI: <https://doi.org/10.1371/journal.pone.0194768>
- Fong, BL and Wang, M.** 2015. Required Data Management Training for Graduate Students in an Earth and Environmental Sciences Department. *Journal of eScience Librarianship*, 4(1): e1067. DOI: <https://doi.org/10.7191/jeslib.2015.1067>
- Gerunov, A.** 2017. Understanding Open Data Policy: Evidence from Bulgaria. *International Journal of Public Administration*, 40(8): 649–657. DOI: <https://doi.org/10.1080/01900692.2016.1186178>
- Gui, W.** 2007. Facing the New Stage of Scientific Databases Development: Review and Reflection on The Development of Scientific Databases in Chinese Academy of Sciences for 20 Years. *Bulletin of Chinese Academy of Sciences*, 22(1): 83–85 (In Chinese).
- Guo, H.** 2017. Big Earth Data: A new Frontier in Earth and Information Sciences. *Big Earth Data*, 1(1–2): 4–20. DOI: <https://doi.org/10.1080/20964471.2017.1403062>
- Holub, P, Kohlmayer, F, Prasser, F,** et al. 2018. Enhancing Reuse of Data and Biological Material in Medical Research: From FAIR to FAIR-Health. *Biopreservation and Biobanking*, 16(2). DOI: <https://doi.org/10.1089/bio.2017.0110>
- Hu, L, Li, J, Liu, N,** et al. 2012. Practice and Some Thoughts on Quality of Scientific Data. *eScience*, 3(2): 10–18 (In Chinese).
- Joseph, H.** 2016. The Evolving U.S. Policy Environment for Open Research Data. *Information Services & Use*, 36: 45–48. DOI: <https://doi.org/10.3233/ISU-160806>
- Karcher, S, Kirilova, D and Weber, N.** 2016. Beyond the Matrix: Repository Services for Qualitative Data. *International Federation of Library Associations and Institutions*, 42(4): 292–302. DOI: <https://doi.org/10.1177/0340035216672870>
- Liu, F, Chen, X, Li, J,** et al. 2016. Large Scale Distributed Scientific Data Management and Service Technology Framework and System. *Big Data Research*, 2(6): 14–24. DOI: <https://doi.org/10.11959/j.issn.2096-0271.2016062> (In Chinese).
- Li, J, Wu, C, Zhang, L,** et al. 2016. Survey and Analysis of Scientific Data Publishing. *China Scientific Data*, 1.
- Li, J and Zhang, L.** 2017. Open Research Data Policies and Practices in China. *Presented to the Board on Research Data and Information/US CODATA Symposium, International Coordination for Science Data Infrastructure Board Meeting*. Washington, DC: National Academy of Sciences. November 2, 2017.
- Li, J, Zou, Y, Hu, L,** et al. 2016. Scientific Data Cloud Construction and Service of Chinese Academy of Sciences. *Big Data Research*, 2(6): 3–13. DOI: <https://doi.org/10.11959/j.issn.2096-0271.2016061> (In Chinese).
- Li, Y, Huang, J, Qin, S,** et al. 2017. Big Data Model of Security Sharing Based on Blockchain. *2017 IEEE 3rd International Conference on Big Data Computing and Communications*, Apr. 2017, San Francisco Bay, USA. DOI: <https://doi.org/10.1109/BIGCOM.2017.31>

- Lin, D, Crabtree, J, Dillo, I, et al. 2020. The TRUST Principles for Digital Repositories. *Scientific Data*, 7: 144. DOI: <https://doi.org/10.1038/s41597-020-0486-7>
- Mauthner, NS and Parry, O. 2013. Open Access Digital Data Sharing: Principles, Policies and Practices. *Social Epistemology*, 27(1): 47–67. DOI: <https://doi.org/10.1080/02691728.2012.760663>
- Ministry of Finance of the People's Republic of China (MOF). 2013. Measures of Post-subsidy Management for National Science and Technology Plan and Special Funds. Available at: [http://www.mof.gov.cn/mofhome/jiaokewensi/zhengwuxinxi/zhengcefabu/201312/t20131204\\_1020452.html](http://www.mof.gov.cn/mofhome/jiaokewensi/zhengwuxinxi/zhengcefabu/201312/t20131204_1020452.html) [Accessed : 14 Jul 2018] (In Chinese).
- Ministry of Science and Technology (MOST). 2016. Notice of the Ministry of Science and Technology on the Release of the 2016 National Project Application Guidelines for Major National Chronic Non-Communicable Diseases Prevention and Control Research. Available at: [http://www.most.gov.cn/tztg/201604/t20160407\\_125054.htm](http://www.most.gov.cn/tztg/201604/t20160407_125054.htm) [Accessed : 27 Aug 2018].
- MOST, MOF. 2019. The Ministry of Science and Technology and the Ministry of Finance Released a List of Optimized National Science and Technology Resource Sharing Service Platforms. Available at: [http://www.gov.cn/xinwen/2019-06/11/content\\_5399105.htm](http://www.gov.cn/xinwen/2019-06/11/content_5399105.htm) [Accessed: 6 Aug 2020] (In Chinese).
- Montgomery, L and Ren, X. 2018. Understanding Open Knowledge in China: A Chinese Approach to Openness? *Cultural Science Journal*, 10(1): 17–26. DOI: <https://doi.org/10.5334/csci.106>
- Nardi, A and O'Day, VL. 2000. Information Ecologies: Using Technology with Heart. Cambridge, Massachusetts: The MIT Press.
- Natural Science Foundation of China (NSFC). 2018. Nature Talks about China's Restructure of the Ministry of Science and Technology. Available at: <http://www.nsfc.gov.cn/csc/20340/20289/24107/index.html> [Accessed: 31 Jan 2019] (In Chinese).
- Nature Editorial. 2018. Everyone Needs a Data-management Plan. *Nature*, 555: 286. DOI: <https://doi.org/10.1038/d41586-018-03065-z>
- National Science and Technology Infrastructure Center of Ministry of Science and Technology of the People's Republic of China (NSTI, MOST). 2019. National Scientific Data Resources Development Report 2018. Beijing, China: The Scientific and Technological Documentation Press (In Chinese).
- Naughton, L and Kernohan, D. 2016. Making Sense of Journal Research Data Policies. *Insights*, 29(1): 84–89. DOI: <https://doi.org/10.1629/uksg.284>
- Naz, M, Al-zahrani, F, Khalid, R, et al. 2019. A Secure Data Sharing Platform Using Blockchain and Interplanetary File System. *Sustainability*, 11(24): 7054. DOI: <https://doi.org/10.3390/su11247054>
- Noorden, RV. 2014. Funders Punish Open-access Dodgers. *Nature*, 508(7495): 161. DOI: <https://doi.org/10.1038/508161a>
- OECD. 2017. Business Models for Sustainable Research Data Repositories. *OECD Science, Technology and Industry Policy Papers*, No. 47. Paris: OECD Publishing. DOI: <https://doi.org/10.1787/302b12bb-en>
- OpenAIRE. 2014. Chinese Funders Issue OA Policies for Research. Available at: <https://www.openaire.eu/chinese-academy-of-sciences-national-natural-science-foundation-issue-open-access-policies> [Accessed : 16 Jul 2018].
- OpenAIRE. 2015. OpenAIRE Guidelines for Data Archives. Available at: <https://guidelines.openaire.eu/en/latest/data/index.html> [Accessed: 17 Jul 2018].
- Ou, S and Zhou, Y. 2016. Current Status of Scientific Data Curation Research and Practices in Mainland China. *Libres*, 16(1): 73–88.
- Parsons, MA. 2013. Data Policy. *Data Science Journal*, 12: GRDI43–GRDI50. DOI: <https://doi.org/10.2481/dsj.GRDI-008>
- Parsons, MA, Godøy, Ø, LeDrew, E, et al. 2011. A Conceptual Framework for Managing Very Diverse Data for Complex, Interdisciplinary Science. *Journal of Information Science*, 37(6): 555–569. DOI: <https://doi.org/10.1177/0165551511412705>
- Peng, G. 2018. The State of Assessing Data Stewardship Maturity – An Overview. *Data Science Journal*, 17: 7. DOI: <https://doi.org/10.5334/dsj-2018-007>
- Pritchard, MJ and Martel, JC. 2019. Information System Ecology: An Application of Dataphoric Ascendancy. *Information systems*, 1–14. DOI: <https://doi.org/10.1016/j.is.2019.101486>
- Qiu, J. 2014. China Goes Back to Basics on Research Funding. *Nature*, 507(7491): 148–149. DOI: <https://doi.org/10.1038/507148a>
- Ragusa, JM. 2010. The Lifecycle of Public Policy: An Event History Analysis of Repeals to Landmark Legislative Enactments, 1951–2006. *American Politics Research*, 38(6): 1015–1051. DOI: <https://doi.org/10.1177/1532673X10374170>
- RDA COVID-19 Working Group. 2020. RDA COVID-19 Recommendations and Guidelines on Data Sharing (Version 1.0). DOI: <http://doi.org/10.15497/rda00052>
- Science China. 2014. Enlightenment of Citizen Science. Available at: [http://science.china.com.cn/2014-01/06/content\\_31083060.htm](http://science.china.com.cn/2014-01/06/content_31083060.htm) [Accessed: 5 Dec 2018] (In Chinese).
- Shen, C, Riaz, Z, Palle, M, et al. 2015. Open Data Landscape: A Global Perspective and a Focus on China. *14th Conference on eBusiness, e-Services and e-Society (I3E)*, Oct 2015, Delft, Netherlands. Lecture Notes in Computer Science, LNCS-9373, pp. 247–260, 2015, Open and Big Data Management and Innovation. DOI: [https://doi.org/10.1007/978-3-319-25013-7\\_20](https://doi.org/10.1007/978-3-319-25013-7_20)

- Sneddon, TP, Li, P and Edmunds, SC.** 2012. GigaDB: Announcing the GigaScience Database. *GigaScience*, 1(1): 11. DOI: <https://doi.org/10.1186/2047-217X-1-11>
- Tenopir, C, Dalton, ED, Allard, S,** et al. 2015. Changes in Data Sharing and Data Reuse Practices and Perceptions among Scientists Worldwide. *PLoS ONE*, 10(8): e0134826. DOI: <https://doi.org/10.1371/journal.pone.0134826>
- Thielen, J, Samuel, SM, Carlson, J,** et al. 2017. Developing and Teaching a Two-Credit Data Management Course for Graduate Students in Climate and Space Sciences. *Science and Technology Librarianship*, spring issue. DOI: <http://doi.org/10.5062/F42Z13HQ>
- Timney, MM.** 2002. Must Policy Making Wait Until All the Data Are In? An Empirical Look at Environmental Justice. *Public Administration Review*, 62(4): 506–508. Available at: <https://www.jstor.org/stable/3110365> [Accessed: 28 Aug 2018]. DOI: <https://doi.org/10.1111/1540-6210.00204>
- Tollefson, J.** 2018. China Declared Largest Source of Research Articles. *Nature*, 533: 390. DOI: <https://doi.org/10.1038/d41586-018-00927-4>
- UNESCO.** 2015. UNESCO Science Report 2030. Paris, France. Available at <http://unesdoc.unesco.org/images/0023/002354/235407e.pdf> [Accessed: 18 Sep 2018].
- Vita, R, Overton, JA, Mungall, CJ,** et al. 2018. FAIR Principles and the IEDB: Short-term Improvements and a Long-term Vision of OBO-foundry Mediated Machine-actionable Interoperability. *Database*, pp. 1–9. DOI: <https://doi.org/10.1093/database/bax105>
- Wang, X, Guo, Y, Yang, M,** et al. 2015. Information Ecology Research: Past, Present, and Future. *Inf Technol Manag*. DOI: <https://doi.org/10.1007/s10799-015-0219-3>
- Wellcome.** 2020. Sharing Research Data and Findings Relevant to the Novel Coronavirus (COVID-19) Outbreak. Available at: <https://wellcome.ac.uk/coronavirus-covid-19/open-data> [Accessed: 13 Aug 2020].
- World Bank's Open Government Data Working Group.** 2013. Open Data Readiness Assessment Tool. Available at: <https://data.worldbank.org/ogd> [Accessed: 15 Jul 2018].
- Wilkinson, M, Dumontier, M, Aalbersberg, I,** et al. 2016. The FAIR Guiding Principles for Scientific Data Management and Stewardship. *Nature Scientific Data*, 3. DOI: <https://doi.org/10.1038/sdata.2016.18>
- Wyatt, S.** 2017. Making Policies for Open Data: Experiencing the Technological Imperative in the Policy World. *Science Technology & Human Values*, 42(2): 320–324. DOI: <https://doi.org/10.1177/0162243916687645>
- Whyte, A and Wilson, A.** 2010. How to Appraise and Select Research Data for Curation. Edinburgh, UK: Digital Curation Centre. Available at: <http://www.dcc.ac.uk/resources/how-guides> [Accessed: 11 July 2018].
- Zhang, J and Li, W.** 1997. Construction Application and Developing of the Scientific Databases of the Chinese Academy of Sciences. *Modern Library and Information Technology*, 2: 3–9, 15. (In Chinese).
- Zhang, N and Yuan, Q.** 2017. A review of Data Governance Research. *Journal of Intelligence*, 36(5): 131–136, 163 (In Chinese).
- Zhang, P, Li, S, Lin, M,** et al. 2016. Database of Cetacean Stranding Records around Hainan Island (1993–2015). *China Scientific Data* 2. DOI: <https://doi.org/10.11922/csd.170.2015.0029>
- Zhang, X.** 2014. Development of Open Access in China: Strategies, Practices, Challenges. *Insights*, 27(1): 45–50. DOI: <https://doi.org/10.1629/2048-7754.111>
- Zhang, Z.** 2018. Database Resources of the BIG Data Center in 2018. Available at: [http://english.big.cas.cn/rh/rp/201711/t20171103\\_185633.html](http://english.big.cas.cn/rh/rp/201711/t20171103_185633.html) [Accessed: 16 Jul 2018].
- Zuiderwijk, A and Janssen, M.** 2014. Open Data Policies, Their Implementation and Impact: A Framework for Comparison. *Government Information Quarterly*, 31(1): 17–29. DOI: <https://doi.org/10.1016/j.giq.2013.04.003>

#### TO CITE THIS ARTICLE:

Zhang, L, Downs, RR, Li, J, Wen, L and Li, C. 2021. A Review of Open Research Data Policies and Practices in China.. *Data Science Journal*, 20: 3, pp. 1–17. DOI: <https://doi.org/10.5334/dsj-2021-003>

Submitted: 01 February 2019

Accepted: 10 January 2021

Published: 29 January 2021

#### COPYRIGHT:

© 2021 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC-BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. See <http://creativecommons.org/licenses/by/4.0/>.

*Data Science Journal* is a peer-reviewed open access journal published by Ubiquity Press.