Global Visibility of Open Research Data Repositories: A Case Study of China

Yuan Ye*, Noorhidawati, A., and **Yanti Idaya Aspura, M.K**
Department of Library and Information Science,
Faculty of Arts and Social Sciences,
Universiti Malaya, Kuala Lumpur, MALAYSIA
e-mail: s2179270@siswa.um.edu.my; noorhidawati@um.edu.my; yanti@um.edu.my

ABSTRACT

The operation of an open research data repository is an important element in measuring the extent of open science. China is one of the most populous countries and contributes significantly to research output. This paper explores the current status of open research data repositories in China by analyzing the characteristics of the repositories and reflecting on the process of open science in China.The data were obtained from the re3data database (re3data.org), and the search was limited to China. The data were further analyzed and tabulated according to the following criteria: subject, type of content, keywords, language, software, type of repository, institution responsibility, institution type, and type of provider.The study shows that China ranks 9th in the world in terms of the number of repositories, falling short of expectations. Of the 81 open research data repositories surveyed, 79 are nonprofit organizations, with a focus on life sciences, natural sciences, and medicine as the main subject repositories. The diversity of content types and the universality of the language interface also reflect the richness of the repository. However, the state of use of the software is promising and most of it is "unknown"; the mode of delivery is dominated by the provision of data and a lack of services.This paper has examined OA data repositories in China and provides an important reference for understanding open science initiatives in China by highlighting the current status of open data repository development.

Keywords: Open access, China, Open Research Data Repositories, Re3data

INTRODUCTION

The advent of digital technology and ICT has revolutionized the entire academic environment, increasing the visibility and availability of scholarly output on a wide scale (Singh, 2016). The rise in subscription and licensing fees and increased journal requirements forcing researchers to rely solely on pre-existing data for their research is undoubtedly a major burden and challenge. To remove this obstacle, the concept of open access was developed. Open access is a useful and affordable way to share knowledge. According to Filipi Matutinović, open access means that "any user anywhere in the world with Internet access has the right to read, download, store, print, and use the digital content of open access publications simply by citing them correctly" (Filipi Matutinović S, 2014). Open access for research data refers to the right

to access and reuse digital research data (Uribe-Tirado et al., 2020). In recent years, open access has received much attention from various industries and academic experts, and it has become a hot topic of discussion and investigation.

Research data is defined as "factual records (numerical results, textual records, images, and sounds) that serve as the primary source of scientific research. In the scientific research process, personal habits or constraints imposed by knowledge background lead to a diversity of formats for managing research data and the use of tools, a phenomenon that reflects a large number of researchers and outputs, but is not conducive to data sharing and collaboration among researchers (Ghosh & Bijan Kumar, 2022). The establishment of a research data management repository will solve this challenge. Research data repositories (Björk, 2013) unify metadata standards, content types, and software usage, enabling accessibility and usability of research data, thereby increasing transparency and credibility of the research process, improving citability, validating findings by re-analysis of data, facilitating reuse and repurposing of research data to answer different research questions, facilitating discovery, reducing duplication of effort and its ancillary costs, creating new collaborations between data users and data producers, and increasing the number of publications by authors(Misulis & Frisse, 2019).

China, recognized as one of the foremost technological leaders with unparalleled global productivity, assumes a pivotal role in driving data policies and practices (Tollefson, 2018). The "Law of the People's Republic of China on Science and Technology Progress," established in 2008, laid down foundational guidelines for research data management. Subsequently, the revised "Copyright Law of the People's Republic of China" enacted in 2010, along with the "Law of the People's Republic of China on Promoting the Transformation of Scientific and Technological Achievements" revised in 2015, further refined the legal framework. The "Measures for the Management of Scientific Data," introduced in 2018, propelled research data management and sharing to unprecedented heights. By the conclusion of 2022, 15 provincial-level administrative regions, including Hubei, Tianjin, and Hainan, had implemented regional governance regulations for scientific data in alignment with national-level laws.

Concurrently, various national-level research projects are actively advancing and being implemented. These include the National Science and Technology Infrastructure Platform, the Chinese Academy of Sciences Scientific Data Program, and the Chinese Academy of Sciences Earth Data Science Project. To cater to the requirements of institutions and disciplines for data management and services, corresponding repositories have been established, upholding stringent quality control and data-sharing practices. Notable examples encompass the Earth Science Data Repository GSCloud (www.gscloud.cn), the Spectroscopy Data Center GSA (big.big.ac.cn/GSA/), and institutional repositories such as the Peking University Open Research Data Platform (opendata.pku.edu.cn).

Re3data (http://www.re3data.org/) serves as a global research data registry center, indexing over 4,000 interdisciplinary data repositories. The Re3data database records and displays characteristics of registered research data resources, such as subject distribution, content

types, and institutional responsibilities. It has become an invaluable tool for researchers to understand and navigate open-access processes across different countries and regions.

Given the exploration and contributions in the open access domain, as well as the research gap concerning a comprehensive overview of Chinese research data repositories, an investigation into the features and operational status of these repositories in China has become particularly necessary and urgent.

LITERATURE REVIEW

To date, numerous studies have been conducted on repositories, focusing mainly on two types of repositories: Open Access repositories and open research data repositories, although some studies on Open Access institutions fall into the first category. Open Doar and re3data are the most widely used databases for repository registration. OpenDOAR (Open Access Directory Repository Catalog) is managed by the SHERPA service and is primarily intended to promote scholarship and research. The initial development of OA repositories was concentrated in the United States, Canada, and Australia, and was primarily institutional, multidisciplinary, and English-based, with access licensing likely to be a key element in the future development of repositories (Pinfield et al., 2014). After 2010, OA repositories emerged in various regions, especially in Asia, and the development of OA repositories in Asia has become an important research topic for scholars. Japan has the largest number of OA repositories but lags in terms of Web 2.0-enabled repositories. Most Web 2.0 repositories were found in Turkey, followed by China and India. It was also recommended that LIS schools should introduce "Web Design" as one of their graduate-level subjects so that LIS professionals can learn these skills and not depend on IT professionals (Khan et al., 2022). The most prestigious universities in Asia are not actively promoting the OA movement(Abrizah A. et al., 2010). Parray(2023)comparison of the current status of OA repository development in China and India based on characteristics such as repository type, software usage, repository interface language, year of development, subject coverage, content coverage, and use of Web 2.0 tools by repositories revealed that India and China focus primarily on institutional repositories, with DSpace still the preferred choice in both countries and India only English is used in repositories, while in China half of the repositories have a mixed language interface (English and Chinese). It is recommended that both India and China should implement the Turkish and Indonesian models to accelerate their growth (Singh, 2016).

Armbruster and Romary (2010) distinguish between different types of repositories, namely disciplinary repositories, research repositories, institutional repositories, and national repository systems.Re3data.org (a registry of open research data repositories) provide researchers with guidance in the heterogeneous landscape of RDR. It provides users with information about their different roles as data producers and data users(Schöpfel, 2022). Sheng et al. (2021) show that open research data repositories can effectively make research outputs visible and share them globally, providing important opportunities for the research community.

The U.S. government directs its national research agencies to maximize access to digital research data, (Office of Science and Technology Policy, 2013). The European Commission is planning a similar requirement in its Eighth Framework Program HORIZON 2020 (European Commission, 2012). In China, provinces have issued requirements and documents for building research data repositories following the release of the Measures for the Management of Scientific Data in 2018(Zhong & Jiang, 2016). In the 2019 survey of global open-access research data repositories, the United States, Germany, the United Kingdom, and Canada were the leading countries, with subject categories being the main repository type and most not following metadata standards(Bhardwaj, 2020). A survey of the BRICS countries (Brazil, Russia, India, China, and South Africa) shows that India is the pioneer and largest contributor to research repositories, and while China has done quite well over the last three decades, the country seems very reluctant to release research data through open platforms or under Creative Commons licenses (Misgar et al., 2020). However, the study only looks at characteristics such as the type of content, the language of the interface, and the distribution of disciplines, focusing more on comparing characteristics across these countries. Of the 45 RDRs in India, only 30 (67%) are open, with a concentration of disciplines in the life sciences (28), and only 20% of data repositories use metadata standards in their metadata (Khan, 2020).

In summary, the United States, Germany, and Canada lead in both types of open access and types of open research data, while India is somewhat more developed in the Asian region, and in terms of characteristics, English is the main global interface language, with difference differentiation in the use of software, disciplines, and so on. In the case of the study on open research repositories in China, it is only a comparison of the general direction of the characteristics of content types, interface languages, and disciplinary distribution, and no specific parameters such as software use and types of institutional responsibility have been addressed yet Have the content types been adjusted? Has the interface language remained consistent? These questions are also the focus of the research in this paper.

Objectives of the study

This study aims to analyze the current state of Open Research Data repositories and to describe their characteristics and performance. It addresses the following research objectives:

- To determine the subject coverage and content type archived by repositories.
- To ascertain software used and repository language being used in repositories.
- To discover the different types of repositories and Keywords.
- To identify the institution responsibility type and institution type of repositories.

RESEARCH DESIGN

The descriptive methodology used in this study is based on the following phases:

Phase 1: Selection of the data source

The research data is sourced from the world's largest indexed repository of research data, known as the Registry of Research Data Repositories (http://re3data.org/). This registry encompasses a variety of research data repositories from different countries across the globe. The research data utilized in this article was extracted from the re3data registry on April 27, 2023.

Phase 2: Data extraction, refinement and visualization

The re3data registration portal offers a range of parameters for investigating the listed data repositories. However, this study specifically gathered chosen metadata from these repositories, including subject domain, content type, repository language, software utilization, repository classification, keywords, institutional responsibility, and institutional and provider types.

The browsing tool provides three main avenues for exploring the database: browsing by subject, content type, and country. National contributions are directly derived from the repository's browsing section, yielding data that characterizes Chinese research data repositories. Based on the selected data types for this research, a targeted search and refinement process was carried out within the descriptions of Chinese research data repositories. The findings were meticulously recorded using MS EXCEL 2010 software.

In conclusion, the collected information was streamlined by eliminating extraneous details, resulting in a structured table that is primed for further analysis.

RESULTS

To the objectives of this study, the relevant data were selected for statistical analysis in order to draw conclusions and findings.

Country-wise distribution

As of April 27, 2023, a total of 4140 open-access repositories had been established worldwide. Most repositories are located in the United States (1169), followed by Germany (495), Canada (393), the United Kingdom (314), the European Union (286), France (126), Australia (101), Switzerland (84), and China (81).

What is clear is that the number of repositories is concentrated in the developed world, with the U.S., in particular, leading the world. China is in ninth place but is still far from the first-mentioned countries.

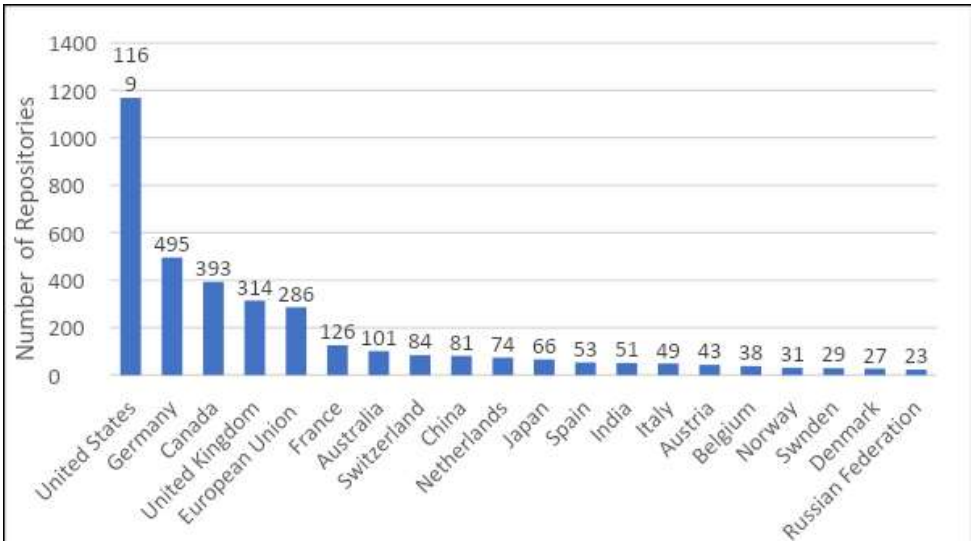

Figure 1: Distribution of Top 20 Countries in terms of Number of Open Research Data Repositories

Subject coverage and Content type

This study examined the distribution of subjects and types of content in Chinese repositories listed in re3data. Figure 2 shows that the distribution of disciplines is mainly in life sciences (44%) and natural sciences (46%), less in humanities and social sciences and engineering, with the proportion of life sciences decreasing and the proportion of engineering increasing compared to the 2019 survey. At the same time, Figure 3 shows that the distribution of specific disciplines is diverse and cross-fertilized, with genetics and cell biology generally having the most data.

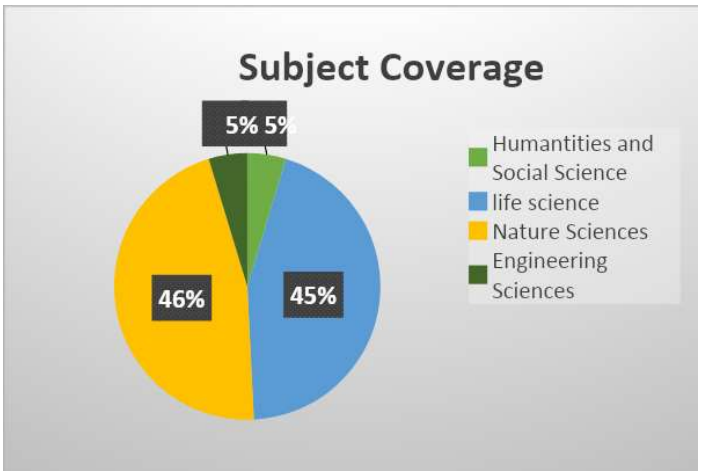

Figure 2: Discipline Distribution of China's Open Research Data Repositories

A total of 52 specific disciplines are counted, with the number of disciplines reaching 157, showing the diversity and interdisciplinarity of the distribution. The richness of the repositories is also reflected in the fact that some of them may contain multiple disciplines. Among them, general genetics (16, 10.18%), cell biology (11, 7.01%), geology, and paleontology (7, 4.46%) occupy the top three places numerically. The statistics on specific disciplines are also one of the gaps filled in the last survey.

Table 1. Discipline-specific distribution of China's open research data repositories

Subjects	**Numbers**	**Percentage**	**Subjects**	**Numbers**	**Percentage**
General Genetics	16	10.18%	Systematics and Morphology	1	0.63%
Cell Biology	11	7.01%	Biochemistry and Animal Physiology	1	0.63%
Geology and Palaeontology	7	4.46%	Microbial Ecology and Applied Microbiology	1	0.63%
Animal Ecology, Biodiversity and Ecosystem	5	3.20%	Public Health, Health Services Research, Social Medicine	1	0.63%
Human Genetics	5	3.20%	Physiology	1	0.63%
Oceanography	5	3.20%	Pathology and Forensic Medicine	1	0.63%
Geochemistry, Mineralogy and Crystallography	5	3.20%	Pharmacy	1	0.63%
Metabolism, Biochemistry and Genetics of Microorganisms	4	2.58%	Reproductive Medicine/Biology	1	0.63%
Geophysics	4	2.58%	Cellular Neuroscience	1	0.63%
Geodesy, Photogrammetry, Remote Sensing, Geoinformatics	4	2.58%	Developmental Neurobiology	1	0.63%
Geography	4	2.58%	Soil Sciences	1	0.63%
Water Research	4	2.58%	Plant Cultivation	1	0.63%
Astrophysics and Astronomy	3	1.91%	Ecology of Agricultural Landscapes	1	0.63%
Atmospheric Science	3	1.91%	Plant Breeding	1	0.63%
Social Sciences	2	1.27%	Basic Forest Research	1	0.63%
Economics	2	1.27%	Biological Chemistry and Food Chemistry	1	0.63%
Biochemistry	2	1.27%	Optics, Quantum Optics and Physics of Atoms,	1	0.63%

			Molecules and Plasmas1		
Anatomy	2	1.27%	Geochemistry, Mineralogy and Crystallography	1	0.63%
Plant Ecology and Ecosystem Analysis	2	1.27%	Thermal Engineering/Process Engineering	1	0.63%
Epidemiology, Medical Biometry, Medical Informatics	2	1.27%	Heat Energy Technology, Thermal Machines, Fluid Mechanics	1	0.63%
Particles, Nuclei and Fields	2	0.63%	Materials Science	1	0.63%
Fine Arts, Music, Theatre and Media Studies	1	0.63%	Computer Science	1	0.63%
Structural Biology	1	0.63%	Software Technology	1	0.63%
Plant Systematics and Evolution	1	0.63%	Systematics and Morphology	1	0.63%
Inter-organismic Interactions of Plants	1	0.63%	Biochemistry and Animal Physiology	1	0.63%
Plant Biochemistry and Biophysics	1	0.63%	Microbial Ecology and Applied Microbiology	1	0.63%

The diversity of data content types enhances the popularity among researchers of different disciplines. From Figure 3, it can be concluded that Chinese open research repositories contain a total of 14 categories of content types, numbering 314, reflecting the fact that a repository may contain a variety of content types. Of these, scientific and statistical data formats (59), images (40), structured text (38), standard office documents (32), and raw data (29) are the main content types.

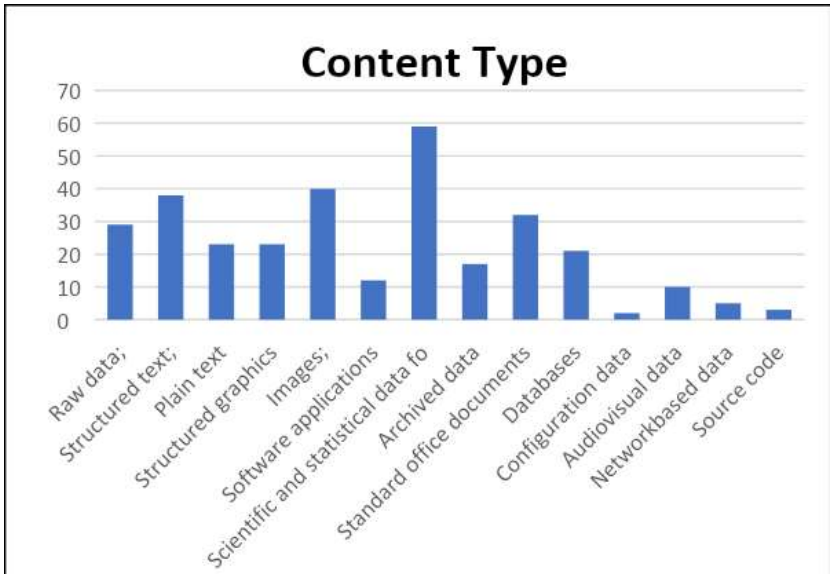

Figure 3: Types of Content in China's Open Research Data Repositories

Software used and Repostories Language

Research data repositories have been created in multiple languages to support users in their Language interface . A total of 24 languages have been observed across the 81 repositories. Of these, English, as the most spoken language worldwide, has the highest number, reflecting the wide applicability of the repositories; Chinese, as the native language of China, is in second place; the remaining language is only one due to "Fishbase", a global species repository with mirror sites in English, German, French, Spanish, Portuguese, French, Swedish, Chinese and Arabic, Swedish, Chinese and Arabic mirror sites. (However, the use of software is not encouraging, with most being "unknown" and only a few repositories using software such as MySQL and DataVerse tools.

Table 2. Language Interface and Software Usage of Open Research Data Repositories in China

Repostories languages	Numbers	Soft ware	Numbers
English	61	MySQL	6
Chinese	56	DataVerse	2
Arabic	1	other	3
Bengali	1	unknown	23
German	1		
Greek, Modern	1		
Persian	1		
French	1		
Gujarati	1		
Hindi	1		
Indonesian	1		

Italian	1
Japanese	1
Kannada	1
Lao	1
Malayalam	1
Marathi	1
Dutch	1
Portuguese	1
Spanish; Castilian	1
Swedish	1
Tamil	1
Telugu	1
Thai	1

Types of Repositories and Keywords

There are mainly three types of repositories enlisted in re3data.org: Other (a repository whose type has not yet been assessed), Institutional (an institutional or departmental repository), Disciplinary (a cross-institutional subject repository). The number of disciplinary categories is as high as 73, accounting for 90.12%; 6 are institutional categories, and 5 others. "Inspire-HEP", "Protein Lysine Modification Database", and "eLibrary of Microbial Systematics and Genomics "are both disciplinary and institutional types. The top three repositories are Genomics, Genetics, and Ecology, which are consistent with the number of specific disciplines.

Table 3. China's Open Research Data Repository Keywords

Keywords	Numbers	Percentage	Keywords	Numbers	Percentage
genomics	11	10.76%	environmnet	3	2.94%
genetics	9	8.82%	genes	3	2.94%
ecology	5	4.90%	germplasm	3	2.94%
FAIR	4	3.93%	metabolomics	3	2.94%
archaea	4	3.93%	molecular biology	3	2.94%
bacteria	4	3.93%	BLAST	2	1.96%
bioinformatics	4	3.93%	DNA	2	1.96%
genome	4	3.93%	atmosphere	2	1.96%
proteomics	4	3.93%	biomedicine	2	1.96%
RNA	3	2.94%	climate	2	1.96%
agriculture	3	2.94%	diseases	2	1.96%
algae	3	2.94%	epigenomics	2	1.96%
biology	3	2.94%	fungi	2	1.96%
cell lines	3	2.94%	geography	2	1.96%
chromosome	3	2.94%	geology	2	1.96%

Institution Type and Institution Responsibility Type

The nature of the institution and the responsibilities of its members are an important basis for understanding the internal organizational framework of a repository. Currently, there are 79 non-profit repositories and 3 commercial types, while "GigaDB" consists of the non-profit - Beijing Genomics Institute and GigaScience Press and the commercial - China National Genebank together. The types of institutional responsibility are divided into general, technical, and funding and sponsorship, with the percentages shown in Figure 4.

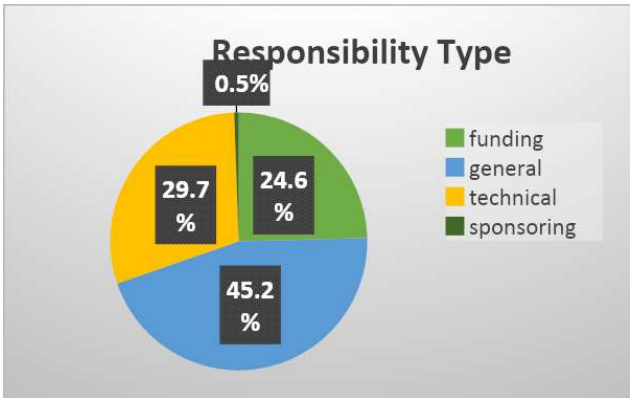

Figure 4: Institutional Responsibility for Open Research Data Repositories in China

Provider Type

The study analyzed the type of provider to determine whether the research data repository was a data provider or a service provider. Figure 5 shows that the majority of research data repositories in China are data providers 78 (96%), containing data only 40 (49%) and data and services together 38 (47%); and service providers 41 (51%), most of which are co-providers, while only three (4%) provide services only. This indicates that data provision is the primary focus. However, to ensure sustainability and encourage researchers to use the database, service capacity must also be strengthened, services expanded, user needs met, and necessary support and training provided.

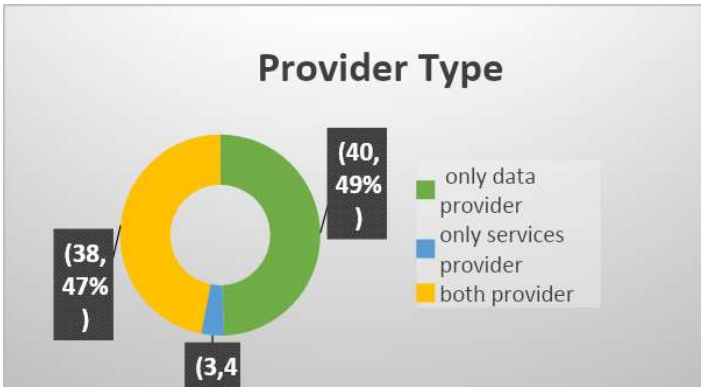

Figure 5: Provider Types of Open Research Data Repositories in China

CONCLUSION

The collection and recording of research data take on a diverse range of forms, but due to the lack of consistent standards, sharing can be challenging. Re3data.org serves as a registered platform for managing and storing scientific research data, providing us with a valuable data source for our research endeavors. Analysis reveals that China, as one of the most populous countries with a significant number of researchers, still lags behind developed nations like the United States, Canada, and Germany in terms of the quantity of registered research data repositories.

The distribution of disciplines and content types is varied, yet there exists an uneven distribution across disciplines. The predominant types are in the realm of life sciences and natural sciences, whereas humanities, social sciences, and engineering categories are relatively underrepresented. This phenomenon can be attributed to the contributions and leadership of repositories established by the Chinese Academy of Sciences and its affiliated institutions in China's open-access initiatives. These repositories predominantly cover life sciences and natural sciences, particularly focusing on biology and geography. Encouraging universities and other higher education institutions to establish more databases catering to humanities and social sciences is recommended. This initiative would strengthen data integration and sharing among repositories of different types, thereby enhancing overall accessibility.

As a country with Chinese as its native language, Chinese language repositories are expected to be prevalent. However, the number of repositories with English interfaces surpasses those with Chinese interfaces. This dual-language approach can be attributed to the usage of English as the international interface language for data repositories, catering to the world's largest linguistic community. It also underscores the broad applicability of Chinese repositories, as most of them utilize both English and Chinese languages. This approach fosters openness and inclusivity for sharing and exchanging data. Nonetheless, the state of software usage isn't entirely optimistic. The lack of uniformity and commonality in software usage can pose barriers to data integration and utilization, potentially leading to negative attitudes among users toward sharing data. It's advisable for disciplinary or industry associations, possessing authority and influence, to endorse software tools relevant to their respective fields or industries.

Repository types primarily align with disciplines, constituting a substantial 90.12%. This emphasizes the importance of disciplinary associations, higher education institutions, and similar entities as the main consumers and producers of research data. Some repositories feature dual types—discipline and institution—which reflect collaboration and sharing in aspects such as data utilization and platform establishment. The distribution of keywords mirrors the distribution of disciplines, reaffirming the pivotal role and status of discipline-specific repositories.

Among the 81 registered data repositories, 79 are non-profit organizations. This conclusion is closely tied to repository types. Higher education institutions, being the primary entities

associated with disciplinary types, are predominantly non-profit organizations and pivotal drivers of disciplinary advancements. It is anticipated that more commercial entities will sign up with a responsibility type focused on funding or technical support, enriching the landscape of research data management. The responsibility types for institutions include "general," "technical," and "funding," with instances where the "general" type encompasses the other two or is ambiguous. Given the functionalities and goals of repositories, data providers are the primary provider type. The proportion of both data and data service offerings is expected to rise in the future, aligning with the maturation of database establishment and catering to the deeper needs of users.

In the context of open access in China, this study utilizes data from the re3data website and draws insights from current research on similar repositories in different countries and regions to examine the characteristics of Chinese research data repositories. In comparison to prior studies on the development of repositories in various nations, this article engages in a more comprehensive exploration of specific aspects of repositories, including disciplines, repository types, and institutional responsibilities. Through data analysis, the research reflects the current state and trajectory of research data repositories in China, unveiling challenges such as uneven disciplinary distribution and suboptimal software utilization. This study provides a valuable reference for the development of repositories within the context of open access initiatives and offers a dataset for future academic research.

However, certain limitations should be acknowledged. First, this study does not extensively delve into repository policies. Although it touches on the existence of policies, a more comprehensive investigation necessitates content and textual analysis across multiple data sources to identify policy types and key points. Second, legal considerations, such as data licenses, database licenses, and data download availability, were not extensively explored. These legal factors dictate how users can interact with data within repositories.
To expand the study's scope, future research could consider the following directions:

Policy Inclusion: Include an in-depth analysis of repository policies, involving content and textual analysis across various sources to categorize policy types and summarize key aspects.

Legal Considerations: Expand the investigation of legal dimensions, including data licenses, database licenses, and data download availability, as these significantly influence how users can engage with repository data.

Comparative Studies: Conduct cross-regional and international comparative studies, with a focus on distinct regions and international organizations. Additionally, deeper longitudinal investigations in countries where open access efforts are just beginning, especially in developing nations, can contribute to creating an environment conducive to internationalization, disciplinary specialization, and integrated sharing within open access initiatives.

In conclusion, this study comprehensively explores Chinese research data repositories within the framework of open access. While some areas warrant further exploration, the research

findings contribute to a better comprehension of the present landscape and challenges, offering insights for advancing open-access initiatives.

REFERENCES

Abdala, C. V., and Rosane Taruhn. 2021. Access to Health Information in Latin America and the Caribbean. *Electronic Resources in Medical Libraries*, 41-50. doi:10.4324/9780203051948-5.

Armbruster, Chris, and Laurent Romary. 2010. Comparing Repository Types. *International Journal of Digital Library Systems,* 1(4), 61-73. doi:10.4018/jdls.2010100104.

Bhardwaj, Raj K. 2020. Open Research Data Repositories: A Content Analysis to Comprehend Data Equitable Access. *Journal of Scientometric Research,* 8(3), 135-142. doi:10.5530/jscires.8.3.29.

Björk, Bo-Christer. 2013. Open access subject repositories: An overview. *Journal of the Association for Information Science and Technology,* 65(4), 698-706. doi:10.1002/asi.23021.

China, People's Republic of - Discrimination. (n.d.). Foreign Law Guide. https://doi.org/10.1163/2213-2996_flg_com_323176

Department of Defense Washington DC. 2015. *Department of Defense Plan to Establish Public Access to the Results of Federally Funded Research*. doi:10.21236/ada614321.

Khan, Aasif M., Fayaz A. Loan, and Syed A. Ahmad Andrabi. 2022. Web 2.0 Applications in Open Access Institutional Repositories of Asia. *DESIDOC Journal of Library & Information Technology,* 42(3), 149-153. doi:10.14429/djlit.42.3.17473.

Law of the People's Republic of China on Science and Technology Progress, Standing Committee of theNational People's Congress, P.R.C., Art.65 2008. Available at: http://www.gov.cn/ziliao/flfg/2007-12/29/content_847331.htm [Accessed: 8 Sep 2020].

Misgar, Safat M., Ajra Bhat, and Zahid A. Wani. 2020. A study of Open Access research data repositories developed by BRICS countries." *Digital Library Perspectives,* 38(1), 45-54. doi:10.1108/dlp-02-2020-0012.

Pulst, P. 2021. The copyright law of the People's Republic of China. LIT Verlag Münster.

Parray, Umer Y., Aasif M. Khan, Aasif A. Mir, and Shahid M. Mir. 2023. Unveiling the present status of open access repositories: a comparative analysis of India and China. *Library Management,* 44(½), 120-132. doi:10.1108/lm-09-2022-0084.

Pinfield, Stephen, Jennifer Salter, Peter A. Bath, Bill Hubbard, Peter Millington, Jane H. Anders, and Azhar Hussain. 2014. Open-access repositories worldwide, 2005-2012: Past growth, current characteristics, and future possibilities. *Journal of the Association for Information Science and Technology,* 65(12), 2404-2421. doi:10.1002/asi.23131.

Schöpfel, Joachim. 2022. Issues and Prospects for Research Data Repositories." *Research Data Sharing and Valorization*, 211-229. doi:10.1002/9781394163410.ch12.

Sheng, Jie, Joseph Amankwah-Amoah, Zaheer Khan, and Xiaojun Wang. 2020. COVID-19 Pandemic in the New Era of Big Data Analytics: Methodological Innovations and

Future Research Directions. *British Journal of Management,* 32(4), 1164-1183. doi:10.1111/1467-8551.12441.

Singh, Prerna. 2016. Open access repositories in India. *IFLA Journal,* 42(1), 16-24. doi:10.1177/0340035215610131.

Tapfuma, Mass, and Ruth Hoskins. 2020. Visibility and Accessibility of Indigenous Knowledge on Open Access Institutional Repositories at Universities in Africa. *Digital Libraries and Institutional Repositories*, 454-472. doi:10.4018/978-1-7998-2463-3.ch028.

Tollefson, J. 2018. China declared world's largest producer of scientific articles. *Nature*, 553(7689), 390-390. https://doi.org/10.1038/d41586-018-00927-4

Uribe-Tirado, Alejandro, Gimena Del Río Riande, Sandra Raiher, and Jaider Ochoa-Gutiérrez. 2020. Open Science since COVID-19: Open Access + Open Data. *SSRN Electronic Journal*, doi:10.2139/ssrn.3621047.

Zhong, Jing, and Shuyong Jiang. 2016. Institutional Repositories in Chinese Open Access Development: Status, Progress, and Challenges. *The Journal of Academic Librarianship,* 42(6), 739-744. doi:10.1016/j.acalib.2016.06.015.