Springer Nature Research Data Policy Standardization Practice and Implications Postprint

Authors: Wang Dandan, Liu Qinghua, Ge Liyun

Date: 2023-04-01T16:16:01+00:00

Abstract

[Purpose/Significance] Springer Nature has made pioneering contributions to the standardization and unification of research data policies. This study investigates the research data policy standard framework proposed by Springer Nature and its promotion and application, aiming to provide references and insights for Chinese scientific journal management institutions to formulate and implement a research data policy standard framework suitable for Chinese scientific journals. [Method/Process] Through web searching, text analysis, website investigation, and comparative analysis, this study deconstructs Springer Nature's policy standard framework, examines the support services provided by Springer Nature for promoting the policy standard framework, systematically reviews and summarizes Springer Nature's practical efforts in advancing research data policy standardization, and reveals its reference significance. [Results/Conclusion] Springer Nature's policy standardization framework can serve as a reference for relevant Chinese departments in developing research data policy standard frameworks. The third-party monitoring services developed by Springer Nature to promote its policy standard framework demonstrate a practical pathway for publishers to participate in research data management services. It is recommended that Chinese scientific journal management institutions take the lead, convene stakeholder groups, draw on international advanced experiences to construct a scientifically sound policy framework, and design supporting services to effectively promote the application of the policy framework.

Full Text

Springer Nature's Practice in Standardizing Research Data Policies and Its Implications

Authors: Wang Dandan, Liu Qinghua, Ge Liyun
Affiliation: School of Management, Henan University of Science and Technology, Luoyang 471023

Abstract

[**Purpose/Significance**] Springer Nature has made pioneering contributions to promoting the standardization and unified development of research data policies. This study examines the standard framework for research data policies proposed by Springer Nature and its implementation, aiming to provide references for Chinese scientific journal management agencies to develop and promote suitable research data policy frameworks for Chinese journals. [**Method/Process**] Through web searches, text analysis, website visits, and comparative analysis, this paper analyzes Springer Nature's policy standardization framework, investigates the supporting services provided to facilitate its implementation, and systematically summarizes the practices and implications of Springer Nature's efforts to standardize research data policies. [**Result/Conclusion**] Springer Nature's policy standardization framework can serve as a reference for relevant Chinese departments in formulating research data policy frameworks. The third-party curation services developed by Springer Nature to promote its framework demonstrate a practical pathway for publishers to participate in research data management services. It is recommended that Chinese scientific journal management agencies take the lead in convening stakeholder groups, draw on international advanced experience to construct scientifically sound policy frameworks, and design supporting services to effectively promote their application.

Keywords: publisher; research data; policy; service

In today's data-intensive scientific environment, technological innovation increasingly depends on access to and analysis of research data, making the storage and sharing of research data ever more critical. Research data policies significantly influence researchers' willingness to share data to varying degrees [1], and an increasing number of funding agencies have issued policies on research data sharing, including the National Institutes of Health (NIH), Gates Foundation, EU Horizon 2020 program, Wellcome Trust, and seven UK research councils. These agencies' policies vary considerably—some require researchers to prepare data management plans, while others demand evidence of data archiving in published papers, such as the Engineering and Physical Sciences Research Council (EPSRC).

Since 2015, an increasing number of journals and publishers have recognized and begun implementing the Transparency and Openness Promotion (TOP) Guidelines. The TOP Guidelines represent a comprehensive yet aspirational set of journal policies comprising eight modular standards [2]. Journals and publishers treat full compliance with the TOP Guidelines as a long-term goal, adjusting their policies to progressively approach this target. Research has shown that publishers and journals can help raise researchers' awareness of data sharing issues by developing and continuously improving their policies [1]. Although the number of publishers and journals with research data policies is growing, these policies often differ substantially in content, discoverability, interpretability,

infrastructure integration, and compliance support because they are typically developed based on the needs and norms of individual research communities [3]. This lack of unified standards makes it difficult for journal editors to formulate and support data policies, challenging for researchers to understand and comply with them, and complex for infrastructure providers and research support staff to assist authors in policy adherence [4].

This has made research on standardized and relatively unified research data policies a new hotspot. Discipline-specific initiatives have emerged in chemistry, high-energy physics, and medicine to standardize and unify research data policies across journals and publishers. In 2017, the International Committee of Medical Journal Editors (ICMJE) introduced standardized data sharing policies for its member journals (including *BMJ*, *The Lancet*, *JAMA*, and *New England Journal of Medicine*) [5]. While not mandating data sharing, the ICMJE policy requires data availability statements. Such statements are common elements in journal and publisher data policies—*PLOS ONE*, *Nature*, and *BMC* all require them [6]. Some research funding agencies, such as the UK's seven research councils, also require data availability statements in published articles [7]. Data availability statements are also mandated by the UK Research and Innovation (UKRI) Common Principles on Data Policy [7]. Studies have shown that journal policies reflected in submission guidelines help improve the transparency of authors' research [8-9], and journals with higher impact factors tend to adopt stronger data sharing policies [10]. Stringent data policies not only require verification of authors' data sharing practices but also demand data availability statements.

Springer Nature has conducted valuable exploratory practices in policy standardization by introducing a research data policy standardization framework. On one hand, this framework provides journals with references and foundations for quickly developing data policies that follow best practices and can be effectively implemented; on the other hand, it offers a communication platform for stakeholder groups [11]. To facilitate framework implementation, Springer Nature has also launched data availability guidelines and developed third-party data curation services. The effectiveness of Springer Nature's policy standardization efforts has been remarkable. The authors simultaneously investigated other foreign publishers such as Elsevier, Wiley, PLOS, and Taylor & Francis to understand their progress in journal policy standardization. While these publishers have recognized the importance of standardization and undertaken similar initiatives—the Research Data Alliance Interest Group has even developed a policy framework to coordinate data policies across all publishers [12]—their standardization practices lag behind Springer Nature's in terms of systematicity and completeness, and their frameworks have not achieved comparable implementation success. Therefore, by analyzing Springer Nature's policy standardization framework, examining its supporting services for framework promotion, and systematically summarizing its practices, this study aims to provide references for Chinese scientific journal management agencies to develop and implement suitable research data policy frameworks.

2. Springer Nature's Research Data Policy Standardization Framework

2.1 Background

Springer Nature observed that many of its journals still lacked clear research data policies, and those with policies often had ambiguous mandatory or optional requirements and lacked clear implementation guidance. The publisher recognized that journals and publishers have a responsibility to support researchers in complying with funding agencies' policies (whether for open access or open data). Beyond policy compliance, linking data to scholarly publications can increase paper visibility and citation rates, help publishers produce richer content, and improve data access and understanding to drive further research—practices in crystallography, genetics, archaeology, and linguistics have demonstrated these benefits [13]. To achieve these benefits and address the problem of diverse and complex journal policies, Springer Nature launched a research data policy standardization framework applicable to over 2,500 journals in its publishing group in 2016 [14]. Each journal under its umbrella can formulate and optimize its research data policy based on this framework while tailoring it to the characteristics of its discipline and community.

The framework's purposes are: (1) to provide journals with evidence-based guidance for policy formulation and optimization, ensuring both uniformity and personalization; (2) to help researchers meet data sharing requirements of research institutions and funding agencies; and (3) to ensure all journals and authors can publish results following best practices in data sharing and archiving for their disciplines. To achieve these goals, Springer Nature first convened an expert working group that consulted over 30 journal editors across all disciplines, as well as funding agencies and librarians, to identify nine common elements for Springer Nature journals' research data policies: (1) support for data sharing via repositories; (2) permitting data citation; (3) publisher helpdesk; (4) public data deposition and dataset identifier checks for specific data types; (5) data availability statements; (6) requiring and verifying public data deposition and dataset identifiers; (7) data citation standards; (8) peer review of data; and (9) integrated data repositories. Second, the group solicited extensive feedback on these nine elements from journal editors, librarians, and research funders. Finally, considering the differential applicability of these elements across journals and disciplines, four types of policy templates were proposed (see). Regardless of discipline, journals can select the policy type best suited to their needs based on available resources in their disciplinary communities.

**** Four Types of Policy Templates in the Common Policy Framework [11]

- **Type 1:** Encourages data sharing and data citation
- **Type 2:** Encourages data sharing and research materials sharing
- **Type 3:** Encourages data sharing and requires data availability statements
- **Type 4:** Requires data and research materials sharing; shared data must

undergo peer review; and provides availability statements

Individual journals and research communities may be at different stages of readiness to support data sharing—some have already mandated open data, while others are just beginning to discuss whether to share data. The four policy template types respect these developmental differences while providing a common and easily understandable framework to encourage good data sharing practices across all research fields. The stringency of data policies increases progressively from Type 1 to Type 4. Type 1 encourages data sharing and citation and provides researchers with a list of data repositories; Type 2 recommends providing information on preparing data availability statements; Type 3 requires data availability statements; and Type 4 mandates open data and requires data peer review.

**** Implementation of Four Policy Types Across Nine Common Elements [11]

Element	Type 1	Type 2	Type 3	Type 4
1. Support for data sharing via repositories	Detailed information on sharing via repositories provided in author guidelines	Detailed information on sharing via repositories provided in author guidelines	Datasharingviarepositoriessupported	Datasharingviarepositoriessupported
2. Permit data citation	Datacitationpermitted	Datacitationpermitted	Datacitationpermitted	Datacitationpermitted
3. Publisher helpdesk	Helpdesk contact information included in author journal information	Helpdesk contact information included in author journal information	Helpdesk contact information included in author journal information	Helpdesk contact information included in author journal information
4. Public data deposition and dataset identifier checks for specific data types	No requirement	Recommended	In communities with established mandatory requirements, data deposition is checked as part of the publication process	In communities with established mandatory requirements, data deposition is checked as part of the publication process

Element	Type 1	Type 2	Type 3	Type 4
5. Data availability statements	No requirement	Recommended	Required	Required
6. Require and verify public data deposition and dataset identifiers	No requirement	No requirement	Required	Required
7. Data citation standards	Provide and verify relevant dataset citations in reference lists	Provide and verify relevant dataset citations in reference lists	Provide and verify relevant dataset citations in reference lists	Provide and verify relevant dataset citations in reference lists
8. Peer review of data	No requirement	No requirement	No requirement	Peer reviewer guidelines and processes provide information on accessing and reviewing data files
9. Integrated data repository	No requirement	No requirement	No requirement	Submission system/review process integrated with journal-specific or general repositories

All four policy types mandate "support for data sharing via repositories" and "permitting data citation." Types 1 and 2 have relatively lenient requirements, giving researchers considerable autonomy, while Types 3 and 4 are more stringent, primarily targeting high-value data to ensure maximum development and reuse through mandatory requirements.

2.3 Element Analysis

To support framework application, Springer Nature launched a research data helpdesk in July 2016 [15]. The helpdesk provides free email-based advice to journal editors, authors (researchers), and support staff to help them implement and comply with journal data policies. M. Astell et al. found that by October 2017, the helpdesk had received over 300 queries. More than half (53%) of those seeking help were researchers (authors), 43% were journal editors, and the remaining 4% were other stakeholders such as librarians and repository managers. The four most common query categories were policy implementation (122), data repositories (53), policy compliance (46), and writing data availability statements (31) [16]. In response, Springer Nature compiled and provided a list of recommended data repositories [17] and data availability statement guidelines [18] on its website.

This section uses the data availability statement element as an example to reveal Springer Nature's specific practices. Data availability statements, written by authors, aim to provide information about where data supporting reported results can be found and how to access them. Although these statements have no mandatory format, journals should generally provide template examples for authors. In September 2016, *Nature* announced that all accepted papers in *Nature* and its life science journals must include data availability statements [19]. By introducing data availability statements, *Nature* aligned its policies with Springer Nature's policy standardization framework.

In Springer Nature's policy framework, Types 2, 3, and 4 encourage or mandate data availability statements. Different data types have varying levels of openness—some can be openly accessed, while privacy-sensitive or confidential data require permission. Therefore, different data types should have their own unique "availability statements." Common examples include: (1) "The datasets generated during and/or analyzed during the current study are available in the [NAME] repository, [PERSISTENT WEB LINK TO DATASETS]"; (2) "The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request"; (3) "All data generated or analyzed during this study are included in this published article (and its supplementary information files)"; (4) "The datasets generated during and/or analyzed during the current study are not publicly available due to [REASON(S) WHY DATA ARE NOT PUBLIC] but are available from the corresponding author on reasonable request"; (5) "Data sharing not applicable to this article as no datasets were generated or analyzed during the current study"; and (6) "The data that support the findings of this study are available from [THIRD PARTY NAME] but restrictions apply to the availability of these data, which were used under license for the current study, and so are not publicly available. Data are however available from the authors upon reasonable request and with permission of [THIRD PARTY NAME]."

If a journal has no explicit requirements for data availability statements, au-

thors can use or adapt the above statements to determine the most suitable one for their paper. Depending on the nature of the research, a combination of statements may be necessary. Springer Nature provides examples of data availability statements supported in journal articles, as shown in .

**** Examples of Data Availability Statements Supported in Journal Articles

Template/Example Text	Description
The [data type, e.g., “sequence”] data supporting this study’s findings have been deposited in [repository name, e.g., “GenBank”] with the [primary] accession code(s) [list linked accession codes, e.g., “kp253039”]	Data deposited in repository with accession codes
The [data type] data supporting this study’s findings are available in [repository name, e.g., “figshare”] at [data DOI], e.g., “http://doi.org/10.6084/m9.figshare.1499292” [reference number]	Data in repository with DOI
The [data type, e.g., “Snowpack depth”] data supporting this study’s findings are available from [repository/resource name, e.g., “GlobSnow”], [hyperlink to dataset/source, e.g., “http://www.GuangS.If/”]	Data from external resource with hyperlink
The authors declare that [all other] data supporting this study’s findings are available within the article [and its supplementary information files]	Data included in article/supplementary files
Data supporting this study’s findings are available from the corresponding author [author initials] upon reasonable request. Due to [national restrictions, e.g., “containing information that could compromise research participant privacy/consent”], these data are not publicly available	Restricted data available upon request
Data supporting this study’s findings are available from [third party name] but availability is restricted to current study license, so not publicly available. However, data are available from authors upon reasonable request and with [third party name] permission	Third-party restricted data

Template/Example Text	Description
Datasets derived from public resources and provided with article	Datasets analyzed during current study available in [repository name] repository [data identifier, e.g., doi:10.7910/DVN/HEWGDD] [reference number]. Derived from public domain resources: [list resources and URLs]
Data generated at central large facilities	Raw data generated at [facility name] large facility. Derived data supporting findings available from corresponding author upon request
Data sharing not applicable—e.g., article describing theoretical research with no new data	Data sharing not applicable as no datasets generated/analyzed during current study. If no auxiliary data sources used and no new data created, but data access statement required: "No new data created during study"
Publicly available data	All data created in study publicly available from University of Bath Research Data Archive (insert DOI). All supporting data provided as supplementary information. All data fully provided in article results

Template/Example Text	Description
Citing multiple datasets	Cite multiple datasets in article references section like citing papers; create separate archive record linking all datasets if archived in different locations. Publication supported by multiple datasets publicly available at reference citations
Secondary data analysis	Reanalysis of existing publicly available data at reference citations. Data processing documentation available at University of Bath Research Data Archive [DOI]. Study compiled existing data available upon request with multi-source license restrictions. Details on dataset acquisition in [DOI]

Template/Example Text	Description
Non-digital data	If legal/ethical reasons prevent public data sharing, state reason in data access statement and specify if data fully restricted or accessible under conditions. Supporting data not publicly available for ethical reasons. More information on data and access conditions at University of Bath Research Data Archive [DOI]. Due to study sensitivity (commercial/political/ethical), no respondent consent to retain/share data. More details available from University of Bath Research Data Archive [DOI]. Supporting data available from UK Data Service [DOI] for bona fide researchers but requires registration
Contractual restrictions	If agreements, pending patents, or contractual restrictions limit data sharing, provide this information in data access statement. Due to confidentiality agreements with collaborators, data only available to bona fide researchers under confidentiality agreement. Details at University of Bath Research Data Archive [DOI]

Template/Example Text	Description
Using secondary/third-party data	When using secondary or third-party owned data, unlikely to archive/share original data. Provide data source and access arrangements. For multiple sources, consider documenting all sources, access methods, and processing. Reanalysis of existing publicly available data at reference citations. Processing documentation at University of Bath Research Data Archive [DOI]. Non-digital supporting data stored by corresponding author at University of Bath. Access details in documentation at University of Bath Research Data Archive [DOI]

Data availability statements represent a simple, interoperable mechanism for communicating data availability among stakeholders, publishing platforms, and research disciplines, helping funders, publishers, and research communities monitor policy compliance. Their prevalence is expected to grow. However, adding data availability statements to accepted articles increases editorial processing time. R. Grant et al. found that adding data availability statements adds an average of 10 minutes per article for professional editors, but this time is acceptable given the statements' importance and benefits [20]. For journal editorial offices, disciplinary differences in data sharing practices must be considered when requiring data availability statements. For example, the life sciences have numerous community repositories and long-standing data sharing requirements, making data availability statement requirements in *Nature* journals readily implementable. In high-energy physics, however, generated and analyzed data may be so extensive that online sharing is challenging, making "data available upon reasonable request" a pragmatic option.

3. Supporting Services for Framework Implementation

3.1 Service Overview

To coordinate framework implementation, Springer Nature has introduced supporting services focused on curating datasets for peer-reviewed publications (Springer Nature defines datasets as collections of files supporting research projects, studies, or publications that may contain multiple files and various file types), while addressing the diverse needs of journal editors, peer reviewers, readers, and authors (data creators). These services are categorized into researcher-facing and journal-facing services (see [21]). Researchers and journal editors submit service requests via email, and Springer Nature responds within two working days. Services are divided into free data services (for researchers and all journal editors) and paid data services (for researchers only). Simultaneously, Springer Nature encourages researchers to publish data in data journals such as *Scientific Data* and *BMC Research Notes*, and has established growing research data sharing learning and exchange communities (e.g., data dialogues, resources, and event discussion communities [22]) to provide interactive platforms where researchers can learn about the latest developments in research data.

**** Springer Nature's Supporting Services [21]

Service Type	Free Data Services	Paid Data Services
Content	For researchers: policy advice and consultation on data policies, repository selection, and writing availability statements and data citations For journal editors: policy framework advice and policy development consultation	Upload datasets to figshare-hosted secure online repository Professional research data editors check file integrity, sensitive data, and alignment with manuscripts Enhance metadata for discoverability, accessibility, interoperability, and reusability Coordinate data publication with author paper publication

3.2 Third-Party Data Curation Services

To better understand the practical challenges researchers face when sharing data, Springer Nature conducted a survey in 2018 that received over 7,700 responses. Respondents identified the primary barrier to data sharing as "organizing data in a presentable and useful way" (46%). Other common challenges in-

cluded “uncertainty about copyright and licensing” (37%), “not knowing which repository to use” (33%), “no time to deposit data” (26%), and “cost of sharing data” (19%). These reported barriers suggest that researchers wishing to curate, share, manage, and archive data may lack skills, support, or resources. Interviews with funding agencies (n=13) and higher education institutions (n=15) revealed that metadata skills and expertise are commonly considered barriers to data sharing [23], consistent with previous surveys by publishers Wiley and Elsevier [24-25]. Springer Nature believes researchers should have faster, easier pathways to data deposition, but training researchers to become data management experts is not its goal. Instead, Springer Nature views close collaboration with stakeholder groups to find simplified data deposition workflows and develop third-party curation services (paid services) as the path to success.

Springer Nature’s introduction of paid curation services provided by independent third-party institutions aims to: (1) help researchers more easily meet data sharing requirements of publishers, institutions, and funders; (2) enhance data discoverability and reusability; (3) reduce researchers’ time spent on data deposition, sharing, and management; (4) enable scientific presentation of data; and (5) increase researcher visibility, impact, and likelihood of research utilization.

In 2017, Springer Nature launched a pilot service [24] to provide more support for researchers wishing to share their data. This service offers hands-on assistance with uploading data to repositories, selecting appropriate licenses, enhancing metadata, and cross-referencing data with associated publications. Data are hosted and preserved in Springer Nature’s figshare repository. Initially available to authors submitting to selected BMC journals, the service expanded to nearly 100 journals and, from January 2018, became available to any researcher or institution with published results. The service focuses on metadata management (drafting complete, accurate, and appropriate metadata for researchers to comply with FAIR principles) while preventing inappropriate publication formats and licensing agreements for sensitive data. Publishing and sharing sensitive data is challenging; although service staff check submissions against established data anonymization guidelines and may request detailed participant consent information for data publication, authors remain ultimately responsible. Therefore, the best approach is recommending that authors deposit “non-anonymized data” in protected repositories with managed or restricted access, such as the UK Data Archive.

The service completed its pilot phase in early 2018 and now charges fees to ensure sustainability, though development and testing continue—for example, considering data management support for conferences. Researchers wishing to use the service must complete a service inquiry form (see) providing data characteristics and confirming compliance with service standards (such as “Types of Research Data We Can Help Organize and Share” [28], “Research Data Support Submission, Checking, and Output” [29], and “Research Data Support Terms and Conditions” [30]). During service provision, all time-consuming data management tasks are handled by professional data editors, including: (1) iden-

tifying data repositories and providing policy compliance advice; (2) advising on writing data availability statements and data citations; (3) uploading data to figshare (up to 50GB); (4) checking for sensitive data; (5) ensuring end-user data accessibility; (6) ensuring data files match metadata and associated publications; (7) organizing and collecting files structurally; (8) optimizing submitted metadata (including titles and dataset descriptions); (9) classifying content according to industry standards; (10) listing authors and corresponding papers; (11) generating DOIs for each dataset; (12) configuring anonymous dataset access; (13) linking and synchronizing data with related papers; and (14) providing evaluation reports and data editor feedback. For datasets exceeding 50GB, additional storage fees apply.

**** Paid Service Inquiry Form [27]

- Is the data related to a manuscript (research paper, article, chapter, or monograph)?
- Is the related manuscript published?
- Agree to be contacted by data support service
- Agree to be contacted by data support service and receive Springer Nature marketing communications

**** Advantages of Springer Nature's Paid Data Services

For Researchers	For Research	For Research Field
• Professional management by experienced data editors ensures data is easily found and reused	• Publishing research data linked to papers enhances research content	• Publishing research data drives change, inspires and promotes scientific innovation
• Saves time on storing, sharing, and managing data; organizes and presents scientifically • Increases team exposure in research field • More easily meets data sharing requirements of publishers, institutions, and funders • Positions data-sharing researchers as pioneers of open science	• Ensures paper-related data is available and findable, increasing article visibility	• Makes data more accessible and discoverable, improving research transparency and reproducibility

The service costs €300/£265/$340 (excluding VAT and local taxes) and has already helped publish over a hundred datasets across Springer Nature journals including *Nature* and *BMC Ecology*.

In fact, academic publishers providing research data management support represents a new development in research data management services. Previously, researchers and journal editors typically accessed such services through institutional research data management teams, research funders, professional data curation organizations like Jisc or the UK's Digital Curation Centre (DCC), or directly through specialized data repositories. However, not all researchers can access services through these channels. Future research data management services will certainly adopt a hybrid model combining services from different stakeholder groups, with no single optimal solution. Publisher-provided research data management services serve as valuable supplements to institutional services, particularly when institutions cannot expand or provide sufficiently flexible services due to personnel or resource constraints.

4. Application Status of Springer Nature's Policy Standardization Framework and Its Implications

4.1 Application Status

The policy framework was introduced to Springer Nature journals in Q2 2016, including *Nature*, BioMed Central journals, and journals hosted on SpringerLink [31]. As of September 2019, over 1,600 journals in the Springer Nature publishing group had developed their own data policies based on the framework. Manual classification of these journals by discipline revealed that natural science journals accounted for the highest proportion at 71.5%, with life science journals predominantly adopting Types 2 and 3. Interdisciplinary journals ranked second at 15.5%, while social science journals were the fewest at only 13%. In the social sciences, as of September 2019, Type 1 was most common (72.4%), followed by Type 2 (21.7%), Type 3 (4.5%), and 1.4% of social science journals remained undecided with policies still pending selection.

Research shows that beyond disciplinary factors, other considerations affect journal policy type selection. For example, journals with insufficient editorial staff to enforce policy compliance may opt for Type 1 or 2, as may those previously lacking data sharing policies. Type 4, which requires open data for every publication, is currently the least adopted policy type, having only been implemented by data journals and/or author communities with well-established open science cultures and open data sharing repositories. As research data sharing policy stringency increases, so do publication process costs. Types 3 and 4 require additional checks of submitted or accepted manuscripts, potentially generating costs that journals and publishers must understand.

4.2 Implications of Springer Nature's Policy Standardization Efforts

4.2.1 Government Agencies Should Lead in Issuing Research Data Policy Frameworks In scientific data open sharing systems, publishers are key stakeholders. Publishers collaborate with funding agencies, data centers, repositories, and research institutions to promote a culture of research data open sharing. Although Chinese publishing institutions have recognized the importance of research data open sharing, they started later than foreign publishers. Both journals and authors have insufficient awareness of the benefits of sharing research data, and publishers' participation in open research data sharing practices has not yet reached scale. Rapidly developing effective, actionable research data policies with clear specifications and recommendations for key processes like data submission, review, storage, and citation, along with clear supervision mechanisms, remains challenging. Therefore, relevant government departments should leverage China's institutional advantages to promptly establish a legal system for research data sharing that clearly defines intellectual property rights and legal consequences for malicious data use. Simultaneously, it is recommended that the China Association for Science and Technology, as the journal management organization, take the lead in convening stakeholder groups in research data sharing to discuss, formulate, and launch a policy framework for scientific journals that guides and helps journals develop effective, personalized data policies that align with international best practices.

4.2.2 Learn from International Experience to Construct Scientific and Reasonable Policy Frameworks National innovation capability is rooted in knowledge creation, aggregation, and dissemination, as well as its ecological environment. Scientific journals inherit human civilization, assemble scientific discoveries, and lead scientific development, representing to some extent a country's international scientific competitiveness and cultural soft power [32]. The November 14, 2018 meeting of the Central Committee for Comprehensively Deepening Reform reviewed and approved the "Opinions on Deepening Reform to Cultivate World-Class Scientific and Technological Journals." Using internationally common languages is an unavoidable important path to building world-class journals. With growing calls for research data sharing and accessibility, an increasing number of international journals are emphasizing research data deposition and publication, requiring authors to disclose relevant raw data when publishing papers [33]. As a major journal publisher, China must follow academic communication development trends, strengthen top-level design in scientific journal construction, leverage policy guidance, and promote the construction of an open science ecosystem. Compared with foreign countries, China has fewer journals with data policies, and journal data policy research has not attracted sufficient attention from relevant institutions and scholars. Therefore, Chinese scientific journal management agencies should actively learn from international experiences and practices. When defining the roles, responsibilities, and obligations of stakeholder groups in research data sharing, they can consider drawing on Elsevier's three-dimensional approach covering data pro-

ducer (researcher) requirements, stakeholder responsibilities, and data citation obligations to refine data ownership rights and clarify obligations. They can also learn from Springer Nature's nine common elements for research data policies, widely consult Chinese scientific journals, librarians, and research funders, and combine disciplinary differences to construct a universal scientific journal policy framework containing several types that guides domestic journals in improving existing data policies and promotes comprehensive, systematic development of journal data policies to enhance their utilization efficiency and scientific rationality.

4.2.3 Design Supporting Services to Effectively Promote Policy Framework Application In the context of "building world-class scientific and technological journals," both scientific and social science journals should pay attention to the latest publishing technologies, tools, and models, actively experiment with them, and strive to align with international journal development trends to improve efficiency and dissemination effectiveness [34]. For both editorial offices and researchers, research data publication is an emerging area where current cognitive levels, knowledge reserves, and practical operations fall short of desired best practices. Therefore, supporting services must be designed to effectively promote policy framework application. The researcher-facing and editorial office-facing supporting services simultaneously launched by Springer Nature provide excellent examples. However, it must be noted that China lags significantly behind foreign countries in the number of both public comprehensive data repositories and disciplinary data repositories. If journals launch research data policies, they will inevitably face substantial data maintenance costs and burdens, causing many journals to hesitate. Therefore, government agencies must lead in promoting research data infrastructure construction. It should be noted that research data infrastructure construction is a systematic project whose service capacity depends not on the scale, performance, and metrics of individual projects or systems, but on a scientific long-term operational mechanism [35]. The goal is not to build a super data storage center but to construct a data infrastructure that enables data sharing, promotes researcher collaboration and exchange, accelerates innovation, and provides momentum for development [36-37]. Only by developing supporting services based on research data infrastructure can services truly be promoted, their value and significance demonstrated, and their continuous deepening and improvement achieved, thereby facilitating framework application, promoting policy formulation, improvement, and effective implementation, and achieving the desired outcomes of open research data.

References

[1] SCHMIDT B, GEMEHOLZER B, TRELOAR A. Open data in global environmental research: the Belmont Forum's open data survey[J]. PLOS ONE, 2016, 11(1): e0146695.

[2] CENTER for OPEN SCIENCE. Transparency and openness promotion guidelines[EB/OL]. [2020-03-10]. https://cos.io/top/.

[3] JONES L, GRANT R, HRYNASZKIEWICZ I. Implementing publisher policies that inform, support and encourage authors to share data: two case studies[EB/OL]. [2020-03-10]. https://insights.uksg.org/articles/10.1629/uksg.463/.

[4] NAUGHTON L, KERNOHAN D. Making sense of journal research data policies[J]. Insights the UKSG journal, 2016, 29(1): 84-92.

[5] TAICHMAN DB, SAHNI P, PINBORG A, et al. Data sharing statements for clinical trials[EB/OL]. [2020-03-10]. https://www.bmj.com/content/bmj/357/bmj.j2372.full.pdf.

[6] COLAVIZZA G, HRYNASZKIEWICZ I, STADEN I, et al. The citation advantage of linking publications to research data[EB/OL]. [2020-03-10]. https://arxiv.org/abs/1907.02565.

[7] UK Research and Innovation. Common principles on data policy[EB/OL]. [2020-03-10]. https://www.ukri.org/funding/information-for-award-holders/data-policy/common-principles-on-data-policy/.

[8] COLAVIZZA G, HRYNASZKIEWICZ I, STADEN I, et al. The citation advantage of linking publications to research data[EB/OL]. [2020-03-10]. https://arxiv.xilesou.top/pdf/1907.02565.pdf).4.

[9] GIOFRE D, CUMMING G, FRESCH L, et al. The influence of journal submission guidelines on authors' reporting of statistics and use of open research practices[J]. PLOS ONE, 2017(12): e0175583.

[10] VASILEVSKY NA, MINNIER J, HAENDEL MA, et al. Reproducible and reusable research: are journal data sharing policies meeting the mark?[J]. PeerJ, 2017(5): e3208.

[11] HRYNASZKIEWICZ I, BIRUKOU A, ASTELL M, et al. Standardising and harmonising research data policy in scholarly publishing[EB/OL]. [2020-03-10]. https://ijdc.net/index.php/ijdc/article/download/12.1.65/467.

[12] HRYNASZKIEWICZ I, SIMONS N, HUSSAIN A, et al. Developing a research data policy framework for all journals and publishers[EB/OL]. [2020-03-10]. https://datascience.codata.org/article/10.5334/dsj-2020-005/.

[13] 王丹丹. 英国《开放科研数据协议》的解读与启示 [J]. 数字图书馆论坛, 2018(8): 25-30.

[14] SPRINGER NATURE. Research data policy types[EB/OL]. [2019-12-17]. https://www.springernature.com/de/authors/research-data-policy/research-data-policy-types.

[15] SPRINGER NATURE. Helpdesk[EB/OL]. [2019-12-17]. http://www.springernature.com/gp/group/data-policy/helpdesk.

[16] ASTELL M, HRYNASZKIEWICZ I, GRANT R, et al. Have questions about research data? Ask the Springer Nature Helpdesk[EB/OL]. [2020-03-10]. https://figshare.com/articles/Providing_{{{advice}}}*{{{and}}}*{{{guidance}}}*{{{on}}}*{{{research}}}*{{{data}}}*

[17] SPRINGER NATURE. Recommended repositories[EB/OL]. [2020-03-10]. http://www.springernature.com/gp/group/data-policy/repositories.

[18] SPRINGER NATURE. Data availability statements[EB/OL]. [2020-03-10]. https://www.springernature.com/gp/group/data-policy/data-availability-statements.

[19] NATURE. Announcement-where are the data?[EB/OL]. [2020-03-10]. http://www.nature.com/news/announcement-where-are-the-data-1.20541.

[20] GRANT R, HRYNASZKIEWICZ I. The impact on authors and editors of introducing data availability statements at Nature journals[EB/OL]. [2020-03-10]. http://www.ijdc.net/article/view/614.

[21] SPRINGER NATURE. Research data support service[EB/OL]. [2019-12-17]. https://www.springernature.com/de/authors/research-data/research-data-support.

[22] SPRINGER NATURE. Research data at springernature[EB/OL]. [2019-12-17]. https://researchdata.springernature.com/?utm_{source}=springernature.com&utm_{campaign}=CMTI

[23] STUART D, BAYNES G, HRYNASZKIEWICZ I, et al. Practical challenges for researchers in data sharing[EB/OL]. [2020-03-10]. https://figshare.com/articles/Whitepaper_{{{Practical}}*{{challenges}}*}{{{for}}*{{researchers}}*}{{{in}}*{{a

[24] MARKET RESEARCH. Wiley open science researcher survey 2016[EB/OL]. [2020-03-10]. https://figshare.com/articles/Wiley_{{{Open}}*{{Science}}*}{{{Researcher}}*{{Survey}}*}{2016

[25] BERGHMANS S, COUSIGN H, DEAKIN G, et al. Open data: the researcher perspective-survey and case studies[EB/OL]. [2020-03-10]. https://data.mendeley.com/datasets/bwrn-fb4bvh/1.

[26] HRYNASZKIEWICZ I. New services to support open research[EB/OL]. [2020-03-10]. https://blogs.biomedcentral.com/bmcblog/2017/04/27/new-services-to-support-open-research/.

[27] SPRINGER NATURE. Research Data Support service[EB/OL]. [2019-12-17]. https://www.springernature.com/gp/authors/research-data/research-data-support.

[28] SPRINGER NATURE. Types of research data we can help to organise and share[EB/OL]. [2019-12-17]. https://media.springernature.com/full/springer-cms/rest/v1/content/15541262/data/v1.

[29] SPRINGER NATURE. Organise and share your data faster with research data support[EB/OL]. [2019-12-17]. https://media.springernature.com/full/springer-cms/rest/v1/content/15403688/data/v3.

[30] SPRINGER NATURE. Please read these terms and conditions carefully before using the site and ordering the services[EB/OL]. [2019-12-17]. https://media.springernature.com/full/springer-cms/rest/v1/content/15403690/data/v3.

[31] HRYNASZKIEWICZ I. Promoting research data sharing at springernature[EB/OL]. [2020-03-10]. http://blogs.nature.com/ofschemesandmemes/2016/07/05/promoting-research-data-sharing-at-springer-nature/.

[32] 中国科协中宣部教育部科技部关于深化改革培育世界一流科技期刊的意见 [J]. 分析化学, 2019, 47(9): 1301, 1365, 1381, 1426.

[33] 傅天珍, 陈妙贞. 我国学术期刊数据出版政策分析及建议 [J]. 中国出版, 2014(23): 31-34.

[34] 武晓耕, 胡小洋, 占莉娟, 等. “建设世界一流科技期刊”背景下的社科期刊发展策略探究 [J]. 中国科技期刊研究, 2019, 30(8): 801-805.

[35] 章昌平, 米加宁, 黄欣卓. 超越数据洪流: 第四研究范式下的社会科学研究数据基础设施 [J]. 学海, 2019(3): 55-61.

[36] 王敬, 王彦兵. 国外科研数据基础设施研究及实践的调研与分析 [J]. 情报资料工作, 2016(6): 99-104.

[37] 马骏. 科学数据管理服务推进过程中的合作策略 [J]. 情报资料工作, 2018(05): 62-67.

Author Contributions

Wang Dandan: Conceived the research idea, wrote and revised the final manuscript;
Liu Qinghua: Wrote and revised the manuscript;
Ge Liyun: Assisted with data collection and organization.

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