



Deposited via The University of Sheffield.

White Rose Research Online URL for this paper:

<https://eprints.whiterose.ac.uk/id/eprint/164614/>

Version: Published Version

Article:

Huang, Y., Cox, A. and Saffi, L. (2021) Research data management policy and practice in Chinese university libraries. *Journal of the Association for Information Science and Technology*, 72 (4). pp. 493-506. ISSN: 1532-2882

<https://doi.org/10.1002/asi.24413>

Reuse

This article is distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs (CC BY-NC-ND) licence. This licence only allows you to download this work and share it with others as long as you credit the authors, but you can't change the article in any way or use it commercially. More information and the full terms of the licence here: <https://creativecommons.org/licenses/>

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

Research data management policy and practice in Chinese university libraries

Yingshen Huang¹ | Andrew M. Cox²  | Laura Sbaffi² 

¹Health Science Library, Peking University, Beijing, China

²The Information School, University of Sheffield, Regent Court, Sheffield, UK

Correspondence

Andrew M. Cox, The Information School, University of Sheffield, Regent Court, 211 Portobello, Sheffield, S1 4DP, UK.
Email: a.m.cox@sheffield.ac.uk

Abstract

On April 2, 2018, the State Council of China formally released a national Research Data Management (RDM) policy “Measures for Managing Scientific Data”. In this context and given that university libraries have played an important role in supporting RDM at an institutional level in North America, Europe, and Australasia, the aim of this article is to explore the current status of RDM in Chinese universities, in particular how university libraries have been involved in taking the agenda forward. This article uses a mixed-methods data collection approach and draws on a website analysis of university policies and services; a questionnaire for university librarians; and semi-structured interviews. Findings indicate that Research Data Service at a local level in Chinese Universities are in their infancy. There is more evidence of activity in developing data repositories than support services. There is little development of local policy. Among the explanations of this may be the existence of a national-level infrastructure for some subject disciplines, the lack of professionalization of librarianship, and the relatively weak resonance of openness as an idea in the Chinese context.

1 | INTRODUCTION

On April 2, 2018, the State Council of China formally released a national Research Data Management (RDM) policy “Measures for Managing Scientific Data,” hereafter *Measures* (The State Council of China, 2018a, 2018b). *Measures* was the first attempt to define the responsibilities for RDM of administrative institutions such as the Ministry of Science and Technology and provincial technology departments, as well as of individual research institutions and research data centers. The policy states that local institutions should establish their own policy and create Research Data Services (RDS).

There has been work around RDM before in China. Since its foundation in 1984, the Chinese Committee on

Data (CN-CODATA) has made efforts to improve China’s capability in data curation and begun to promote sharing of research data. In 2001, the “Meteorological Data Sharing Management Regulation” was issued, which was the first data management policy in China focusing on data sharing (China Meteorological Administration, 2008). Nevertheless, considering China’s importance to global scientific production, *Measures* comes relatively late compared to developments in RDM policy in North America, Europe, and Australasia.

In the context of this major government initiative, this article seeks to examine the changing status of RDM in China, by addressing the following questions:

1. What is the level of RDM policy development in Chinese universities?

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2020 The Authors. *Journal of the Association for Information Science and Technology* published by Wiley Periodicals LLC on behalf of Association for Information Science and Technology.

2. What is the level of RDS development in Chinese academic libraries?
3. What are perceived to be the main drivers and challenges for developing RDS in Chinese academic libraries?
4. How are Chinese academic libraries engaging with the agenda around RDS?

The study is based on analysis of websites, a survey, and interviews. Data from previous work on libraries' international development of RDS conducted in 2014 (Cox, Kennan, Lyon, & Pinfield, 2017) and repeated in 2018 (Cox, Kennan, Lyon, Pinfield, & Sbaffi, 2019) are used to provide comparative context.

2 | LITERATURE REVIEW

The last 15 years have seen an increasing international recognition of the importance of RDM in government and funder policies, triggered by the Organization for Economic Co-operation and Development *Principles and guidelines for access to research data from public funding* (OECD, 2007). In response to this emerging policy framework, institutional services have advanced at the local level (Tenopir, Pollock, Allard, & Hughes, 2016; Tenopir et al., 2017; Cox et al., 2017; Cox et al., 2019). As an important stakeholder, libraries are closely involved in RDM in the United Kingdom and North America (Pryor, Jones & Whyte, 2014), and they provide advisory and technical RDS, which together define a model of mature RDS (Cox et al., 2019).

Measures reflects this trend, but there has been previous work around RDM in China. In addition to the early initiatives mentioned in the introduction to this article, there have been policies specific to certain natural sciences that collect big scale data via observation and measurement, with an emphasis on data sharing, submission, and long-term preservation (China Meteorological Administration, 2008; MOST, Ministry of Science and Technology, 2004). The Chinese Academy of Sciences, the national research institution, has built a scientific data cloud store to deposit research data (Li, Yu, Zhang, Liu, & Wu, 2015). There has also been activity at university level, with some institutions creating data platforms for sharing and reuse (Liu & Rao, 2013; Zhang, Yin, Zhang, Guo, & Zhang, 2015; Luo, Zhu, Cui, & Nie, 2016). In 2014, the China Academic Library Research Data Management Implementation Group was established by a number of high-ranking university libraries with the purpose of promoting the development of RDM (Yin & Wang, 2014). However, as in other

countries, there have been significant barriers to developing RDS, such as lack of policy norms, inadequate technical support, and skill gaps (Zhou, Duan, & Song, 2017).

There are some differences between *Measures* and national policies elsewhere. Most EU policies, for example, are advisory, but the Chinese one, as an executive and governmental order issued by the highest research management department of China, is compulsory and mandatory (SPARC Europe, 2017; UKRI, 2016). This might be perceived to be very positive for ensuring that RDM is now a required function of universities. Nevertheless, it is not straightforward to translate *Measures* into practice. As a national guideline, *Measures* sets out the responsibilities of institutions at various levels, but it does not include how to address these responsibilities. Furthermore, in setting out differing responsibilities, *Measures* only names high-level stakeholders, such as national and provincial bodies, research institutes that generate and manage data, and data centers that focus on data curation. *Measures* does not mention nor define the role of other stakeholders, such as researchers, funding organizations, publishers, and data professionals (Erway, 2013). The definition of research data used in *Measures* is also somewhat ambiguous, as it states that "research data is raw and derived data that is generated in the Natural Science and Engineering Technology Science area, and collected using observation, monitoring, inspection, survey, test and detection and intended for use for research activities" (The State Council of China, 2018a). The definition emphasizes natural and engineering science data and does not mention social research data specifically.

3 | METHODOLOGY

This research involved the sequential use of three forms of data collection: website analysis of university policies and services, an online questionnaire, and semi-structured interviews. One year after the release of the *Measures*, a university website analysis was conducted to discover the level of policy implementation at university level and the related services provided by university libraries. However, the results showed that libraries very rarely state on their websites whether they provide RDS, so a questionnaire survey for librarians was launched to investigate to what extent Chinese university libraries have actually been involved in RDM. Semi-structured interviews with librarians were undertaken as a final step in the research to develop a deeper understanding of their perspectives on RDM.

The website analysis sought to identify institutional RDM activities as evidenced in news, blogs, social media posts, users guides, instructions, or introductions, with data being collected from January 1, 2019 to April 1, 2019. A Google search limited by “site:” was used for detecting whether certain phrases such as *Research Data* or *RDM* or *Scientific Data* or *Data Management* were contained in universities' domains. Then the search results were refined and reviewed manually, to assess: (a) whether there were rules, regulations, and plans related to research data; (b) whether there were research data-related services, and if any, which departments are involved in running them. The website analysis was conducted on 151 universities, including the 137 Double First-Class universities, which are approved by China Ministry of Education as key universities (MOE, 2017b), 11 universities in Hong Kong eligible for doctoral degrees, and 3 universities in Macao.

The follow-up questionnaire was sent to the same list of universities in order to confirm the website analysis results and explore the real situation in libraries. An international questionnaire conducted and reported by Cox et al. (2019) was adapted into a Chinese language version, with the main questions retained to enable comparison. The revised questionnaire, composed of 23 questions, was piloted with five librarians from mainland China and then distributed to the target university libraries via invitation emails sent between June and November 2019. Because mainland Chinese academic library staff contact details are not always published, some invitations were sent to the library's public mailbox, and some libraries have no email address published on their websites; thus, the final number of libraries reached by the invitation email was 122.

Semi-structured interviews were also conducted with librarians in order to understand the drivers and challenges for RDS. The same list of 122 libraries was used to send an interview invitation; this included institutions that had not completed the questionnaire. The interview questions focused on what RDS the library currently provides or plans to develop in the future and what are the main drivers and challenges when delivering RDS. The interviews sought to identify the underlying reasons why the library appears to be less involved in RDM in China. Ten interviews were conducted between September and December 2019, nine in Chinese, one in English; three were face-to-face interviews, and the remaining were remote audio calls. All the recordings were transcribed and translated into English, then coded using NVivo 12. Five interviewees were directors of libraries, three were senior managers, and the other two were librarians responsible for RDS directly. The mean length of the interviews was 42 min.

4 | FINDINGS

4.1 | Website analysis

At the time the website analysis was conducted, only one university—Hong Kong University—had a policy for *Research Data and Records Management*. It was an adapted version of the 2012 Oxford University Policy and was found under the Research Integrity section of their website (The University of Hong Kong, 2015). Some universities had posted or forwarded the national *Measures*; however, none of the other universities in the sample appeared to have a policy in place.

Although there was no formal policy in any of the other 150 universities, nine university libraries, five of which were in Hong Kong, had moved forward to provide RDS (Table 1). Most of the services were advisory in nature, seeking to introduce the benefits of RDM and how to achieve good practice. In addition, four libraries had their own data repository, which were open to the public. One repository was for social science data and the other three were comprehensive ones containing data from both social science and natural science. The Peking University Open Research Data Platform (2016) has a user guide document explaining exactly the functions of the platform, how to use the data or data sets, how to share data, and providing the definitions of key concepts such as, Research data, Data file, Dataverse, and Dataset. There were also four academic libraries, which offered practitioner conferences and workshops about RDM.

Overall the results indicated quite a low level of development of RDS in China at an institutional level, but it was recognized that not all activity is necessarily publicized on the open web. To examine the issue further, the questionnaire was developed with questions to obtain more details of current RDS, strategic priorities, and perceptions of crucial skills and factors shaping RDS.

4.2 | Questionnaire respondents

By the end of November 2019, 42 out of 122 invited libraries (108 in mainland China, 11 in Hong Kong, and 3 in Macao) had completed the questionnaire (35 from mainland and 7 from Hong Kong), corresponding to a response rate of 34%. In interpreting the results, this relatively low response rate should be taken into consideration. It is likely that libraries with more activity or interest in RDM would respond and so the data probably exaggerate the level of engagement with the subject. The data from the questionnaire were analyzed through descriptive statistics and comparative analysis. Twenty-four respondents were from universities located in

TABLE 1 Research data services from website analysis

University	Knowledge of RDM	DMP support	Open data resources	Data publish	Data citation	Data repository	Data analysis/mining/visualization	Training
Beijing Institute of Technology	No	No	No	No	No	No	No	Yes
Fudan University	No	No	Yes	No	No	Yes, Dataverse	No	No
Hong Kong Baptist University	Yes	Yes	No	No	Yes	No	Yes	Yes
Peking University	Yes	No	No	No	No	Yes, Dataverse	Yes	No
The Chinese University of Hong Kong	Yes	Yes	Yes	No	Yes	No	No	No
The Hong Kong Polytechnic University	Yes	Yes	No	No	No	No	No	No
The Hong Kong University of Science and Technology	Yes	Yes	No	No	No	No	No	Yes
The University of Hong Kong	Yes	Yes	Yes	Yes	No	Yes, plug in the institutional repository	No	Yes
Wuhan University	No	No	Yes	No	No	Yes, DSpace	No	No

Beijing (17 libraries) and Hong Kong (7 libraries). As intended by the method of circulation of the survey, 32 (75%) participants were part of the senior management team of the library and around 21 (50%) were from library directors, who are likely to be responsible for the overall future planning of their library at strategic level and understand the priorities for university development.

4.3 | RDM policy and RDS

The survey asked questions about the RDM policy in universities and which departments were involved in developing the policy. Four respondents (10%) stated that their institution has an RDM policy even though there were no formal policies, rules, or guidelines to be found on the university website (except for Hong Kong University). A further 10 institutions (24%) planned to have a policy, but 17 (41%) had no plan. There were 18 respondents (43%) that claimed that their institutions provide some research data-related services and 11 libraries (26%) planned to provide them. In providing RDS, about 18 libraries (66%) cooperate or plan to cooperate with external organizations and use commercial products to deliver RDS.

The questionnaire sought to study the role of libraries in RDM policy development and providing RDS. As shown in Figure 1, libraries are highly involved in both policy development and providing service, especially RDS, then followed by the Research Management Office and IT services. But the pattern in lead roles is different,

libraries are usually taking the leading role for services while research management offices lead on policy making.

The questionnaire also investigated RDS development through a matrix of choices on a wide range of services that could be offered by libraries. Figure 2 shows the current development for each service type. The blue line is always below the 1-point horizontal line, meaning that the mean level of all these services provided by libraries is not reaching the basic requirement, and advisory services are more or less provided at the same level as technical services. Chinese service level (mean = 0.55) is lower than that of western countries in 2018 (mean = 0.78) and 2014 (mean = 0.58) (Cox et al., 2019). The overall trends of the line are of similar except for *Data analysis, mining, and visualization advice services*, which are more developed. *Run a data repository* is the most developed service in China and almost reach the 1-point horizontal line, meaning that almost all respondents considered that they had reached the *Basic service* level in this respect.

The strategic priority given to RDS was evaluated via the same matrix; Figure 3 shows how libraries typically concentrate on certain types of service, and the mean priority score is 0.51 in China, 1.07 in the previous international survey in 2018, and 1.24 in 2014. The overall priority is low in China which may indicate that they are inactive in exploring new RDS. The highest priority in China is *Run a data repository*, which is a technical service. *Advice on DMPs* which was among the highest priorities in western countries seems to be a much lower concern in China.

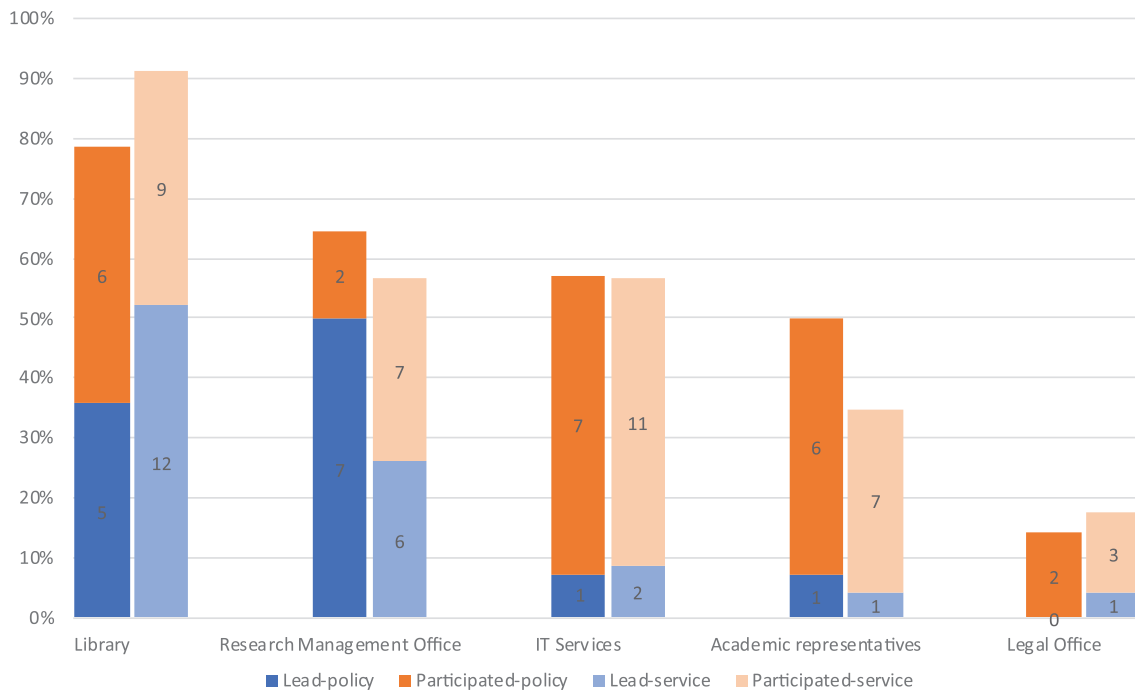


FIGURE 1 Departments involved in development of RDM policy ($n_{policy} = 14$) and RDS ($n_{service} = 23$) [Color figure can be viewed at wileyonlinelibrary.com]

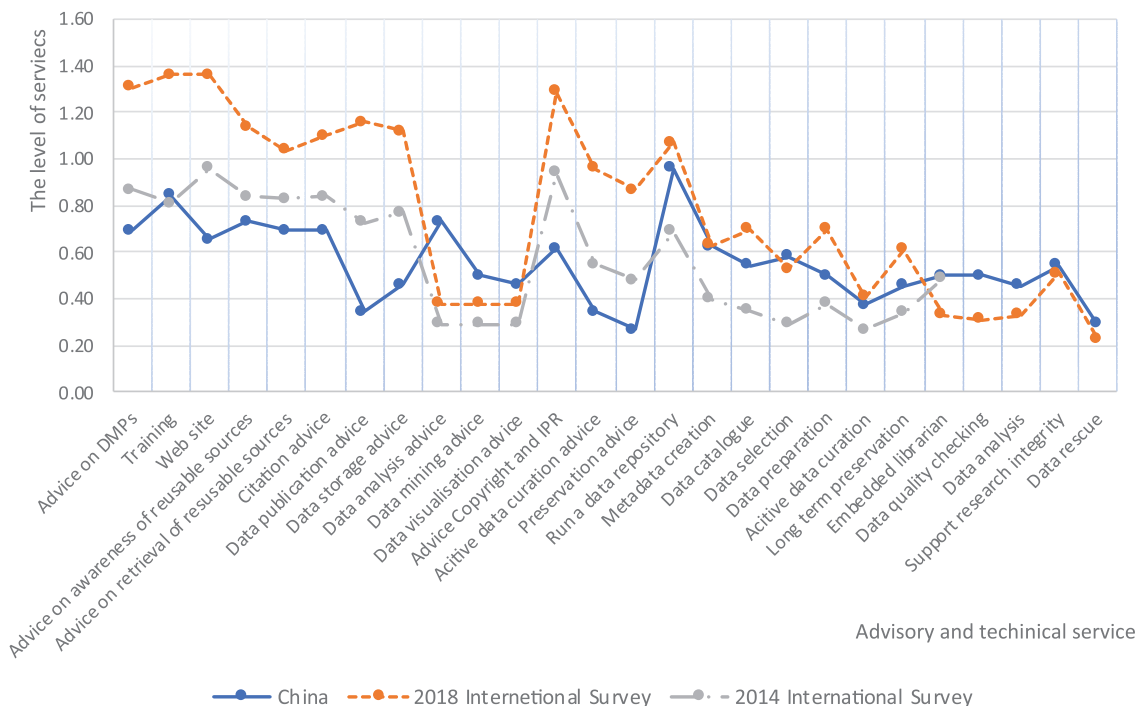


FIGURE 2 Advisory and technical services in China compared with previous surveys [Color figure can be viewed at wileyonlinelibrary.com]

4.4 | Librarians' responsibilities and skills

Within the libraries that have RDS, or plan to, 13 respondents (33%) distributed RDS tasks to a specific new RDM

team and others handing this task to individual or spread across multiple teams. The question that asked about the skills needed for providing RDS was a multiple-choice one, hence the total percentage exceeds 100%, the mean percentage in China (81% of 35 respondents in current

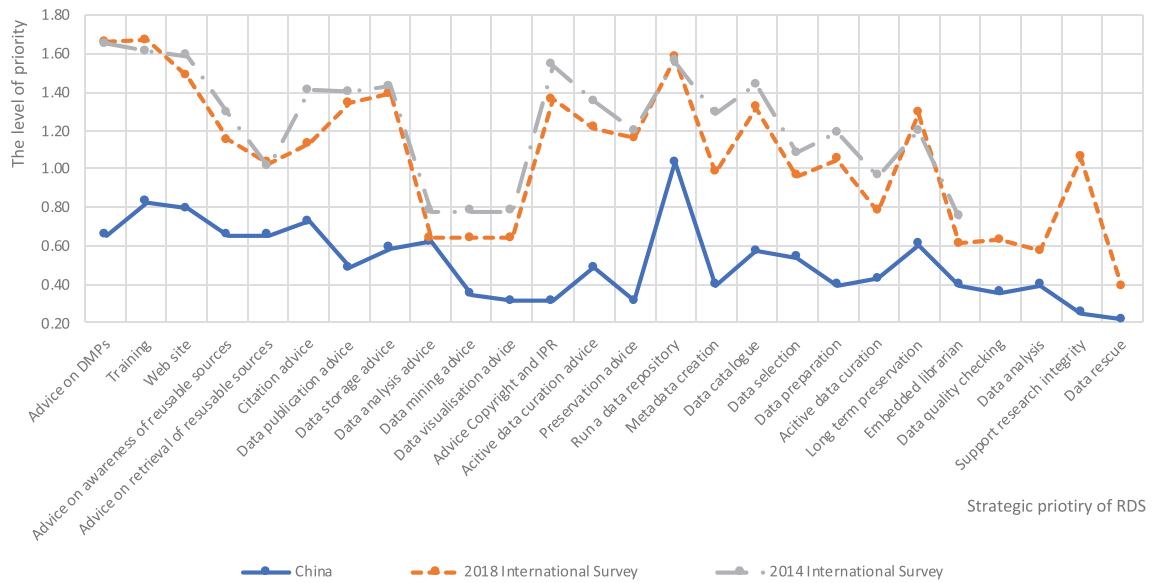


FIGURE 3 Strategic priorities for RDS in China compared with previous surveys [Color figure can be viewed at wileyonlinelibrary.com]

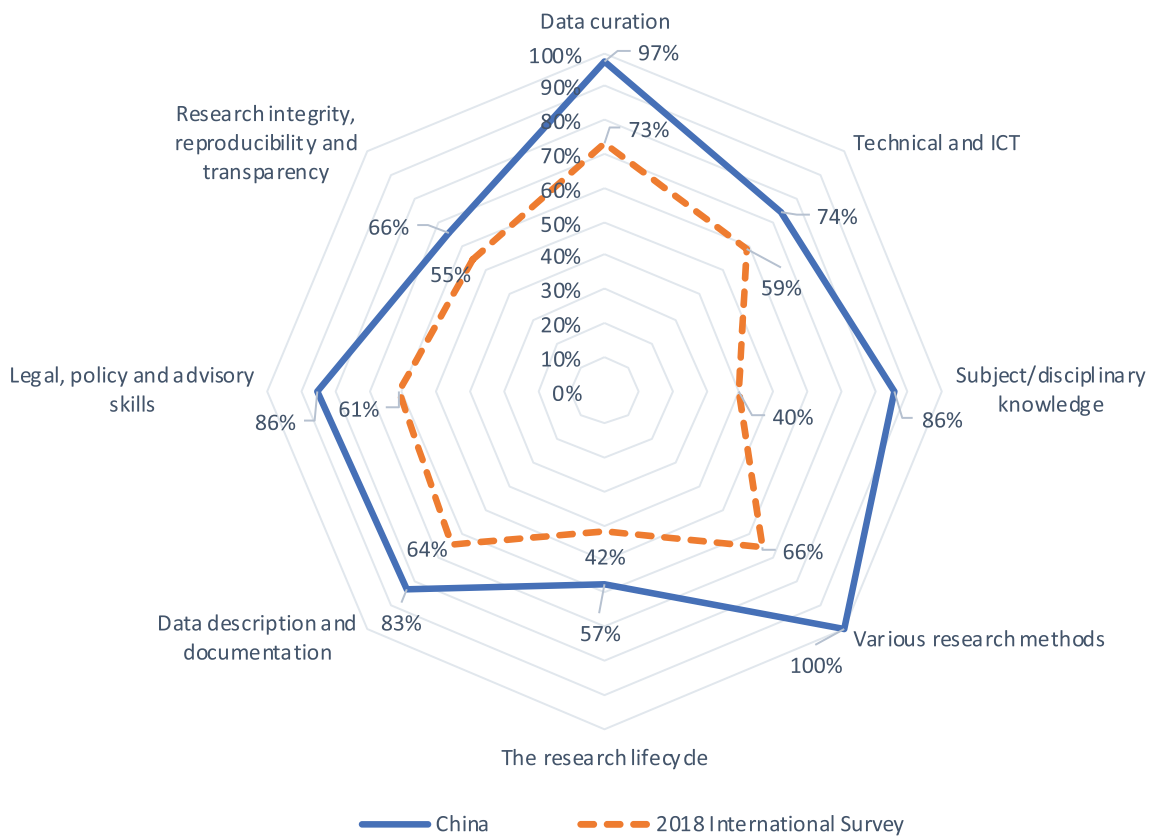


FIGURE 4 Additional knowledge or skills need development ($n_{China} = 35, n_{2018} = 209$) [Color figure can be viewed at wileyonlinelibrary.com]

survey) is much higher than western countries (58% of 209 respondents in the 2018 survey) (Cox et al., 2019), indicating a perception of a deeper lack of knowledge and skills. All 35 respondents thought that knowledge of

research methods (e.g., data analysis or data visualization) is necessary (Figure 4). The second highest skill needed was Data curation. Needs for subject or discipline knowledge differed a lot between the two surveys,

perhaps because librarians with a subject background in China are thought to have higher professional skills than just a general librarian and have deeper understanding of the research of their subjects/majors.

4.5 | Drivers and challenges

The data from both open-ended questions in the questionnaire and the interviews sought to understand the factors that will promote or hinder the library to engage in RDM. In order to categorize the factors, this study used the categories developed in the previous survey conducted by Cox et al. (2019), which was based purely on questionnaires. Sometimes it was hard to differentiate between drivers and challenges when dealing with complicated factors. For example, the respondents sometimes saw the opportunity to improve specific skills as a driver while lack of the same skills was also perceived as a barrier. Therefore, the study included every factor even if they had been mentioned as both drivers and challenges. Table 2 shows the drivers identified in the questionnaire and in the interviews, with the total representing the number of respondents who mentioned each factor.

The most common driver mentioned both in the questionnaire and interviews emphasizes the library's role. All interviewees mentioned the knowledge, skills, and capability of librarians to provide related RDS as a key factor. Six of them thought their library needs to improve but does not have the ability to provide RDS, including the simplest introductory services. Three of them have stepped forward and already provided pilot technical services for 2–3 years. Another side of this first driver is a professional identity crisis, as librarians themselves doubt their own value. Some key university libraries have begun to focus on supporting research, so RDS is a new chance for them to strengthen their presence and the value of their service.

From the perspective of the library, we have the motivation to progress RDM. The library has a sense of crisis. We buy fewer books, there is no more space to store them, and students don't read them. We don't own the electronic resources either. Where is the value of library's survival? I think research data is the next opportunity for the library to seize, because data is also a kind of resource. Other departments within the university

TABLE 2 Major drivers for RDS

	Codes	Questionnaire	Interviews	Total	Percentage (%)
1	Library role—having the skills/need to stay relevant	6	9	15	23
2	Awareness of the value of research data and the benefits of RDM	8	2	10	16
3	Needs of researchers	3	5	8	13
4	Funders' requirements	3	3	6	9
5	Institutional policy	1	4	5	8
6	Needs of university administrative department (research manage office, principal office)	3	1	4	6
7	Openness of research data	2		2	3
8	Reputation of university	2		2	3
9	Publishers	1	1	2	3
10	Sharing of research data	1	1	2	3
11	Drivers-publish and citation		1	1	2
12	Researchers' attitudes	1		1	2
13	Integrity	1		1	2
14	Institutional repository	1		1	2
15	Government's mandatory requirements	1		1	2
16	Drivers-academic rewards system		1	1	2
17	Drivers-data security		1	1	2
18	Drivers-FAIR		1	1	2
	Total	34	30	64	

may be unwilling and unable to carry out this work. This is not a job but an opportunity for the library. [Interview 3]

The second most common driver was the rising awareness of benefits of RDM and value of institutional research data to the university's leadership, library's management, and researchers. Top-down requirements or policies are thought to be more effective than bottom-up ones.

The increased awareness by the university of the importance and value of research data. [Questionnaire]

Whether the library undertakes RDS or not depends on the awareness of university leadership and library director. [Interview 8]

The University's attitude and policies towards RDM. [Questionnaire]

The third major driver was researcher need.

Many departments and research teams start their own management after they get funding although there is no policy ask them to do so. The needs of how to manage the data efficiently will potentially promote the development of research data services and supports. [Interview 8]

Only the first common driver relates to libraries, the other 17 coded drivers concern four other main stakeholders, specifically funders, researchers, publishers, and the research community.

Table 3 shows the challenges identified in the questionnaire and interviews, and the total represents the number of respondents mentioning each factor.

When developing RDS, skills or knowledge is mentioned most frequently as a challenge, although it was also the most common driver.

Skills, skills, skills. We just can't find people with the right ones. [Questionnaire]

NO relevant knowledge. [Questionnaire]

Challenges: knowhow and skills of library staff. [Questionnaire]

The major barrier in questionnaire comments is "skills or knowledge," which had also been mentioned the same

number of times as "lack mandate/rewards" and "Acceptance of RDM role" in the interviews. The rest of the challenges appear to be of similar importance and are closely related to the awareness of RDM. High awareness of their own research data heritage in universities can promote the development of policy that guides researchers to take care of the data they produce. From the perspective of librarians, researchers' reluctance to share data is a great challenge.

The difficulty is that the university does not pay attention to nor value the management of research data, and researchers are not willing to open their own research data. [Interview 8]

Barriers... (3) University's support in terms of budget, human resources and a policy that makes RDM a requirement; (4) Faculty's willingness to share data. [Questionnaire]

The sharing of research data, sensitive data. [Questionnaire]

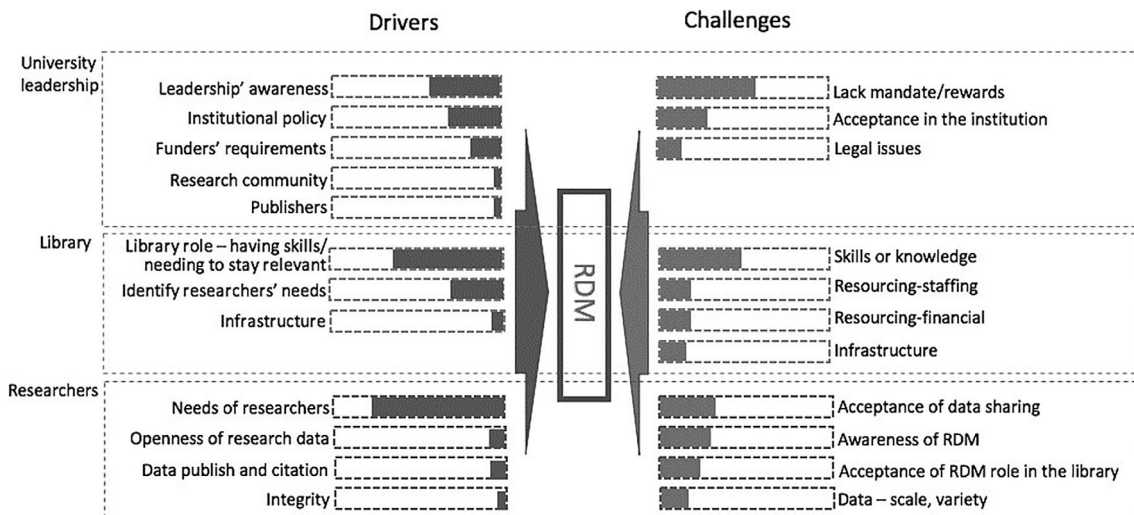
To summarize the drivers and challenges data, a force field analysis visualization was produced. As shown in Figure 5, three main RDM stakeholders determine drivers and challenges for RDS in Chinese universities. University leadership can do more, in terms of developing the policy, arranging RDM activity, conducting audits of the university's data assets or, simply, approving budget and assigning this task to an appropriate organizational unit. The library is a neutral force: it is not the producer nor the user of research data, it can support RDM not only for researchers, but it can help other stakeholders better understand RDM and it can provide different levels of RDS. Whether a library is motivated to explore or is resisting RDM depends on its own quality, technical level, infrastructure, and financial budget. Researchers have the most crucial role in RDM, as they produce, manage, use, share, and publish data. They have their own unique needs, and they must already be managing their own data, so centralized RDM is perceived as an additional job for them, especially when they cannot see any expected benefits for adopting good practice.

4.6 | A "wait-and-see" attitude

The posture of libraries toward RDM from the interview data can be summarized as *wait-and-see*, which encompasses two different approaches. One is a positive *wait-and-see* attitude based on good preparation toward RDM,

TABLE 3 Major challenges for RDM services

	Code	Questionnaire	Interviews	Total	Percentage (%)
1	Skills or knowledge	10	4	14	19
2	Lack mandate/rewards	4	4	8	11
3	Acceptance of data sharing	6	2	8	11
4	Lack institutional policy	4	3	7	9
5	Acceptance of RDM role in the library	1	4	5	7
6	Resourcing—staffing	3	2	5	7
7	Data—scale, variety	1	3	4	5
8	Resourcing—financial	2	2	4	5
9	Awareness on RDM of researchers	2	2	4	5
10	Acceptance in the institution	4		4	5
11	Engagement of academic staff	3		3	4
12	Infrastructure	3		3	4
13	Preservation	1	1	2	3
14	Legal issues	2		2	3
15	Acceptance of the need for RDM	1		1	1
	Total	47	27	74	

**FIGURE 5** Force field analysis of drivers and challenges

which includes high awareness and activities such as theoretical research, literature reviews, case study development, training abroad, and implementing pilot services (Wuhan University Library, 2012).

We started data management work at 2016. Fudan University started earlier, maybe in 2014, and their Director has a team focus on RDM. They went to the United States, visited some universities, research institutes, and data management institutions. At present, in China, our data management work is still

relatively advanced, we are familiar with data management, foreign and domestic RDM situations. [Interview 8]

The service we have been doing is to make data sets discoverable. From 2014, no matter where the research data is generated (our university or other institutions), as long as the data sets are registered in the data repository, we guarantee that the data sets can be found and accessible by our researchers. [Interview 6]

Some library directors are well informed about RDM developments in western countries, so they send staff abroad to investigate how other libraries carry out RDS, what main technologies are needed, and how to operate them. After returning to China, they began to try to find a pathway or *entry point* to promote RDM in their universities, cooperating with some professors and departments, and providing appropriate pilot services:

I am the first person funded in our library for exploring new RDS. I went to Purdue University to participate in the PURR project because our library is going to carry out research data services. [interview 6]

However, although they have provided relevant RDS, such as an integrated open data repository in the library's resource catalogue, metadata creation support, data preservation, and data citation support, and so on, they did not advertise these new services, and had no plan to explore researchers' needs. They thought they have the right preparations to create RDS but seemed to be waiting for the creation of policy at the institutional level and official authorization from the university leadership:

There are two necessary conditions that good RDM needs, one is the awareness of university leaders in terms of developing policy, assigning the work to the appropriate institutional services. The other one is that at least one department in the university should have the knowledge and skills to provide RDM support. These two conditions are indispensable. And I think we have been well prepared with RDM related knowledge, skills, abilities and infrastructures. [Interview 8]

The other type of wait-and-see attitude revealed in our interview data is passive, without preparation and less awareness of RDM, to the point of being confused about the differences between the management of research data and publications. One interviewee, a library vice director, admitted that the interview was the first time she had heard about the national RDM policy and was just starting to understand the concept of RDM.

Another interviewee, also a library vice director, explained that the reason why they had not responded to the questionnaire was that they did not understand the questions. This might represent the views of many other libraries:

Sorry we didn't respond to the questionnaire. I did receive the invitation and took a look at

all the questions, some of which are a bit hard to understand, so, since we have not provided any related services, we decided not to answer the questionnaire. But it is nice to talk with you so that now I understand more about RDM. [Interview 2]

Most of the libraries are the second type of wait-and-see with no preparations in place and simply waiting to see what others do.

We currently have a limited budget and may not be able to carry out such work. We would wait and see for a while, and see how some large and well-resourced libraries conduct RDM, and then we can learn from their experience. [Interview 1]

5 | DISCUSSION

Reflecting on the findings, it is apparent that RDM in Chinese universities remains in its infancy. As evidence of this, firstly, only one university has a publically accessible RDM policy, which has been adapted from the Oxford University policy, and not even revised since the Chinese national policy was issued (The University of Hong Kong, 2015); three other universities have policies in place but they are not openly online. In comparison, as early as 2016, 80 universities in the United Kingdom already had an institutional RDM policy (Horton, 2016), and a significant number of libraries and their institutions in Australia and the Netherlands have research data policy in place or to be implemented within a year (Cox et al., 2017). Secondly, compared to the previous study by Cox et al. (2019), the level of services offered by Chinese institutions was lower in every instance. The low response rate to the survey was also suggestive of a lack of awareness of RDM, as confirmed by comments in the interviews, where library directors often showed little engagement with the topic. This is despite the fact that our target participants were in the best-funded institutions in China, the Double First-Class Universities. One can infer that less well-funded institutions among the 3,000 higher education institutions in China (MOE, Ministry of Education, 2017a, 2017b) would have yielded a picture of even less awareness and activity. Some of the key drivers and barriers are rather familiar from other contexts (Cox et al., 2019), for example, library role, researchers' needs and funder's requirement as drivers, lack of skills as barrier, but the picture in China seems to be still in a posture of "wait and see." We suggest that there may be three main reasons for the low development

of RDM services in academic libraries in China. The first reason is that some data-intensive disciplines already have their own data management infrastructure. Disciplines such as meteorology, geography, population health, and earth science already have a place to deposit and share their data through the National Science & Technology Infrastructure (NSTI, National Science and Technology Infrastructure, 2019). This is a national project hosted by the Ministry of Science and Technology (MOST, Ministry of Science and Technology, 2003) launched in 2002 and which passed its final evaluation in 2013. From 2010 to 2015, Higher Education Institutions (HEI) in China have led more than 80% of National Natural Science Foundation of China (NSFC) projects (MOE, Ministry of Education, 2016), which is the main funding body in China and is administrated by MOST (NSFC, National Natural Science Foundation of China, 2020). Research teams applying for funding from NSFC, both from university and non-university research institutions, have to follow the requirement to deposit and share data. The need for this in such disciplines reflects recognition of the big scale of data collected via large and expensive instrumentation like telescopes that are financed by the state. There is also a national-level priority for the creation of a data service network with various discipline data centers, which is intended to conduct the integration, sorting, classification, mining, and curation of data submitted from national projects and to promote open data sharing (Yuan, 2018). Currently, 20 national data centers and 30 national biological germplasm and experimental material resource banks have been approved to strengthen the construction and implementation of a scientific resources sharing system in order to promote the sharing and opened these resources to the public (MOST, Ministry of Science and Technology, 2019).

Meanwhile, an interviewee in this study who led the library's RDS team, confirmed that some research teams in departments have deposited or shared their data in national discipline data centers and popular data repositories, such as *github* and the local RDS team were seeking to integrate such open source data repositories into the library catalogue, linking data sets with articles. Therefore, in addition to the state-run repositories, some researchers are using open-source databanks to share data. Since RDM needs have already been met in some areas, local institutional level policies and RDS are not always needed.

The second reason why library run RDS have developed less strongly may be due to libraries in China being less influential institutionally than their western counterparts. Unlike in western countries, where libraries have played an important role in leading RDS

(Cox et al., 2019), Chinese university libraries seem passive and are reticent to engage with RDM, citing a lack of knowledge, skills, personnel abilities, human resources, and self-confidence. These are familiar complaints in the western context too. However, libraries in China have traditionally had less ability to influence institutional policies, which in China is typically a top-down process. Librarianship is not professionalized in China, with there being no general or national requirement to have certification to become a librarian. Both the lack of status and the lack of knowledge sharing networks in China may have inhibited developments around RDM. In western countries, knowledge sharing among librarians has played an important part in sustaining momentum for developing services around RDM, but the same level of professional knowledge sharing does not seem to exist in China.

Another aspect of the status of librarians could be that one of the important roles of Chinese academic librarians is to evaluate the novelty of research proposals for certain kinds of funding (MOE, Ministry of Education, 2007). This places them in a relation of evaluation over researchers and perhaps makes them less service-oriented than western librarians and even viewed with some suspicion by researchers. Such existing relationships between researchers and librarians may play a negative role in creating acceptance for new library-based RDS for researchers.

The third reason for the slow creation of RDS in China could be that one of the key policy drivers for RDM in western countries is the principle of "openness," which has a strong positive cultural resonance. Access to information about anything is seen as part of the rights of every citizen and a value that libraries in the west identify themselves with. The word "openness" simply does not have the same resonance in China, where information is usually open with limitations or available only to specific people in a particular area of study. The public and government have begun to recognize the importance of information disclosure (Huang & Wen, 2017), but it has been difficult to change awareness of the value of open research. *Measures* mentions that "research data produced from government budget funding should be stored on the principle of openness as standard and non-openness as an exception" (The State Council of China, 2018a), but in practice, most institutions set strict restrictions for access to their research data through delayed release, technical control, identity control, and so on (Huang, Wang, & Zhou, 2014). It is important to acknowledge that there is a large body of evidence demonstrating that openness is not always simply good and it may be much more to the advantage of scholars in

western countries than to researchers in less advantaged contexts (Bezuidenhout, Leonelli, Kelly, & Rappert, 2017). For researchers, who agree in principle but are reluctant in practice to share or open their data (Ward, Freiman, Molloy, Jones, & Snow, 2010), the newly released Measures has no substantial impact on their daily data management activities, and they do not need to make further changes.

Although there were almost no RDM policies in Chinese universities when the survey was completed, there has been some exploration and services of RDM at university library level since 2009. In the last decade, some libraries have actively studied RDM and tried to engage in it without policy, have directly gone to the technical-level services; supported data catalogues, citation, and metadata, and provided data repository, analysis, and mining services. Where libraries have developed RDS, they seem to have started with technology rather than policy and this is in contrast to the pattern of development suggested by Cox et al.'s (2019) maturity model. One possible reason is that technologies are relatively easy to learn and can be "copied" from country to country. In contrast, policies are harder to transplant directly because they are so context dependent. Libraries find it relatively hard to influence policies. The data platforms have been easy to construct but their sustainability, usage rate, and acceptance among researchers are still developing due to the lack of policy support and mandate requirement from funders.

6 | CONCLUSION

This study is the first systematic attempt to capture the development of RDM in academic libraries in China. Using data from web sites analysis, online questionnaires and interviews, a clear picture of the level of RDS in the more prestigious universities in the country has been formed and provides insights into RDM development's key drivers and challenges. It is apparent that RDM in academic libraries remains in its infancy in the country, with librarians adopting a "wait and see" approach, usually in the sense of simply waiting to follow the rest of the sector. Explanations for this relatively low level of development reflect the existence of alternative infrastructures for data sharing in some disciplines, the low professionalization of librarianship and the weaker cultural resonance of openness as a driver.

This study offers a starting point for stakeholders to identify different strategic priorities and paths when academic libraries begin to develop new RDS, and find more effective ways for libraries to promote RDM in Chinese

universities. Future work will focus on how libraries improve their impact on institutional RDM policy based on closely cooperating with research management offices and university IT departments, and how libraries can dispel researchers' doubts and deliver their new RDS successfully and broadly.

On a practical level, researchers' needs act as an important driver for RDM, which implies that academic libraries should pay more attention to these before implementing new services. Studies of user needs was notably absent in Chinese libraries work around RDM. Understanding needs first instead of implementing the new services without a needs analysis will help to avoid the low usage rate of services. Libraries may well find it useful not just to focus on technical services but promote advisory services because they can gradually build up understanding of the concept of RDM in the research community.

ORCID

Andrew M. Cox  <https://orcid.org/0000-0002-2587-245X>

Laura Sbaffi  <https://orcid.org/0000-0003-4920-893X>

REFERENCES

- Bezuidenhout, L. M., Leonelli, S., Kelly, A. H., & Rappert, B. (2017). Beyond the digital divide: Towards a situated approach to open data. *Science and Public Policy*, 44(4), 464–475.
- China Meteorological Administration 2008. Zhongguo qixiangju di 4 hao ling "qixiang ziliao gongxiang guanli banfa" [China Meteorological Administration Order No. 4 "Meteorological Data Sharing Management Regulation"]. Retrieved from http://www.cma.gov.cn/2011zwxx/2011zflfg/2011zbgmgz/201110/t20111027_135170.html.
- Cox, A. M., Kennan, M. A., Lyon, L., & Pinfield, S. (2017). Developments in research data management in academic libraries: Towards an understanding of research data service maturity. *Journal of the Association for Information Science and Technology*, 68(9), 2182–2200.
- Cox, A., Kennan, M., Lyon, L., Pinfield, S., & Sbaffi, L. (2019). Maturing research data services and the transformation of academic libraries. *Journal of Documentation*, 75(6), 1432–1462. <https://doi.org/10.1108/JD-12-2018-0211>
- Erway, R. (2013). *Starting the Conversation: University-wide Research Data Management Policy*. Retrieved from <https://www.oclc.org/research/publications/library/2013/2013-08r.html>
- Horton, L. (2016). Overview of UK Institution RDM Policies, Version 6 August 2016, Digital Curation Centre. Retrieved from <http://www.dcc.ac.uk/resources/policy-and-legal/institutional-data-policies>.
- Huang, R., Wang, B., & Zhou, Z. (2014). Cujin woguo kexue shuju gongxiang de duice [a study on countermeasures for promoting scientific data sharing in China]. *Tushuguan*, 2014(3), 7–13.
- Huang, R., & Wen, F. (2017). Woguo zhengfu shuju kaifang gongxiang de zhengce kuangjia yu neirong: Guojia cengmian zhengce wenben de neirong fenxi [Policy framework and content of opening and sharing government data in China: A

- content analysis of policy documents at the national level]. *Tushu Qingbao Gongzuo*, 61(20), 12–25. <https://doi.org/10.13266/j.issn.0252-3116.2017.20.002>
- Li, J., Yu, L., Zhang, B., Liu, F., & Wu, Z. (2015). Zhongkeyuan kexue shuju yun jiegou tanxi [Analysis of the Chinese Academy of Sciences scientific data cloud architecture]. *Zhongguo Jiaoyu Wangluo*, 10, 33–34.
- Liu, X., & Rao, Y. (2013). Gaoxiao tushuguan kexue shuju guanli yu fuwu chutan—Wuhan daxue tushuguan anli fenxi [Scientific Data Management and Service in University Library—A case study of Wuhan University Library]. *Tushu Qingbao Gongzuo*, 57(6), 33–38. <https://doi.org/10.7536/j.issn.0252-3116.2013.06.006>
- Luo, P., Zhu, L., Cui, H., & Nie, H. (2016). Jiyu Dataverse de beijing daxue kaifang yanjiu shuju pingtai jianshe [The construction of Peking University Open Research Data Platform based on Dataverse]. *Tushu Qingbao Gongzuo*, 60(3), 52–58. <https://doi.org/10.13266/j.issn.0252-3116.2016.03.008>
- MOE, Ministry of Education. (2007). Jiaoyubu keji chaxin gongzuozhan [The Ministry of Education science and technology innovation check institutions]. Retrieved from http://www.moe.gov.cn/srcsite/zsdwxgk/200706/t20070601_60898.html.
- MOE, Ministry of Education. (2016). Gaodeng xuexiao “shisanwu” kexue he jishu fazhan guihua [“Thirteenth Five—Year plan” for science and technology development in higher education institutions]. Retrieved from http://www.moe.gov.cn/srcsite/A16/moe_784/201612/t20161219_292387.html.
- MOE, Ministry of Education. (2017a). Quanguo gaodeng xuexiao mingdan [The list of China high education institutions]. Retrieved from http://www.moe.gov.cn/srcsite/A03/moe_634/201706/t20170614_306900.html.
- MOE, Ministry of Education. (2017b). “Shuangyiliu” jianshe gaoxiao mingdan [The list of “Double-class” construction universities]. Retrieved from http://www.moe.edu.cn/s78/A22/A22_ztzt/tztl_tjsylpt/sylpt_jsgx/201712/t20171206_320667.html.
- MOST, Ministry of Science and Technology. (2003). Wo keji jichu tiaojian pingtai gongcheng qidong [The launch of the national science and technology infrastructure]. Retrieved from http://www.most.gov.cn/ztzl/qgkjgzh/2003/mtbdzl/200605/t20060509_32055.htm.
- MOST, Ministry of Science and Technology. (2004). 2004–2010 nian guojia keji jichu tiaojian pingtai jianshe gangyao de tongzhi [2004–2010 National Science and Technology Basic Conditions Platform Construction Outline]. Retrieved from http://zwgk.gd.gov.cn/006939748/200909/t20090915_9367.html.
- MOST, Ministry of Science and Technology. (June 5, 2019). Kejibu, caizhengbu guanyu fabu guojia keji ziyuan gongxiang fuwu pingtai youhua tiaozheng mingdan de tongzhi [Notice of the Ministry of Science and Technology and the Ministry of Finance on the release of the list of optimization and adjustment of the national science and technology resource sharing service platform]. Retrieved from http://www.most.gov.cn/mostinfo/xinxifenlei/fgzc/gfxwj/gfxwj2019/201906/t20190610_147031.htm.
- NSFC, National Natural Science Foundation of China. (2020). *Zhineng* [Responsibilities]. Retrieved from <http://www.nsf.gov.cn/publish/portal0/jgsz/02/>.
- NSTI, National Science and Technology Infrastructure. (2019). Guojia keji jichu tiaojian pingtai [National Science and Technology Infrastructure]. Retrieved from <http://www.nsdata.cn/collection/about>
- OECD. (2007). *OECD principles and guidelines for access to research data from public funding*. Paris: OECD Publishing. <https://doi.org/10.1787/9789264034020-en-fr>
- Peking University Open Research Data Platform (2016). *User guide*. Retrieved from <https://opendata.pku.edu.cn/handbook.xhtml>.
- Pryor, G., Jones, S., & Whyte, A. (Eds.). (2014). *Delivering research data management services: Fundamentals of good practice*. London: Facet.
- SPARC Europe. (May 31, 2017). *New SPARC Europe Report provides analysis of Open Data and Open Science Policies in Europe*. Retrieved from <https://sparceurope.org/new-sparc-europe-report-analyses-open-data-open-science-policies-europe/>.
- Tenopir, C., Pollock, D., Allard, S., & Hughes, D. (2016). Research data services in European and North American libraries: Current offerings and plans for the future. *Proceedings of the Association for Information Science and Technology*, 53(1), 1–6. <https://doi.org/10.1002/pr2.2016.14505301129>
- Tenopir, C., Talja, S., Horstmann, W., Late, E., Hughes, D., Pollock, D., ... Allard, S. (2017). Research data services in European academic research libraries. *LIBER Quarterly*, 27(1), 23–44. <http://doi.org/10.18352/lq.10180>
- The State Council of China. (2018a). *Kexue shuju guanli banfa* [Measures for managing scientific data]. Retrieved from http://www.gov.cn/zhengce/content/2018-04/02/content_5279272.htm.
- The State Council of China. (2018b). *Rules to enhance security, openness of scientific data*. Retrieved from http://english.gov.cn/policies/policy_watch/2018/04/05/content_281476101961402.htm.
- The University of Hong Kong. (2015). *Research data and records management*. Retrieved from <http://www.rss.hku.hk/integrity/research-data-records-management>.
- UKRI. (2016). Concordat on open research data. Retrieved from <https://www.ukri.org/funding/information-for-award-holders/data-policy/>.
- Ward, C., Freiman, L., Molloy, L., Jones, S., & Snow, K. (2010). Making sense: Talking data management with researchers. *International Journal of Digital Curation*, 6(2), 8–17.
- Wuhan University Library. (2012). Gaoxiao kexue shuju gongxiang pingtai [China Academic Scientific Data Service]. Retrieved from <http://sdm.lib.whu.edu.cn/jspui/>.
- Yin, S., & Wang, Y. (2014). “Zhongguo gaoxiao tushuguan yanjiu shuju guanli tuijin gongzuozu” zheng shi cheng li [“China Academic Library Research Data Management Implementation Group” was formally established]. *Shanghai Gaoxiao Tushu Qingbao Gongzuo Yanjiu*, 2014(4), 50.
- Yuan, Y. (2018). Rang kexue shuju kaifang gongxiang chengwei changtai [Make the open and sharing of scientific data become the norm]. *Guanming Daily*. Retrieved from http://epaper.gmw.cn/gmrb/html/2018-04/05/nw.D110000gmrb_20180405_7-02.htm
- Zhang, J., Yin, S., Zhang, Y., Guo, Y., & Zhang, Y. (2015). Shehui kexue shuju de gongxiang yu fuwu—yi fudan daxue shehui kexue shuju gongxiang pingtai wei li [Social scientific data sharing and serving—An example of Fudan University Social

Scientific Data Platform]. *Daxue Tushuguan Xuebao*, 33(1), 74–79.

Zhou, L., Duan, X., & Song, Y. (2017). Woguo gaoxiao tushuguan keyan shuju guanli fuwu diaocha yu fenxi [Investigation and analysis of research data management services in Chinese University Libraries]. *Tushu Qingbao Gongzuo*, 61(20), 77–86. <https://doi.org/10.13266/j.issn.0252-3116.2017.20.008>

How to cite this article: Huang Y, Cox AM, Sbaffi L. Research data management policy and practice in Chinese university libraries. *J Assoc Inf Sci Technol*. 2020;1–14. <https://doi.org/10.1002/asi.24413>