



The Creative Commons-Attribution-Noncommercial License 4.0 International applies to all works published by IASSIST Quarterly. Authors will retain copyright of the work and full publishing rights.

Understanding motivations and future needs for data deposits at Korea Social Science Data Archive

Hyowon Kim¹, Do Won Kim², Jungwon Yang³

Abstract

Korea Social Science Data Archive ([KOSSDA](#)) has been integral in archiving and disseminating social science data in South Korea. As it transitioned into an independent research center under Seoul National University, we evaluated current data deposit process of KOSSDA through in-depth interviews with institutional data donors and identified future needs of data depositors. Motivations behind data deposit included recognizing data as a public good, addressing resource shortages in data curation, and enhancing data promotion. Future needs for KOSSDA included developing qualitative data guidelines, strengthening collaborations with donor institutions and academia as a whole, and automating the deposit process. These results guide the long-term strategy of KOSSDA to remain a leader in social science data archiving, ensuring its alignment with the evolving needs of the academic community.

Keywords

Korea Social Science Data Archive, KOSSDA, Social Science Data, Data Deposition

Introduction

The Korea Social Science Data Archive (KOSSDA) has been a pioneer in archiving and sharing social science data since 2006 in South Korea. It has gathered various quantitative and qualitative data that were scattered across individual researchers or research institutions, and made these data publicly available. KOSSDA was designed in accordance with the Open Archival Information System (OAIS) reference model, which is the international standard model. Their data indexing and metadata creation follow the Metadata Encoding and Transmission Standards (METS) and Data Documentation Initiative (DDI) 2.5. This initiative has not only facilitated data reuse by the academic community including researchers and students, but also opened data access to the general public. In 2022, KOSSDA transitioned from being a part of the Asia Center of Seoul National University (SNU) to becoming an independent research center under the Social Science College of SNU. This significant transition highlights the urgent need to develop a long-term strategy for KOSSDA, positioning it as a leading independent research institution at the forefront of social science data.

The importance of establishing long-term strategies is further emphasized by the situations KOSSDA currently faces. Notably, there is a lack of national and institutional foundation for the archiving and management of social science data in South Korea. Even for research funded by national resources, there are no national laws mandating the submission of data management plans. In late 2023, a bill was proposed to establish the 'Act on the Management and Utilization Promotion of National Research Data'. This legislation aimed to promote research data sharing and build an open research

ecosystem, including mandating the submission of Data Management Plans (DMPs). However, this bill was discarded due to the expiration of the legislative term⁴. This legislative setback illustrates the ongoing challenges in establishing systematic management and sharing of research data in South Korea, particularly in the social sciences. Consequently, research data in social sciences continue to lack systematic management and sharing protocols in the country.

In this context of insufficient interest and institutional support by the government, data management and sharing are more dependent on the will and capacity of the research community than on governmental initiatives or regulations. On the other hand, this situation presents an opportunity for KOSSDA to take a leading role, especially at a time when individual researchers and institutions need a focal actor for developing norms about data management, sharing, and reuse.

Moreover, while institutional depositors are primary data contributors to KOSSDA, they have also begun establishing their own data archive agencies. There has been a growing interest in self-archiving within each institution, with an increased emphasis on the transparency, reproducibility, and usability of research results, which would give these institutional depositors direct control and autonomous management of their data. Although KOSSDA has been a part of this open science movement, it is crucial to thoroughly analyze whether its future relationships with these data depositor institutions will lean more towards competition or collaboration.

To summarize, as a crucial step towards establishing KOSSDA as a leading independent research institution in the field of social science data archives, this study aims to review the current data deposit process and identify the future needs of depositors. To this end, this study involved conducting one-on-one, in-depth interviews with individual representatives of institutional data donors. These interviews help us understand the motivations behind data donations, evaluate KOSSDA's strategies for data archiving and curation, and comprehend future expectations of data depositors that KOSSDA needs to meet.

Background

A significant difference between the current situation in South Korea and that of other developed countries is the role of data sharing requirements from national and funding agencies. Countries such as the United States and the United Kingdom are actively pursuing policies to promote the sharing and utilization of research outcomes that have received public funding (Boehm et al., 2023; Kvale & Pharo, 2021; Smale et al., 2020; Williams et al., 2017; Whyte & Tedds, 2011).

In the United States, all federal agencies spending over \$100 million annually on R&D are required to submit a Public Access Plan for publicly funded research outputs (Choi & Lee, 2020). Notably, the National Institutes of Health (NIH) has mandated the submission of a Data Sharing Plan for research projects exceeding \$500,000 since 2003. Similarly, the National Science Foundation (NSF) has required the inclusion of a Data Management Plan (DMP) for all research projects since January 2011 (Choi & Lee, 2020). These policies view data sharing as an "economy driver" (Smale et al., 2020), ensuring that "full value would be realized from the collected data" (Smale et al., 2020). This approach enables follow-up studies and allows for the validation of other research through data reuse. Under such institutional support and attention, researchers find that sharing data creates opportunities for collaboration and recognition of their work as data producers (Boehm et al., 2023; Kvale & Pharo, 2021; Smale et al., 2020; Williams et al., 2017; Whyte & Tedds, 2011).

However, in the South Korean context, norms for managing and sharing research data are still not well-established. The field of social sciences in Korea receives less institutional attention and support compared to scientific and technological fields. Yet, even in the scientific and technological fields that receive relatively more national attention regarding research data management and sharing, the situation is far from ideal. According to the Ministry of Science and the National Research Council of Science & Technology (NST), among four major science and technology institutes in South Korea, only one has established an online repository for storing R&D research data (Cho, 2023). Moreover, none of these institutes have linked their repositories to DataOn, the national research data archive for science and technology fields (Cho, 2023).

The South Korean context is particularly important given the lack of national interest and institutional support, especially in non-scientific fields. Research data management is not prioritized, and data sharing is not mandatory. This situation differs from Europe and the United States, where data sharing has been imposed onto the research community by external forces (Smale et al., 2020).

In this context, the research community in South Korea must take the lead. Social science archive institutions like KOSSDA have taken the initiative to drive the open science movement within the country. KOSSDA work to convince other social science research institutions to deposit their data, which KOSSDA then quality-controls and makes publicly available. This proactive approach positions KOSSDA as a key player in promoting open science practices in the South Korean social science community.

However, research institutions are beginning to establish their own archives for the data they produce. This trend of institutions sharing data through their own proprietary archives raises important questions about how to foster future collaboration with these institutions. In light of this evolving landscape, this study aimed to gather opinions and perspectives from data depositors on how data deposits can be sustained and expanded within this changing relationship.

Method

To evaluate data deposit process of KOSSDA and identify any future needs of data depositors, we conducted in-depth individual interviews with representatives from eight institutions from August 25 to September 2, 2022.

We specifically targeted representatives responsible for data deposits from each institution. The selection was based on institutions that had deposited datasets with the highest download rates within their respective categories. The institutions spanned various organizations, including: four national/public research institutions, one regional research institution, one private research institution, one university research institution, and one national agency.

We prepared interview questions and scripts to gather information on motivations of each institution for data deposits, data sharing and archiving status of each institution, strengths and weaknesses of the current data deposit process with KOSSDA, and expectations about KOSSDA's contribution related to research data management services. The main questions are presented in Table 1. Yet, for in-depth exploration of perspectives of data donors, our interviewers were guided to use a semi-structured interview method while they led a conversation with interviewees. Each interview lasted between 60-90 minutes.

Table 1. Key Questions on Institutional Data Deposition Experience and Expectations for KOSSDA

1. Could you describe the current state of data management practices within your research institution?
2. What are the key motivations driving your institution to deposit research data?
3. What is your personal perspective on data deposition?
4. Since beginning to deposit data, what positive outcomes or benefits has your institution observed?
5. During the data deposition process, what challenges or obstacles does your institution typically encounter?
6. Regarding the data deposition process, what improvements or changes would your institution like to see implemented by KOSSDA?

We assigned one interviewer who would lead the questions to an interviewee and one note taker who would take detailed notes for each interview session. After each interview, we held a debriefing session with the interviewers.

Results

The interviews revealed underlying motivations and incentives for data deposits at both the institutional and individual levels, which were closely related to the data management situation at each institution (see Table 2; the number of participants who mentioned each point is shown in parentheses, along with the corresponding percentage). Specifically, these incentives can be categorized as follows: (1) the norm that research data is a public good, (2) the lack of data-curating resources in each institution, and (3) the desire to promote data.

[1] The recognition and norm that research data is a public good

The ethos that research data is a public good was particularly strong across national or public research institutions. Within these institutions, it is well-acknowledged that data produced by government organizations belongs to South Korean citizens and that the public has the right to be informed about how tax money has been used.

Some institutions are developing or have already developed their own platforms for archiving their research data. However, they do not view KOSSDA as a competitor. Instead, since research institutions acknowledge that research data is a public good, they aim to foster its preservation, sharing, and reuse through multiple channels. In other words, the data archive platform provided by KOSSDA and that developed and managed by each institution are complementary rather than substitutes.

[2] Lack of data curating resources in each organization

While there is widespread agreement within data donor institutions that data is a public good, the necessary infrastructure and resources to uphold this principle are lacking. This lack of specialized resources for data curation has incentivized these institutions to rely on KOSSDA for data curation.

For example, there is a scarcity of researchers who have received professional training in data archiving and curation, and institutional researchers are typically tasked with other responsibilities and administrative workloads, leading to the de-prioritization of data archiving and curation tasks. One of the interviewees had even experienced data loss in their institutions due to insufficient data management resources. Therefore, some institutions address their lack of data curation resources by depositing data with KOSSDA, effectively outsourcing this aspect of data management.

At the same time, a lack of data curation resources may result in fewer data deposits to KOSSDA, rather than encouraging researchers to deposit more. Some research institutions lacked the resources to invest in the data review process, which is required for data deposits to KOSSDA. As a result, reducing or facilitating these required steps would reinforce motivations of institutions to deposit data to KOSSDA.

Table 2. Incentive of data curation

<p>The recognition and norm that research data is a public good</p>	<ul style="list-style-type: none"> - Acknowledgment that government-produced data belongs to citizens (6/8, 75%) - KOSSDA seen as complementary, not competitive in data archiving issues; fostering preservation, sharing, and reuse through multiple channels (8/8, 100%)
<p>Lack of data curating resources in each organization</p>	<ul style="list-style-type: none"> - Insufficient infrastructure and resources despite agreement on data as a public good (5/8, 62.5%) - Scarcity of professionally trained researchers in data archiving and curation (6/8, 75%) - Researchers overwhelmed with their primary research duties; data management and sharing often deprioritized (4/8, 50%) - Researchers experienced instances of data loss due to insufficient management resources (1/8, 12.5%) - Some institutions outsource data management to KOSSDA (5/8, 62.5%) - Lack of resources may also lead to fewer data deposits (5/8, 62.5%)
<p>Data promotion</p>	<ul style="list-style-type: none"> - KOSSDA highly regarded as a social science data archive (4/8, 50%) - Perceived as a channel to promote data produced by research institutions (5/8, 62.5%) - Reported increase in data usage after depositing with KOSSDA (3/8, 37.5%) - Observed diversification in fields of study using data produced by research institutions. For example, data previously used mainly in education is now being utilized in sociology and public administration (2/8, 25%)

[3] Data promotion

Research institutions highly value KOSSDA as a social science data archive, and view it as a valuable channel for promoting their data. In particular, they noted an increase in the use of their data after it was deposited in KOSSDA. Although a clear causal relationship cannot be established, researchers frequently observe a diversification in the fields of study among authors who publish papers based on their data, when comparing before and after data deposition.

For instance, an interviewee mentioned that while their data was primarily used in the field of education in the past, after depositing their data with KOSSDA, they noticed that their data is being used in many other social science domains, including sociology and public administration.

Future needs

Through the interviews, we also identified future needs that both institutions and individuals expect from KOSSDA (see Table 3; the number of participants who mentioned each point is shown in parentheses, along with the corresponding percentage). These can be summarized as follows: (1) setting up guidelines for qualitative data, (2) strengthening collaboration - including spreading citation standards and enhancing academic collaboration between KOSSDA and donor institutions, and (3) automating the data deposit process.

Table 3. Future needs

Establishing guidelines for qualitative data	<ul style="list-style-type: none">- Develop guidelines for handling and disclosing personal information (5/8, 62.5%)- Create comprehensive guidelines for collecting, managing, and sharing qualitative data (3/8, 37.5%)- Implement educational programs for data curators at donor institutions (3/8, 37.5%)
Strengthening collaboration between KOSSDA and the academic community	<p>Spreading Citation Standards:</p> <ul style="list-style-type: none">- Collaborate with journal publishers to establish and spread data citation standards (3/8, 37.5%)- Promote wider adoption of data citation standards in social science journals and theses (3/8, 37.5%) <p>Enhancing Academic Collaboration:</p> <ul style="list-style-type: none">- Integrate KOSSDA's data into methodology training programs (2/8, 25%)- Launch joint conferences or journals to promote data reuse (3/8, 37.5%)- Utilize KOSSDA's seasonal methodology courses for data promotion (2/8, 25%)- Expand KOSSDA's award program for theses using its data collections (2/8, 25%)
Automation of the data deposit process	<ul style="list-style-type: none">- Implement an automated system for easier and more secure data handling (3/8, 37.5%)- Provide easy access to usage statistics for donated data (5/8, 62.5%)

[1] Establishing guidelines for qualitative data

The most prominent demand is evident in dealing with qualitative data generated in social science research. At the institutional level, there is a notable lack of common norms or guidelines for managing qualitative data, particularly concerning ethical issues such as the handling and disclosing personal information. This absence of guidelines not only complicates the internal management of qualitative data but also makes depositing this data challenging. The lack of clear standards across different research institutions leads to confusion among researchers.

Consequently, research institutions and data management representatives often receive inquiries from academic researchers on the use of qualitative data, but they are not equipped with formal training or established protocols to assist them. There is a strong need to create comprehensive guidelines and regulations for collecting, managing, sharing and deposit of qualitative data.

Furthermore, there is a recognized need for an educational program for data curators at donor institutions.

[2] Strengthening collaboration between KOSSDA and the academic community

[2-1] Spreading Citation Standards

KOSSDA currently distributes data citation guidelines⁵ to various users, including data depositors, graduate students, and researchers. However, data donor institutions expressed a desire for these citation norms to be more firmly established and implemented. Particularly, for the data citation norms to be spread and established across different social science fields, KOSSDA should collaborate with not only data depositors but also journal publishers.

In the South Korean context, academic journals lack such guidelines. As of January 25, 2024, prominent journals published by organizations like the [Korean Political Science Association](#) or the [Korean Sociological Association](#) do not include any data citation guidelines in their submission instructions. This disparity highlights the need for wider adoption of data citation standards within academic societies, especially in journals related to social science, as well as in graduate theses and dissertations. As a comparison, in the United States, many academic journals have well-established guidelines for data citation. Currently, there are ongoing efforts to develop citation guidelines for various components, including datasets, measurement instruments, and software in South Korea.

[2-2] Enhancing Academic Collaboration between KOSSDA and donor institutions

There is a need for enhancing academic collaboration between KOSSDA and donor institutions. For example, some interviewees suggested using deposited data in coding sessions of [KOSSDA's methodology training programs](#), and collaborating with KOSSDA to initiate joint conferences or journals, aiming to foster the culture of data reuse. All these specific demands are aimed at fostering cooperation to boost data reuse.

Since 2007, KOSSDA has been at the forefront of promoting social science methodology through its specialized methodology courses conducted across all seasons - Spring, Summer, Fall, and Winter. Each season offers between six to nine classes that cover a wide array of research methodologies, both quantitative and qualitative. Particularly, the quantitative research courses are popular, often attracting more than fifty participants per course. With attendees from varied backgrounds, including students, researchers, and professors, data depositors believed that these courses could serve as an excellent platform to promote their data. For instance, instructors can be encouraged to utilize deposited data in their lab sessions during the methodology training programs. They can teach attendees how to access and download data from KOSSDA's website and analyze it during the lab sessions. Hence, the data is not only reused for educational purposes but also has the potential to lead to further research opportunities for junior scholars and students.

Furthermore, KOSSDA is exploring ways to enhance its contribution to data reuse by establishing joint conferences or journals. Currently, KOSSDA awards outstanding graduation theses that use its data collections. However, there have been suggestions that this award program should further expand its scope and impact further.

[3] Automation of the data deposit process

Lastly, there is a strong desire for improvements in the data deposit process to enhance its efficiency and security. At the time of the interview, the data deposit procedure was manual and conducted via

email. Coupled with the lack of specialized resources for data curation within institutions, there is significant interest in automating this system to enable easier and more secure handling. Additionally, some institutional representatives wished for easy access to usage statistics for donated data, instead of having to submit separate requests to KOSSDA for this information.

Recommendations

These interviews with representatives of institutions that deposit data in the archive have revealed several key areas where KOSSDA, as a social science data archive institution in South Korea, can focus its efforts for long-term development. These areas include:

[1] Strengthening Collaboration with Academic Journals

The interviews reaffirmed the gaps in the Korean social science field, particularly regarding data citation practices. Moving forward, KOSSDA's long-term development strategy should encompass collaboration not only with research institutions that deposit data to KOSSDA but also with academic journals. This approach differs significantly from more advanced academic circles in Europe and the United States, where data citation has already become an established norm.

In these advanced academic environments, citing data in publications when using data produced by others has become a common practice. This shift occurred following the introduction of Digital Object Identifiers (DOIs) for datasets by initiatives like DataCite, and the establishment of the FAIR (Findable, Accessible, Interoperable, Reusable) principles as standard in the mid-2010s (Silvello, 2018). The discussion in these regions has since evolved beyond mere data citation, now exploring comprehensive ways to acknowledge the contributions of data curation and archiving institutions (Buneman et al., 2020).

However, in Korea, even government agencies and higher institutions such as the National Research Foundation, which manage publicly funded research, still lack mandatory and specific requirements for research data management and sharing in social science. Consequently, the practice of citing open data used in research has not yet been firmly established. To establish such norms within the research community and academia, close collaboration with journals would be most effective.

Specifically, KOSSDA should collaborate with major social science journals to incorporate data citation guidelines into their submission requirements. This concrete step would help institutionalize data citation practices within the Korean academic community.

[2] Engaging Individual Researchers

While the current interviews focused on institutional representatives, there is a need to study individual researchers' research data management, sharing, and reuse practices to expand research data sharing and reuse. Moreover, in an environment where Data Management Plans (DMPs) are not standard practice when receiving research funding, KOSSDA should: a) Conduct surveys or interviews with individual researchers to understand their data sharing practices and motivations. b) Develop strategies to encourage data sharing among individual researchers, considering the complex licensing issues that may arise when research is funded by universities, corporations, or national research foundations. c) Provide education and resources to individual researchers on the importance and benefits of data sharing.

[3] Practical Implications for Donor-Friendly Processes

Simultaneously, KOSSDA needs to improve its deposition process and provide better services to data donors. Based on the demand for automation of the data deposit process, the following improvements are suggested: a) Implement an automated, user-friendly, and secure system for data deposition, replacing the current manual, email-based process. b) Leverage DOIs to provide regular analyses of citation counts and usage statistics for deposited data. This information should be easily

accessible to data donors, eliminating the need for separate requests. c) Establish a structured timeline for data deposition requests. Currently, the timing of data deposition is left to the discretion of the donating institutions. Regular reminders and structured timelines can help ensure timely and consistent data deposits.

[4] Expansion of Educational Offerings

While KOSSDA currently provides education in qualitative and quantitative social science research methodologies, there is an opportunity to expand its educational scope in the areas of data curation and management. This expansion is particularly crucial as researchers often lack awareness and skills in data management. It is imperative to develop educational programs tailored to the diverse needs of researchers at different career stages and roles within research teams. These programs should cater to various groups including undergraduate and graduate students, early-career researchers, established researchers leading projects, and professionals responsible for research data management. Such targeted education would address the importance of data management and provide the necessary skills throughout the researcher's lifecycle and within the context of their specific roles in research teams. By broadening its educational offerings, KOSSDA can play a pivotal role in enhancing the data management capabilities of the entire research ecosystem, thereby fostering a culture of effective data curation and sharing within the academic community.

Conclusion

KOSSDA faces the dual challenges of establishing itself as a leader in the data archive field while navigating the evolving landscape of data management and sharing. Through interviews with data deposit institution representatives, it has been established that these discussions provide practical evidence for shaping a specific and actionable long-term development agenda.

Interviews revealed a shared recognition of the importance of data as a public good, the need for improved data curation, and a desire to enhance data visibility and prevent data loss. These insights underscore a clear demand for KOSSDA to establish robust guidelines for qualitative data, strengthen collaborative efforts, especially in creating data citation norms with major journals, and, from a practical standpoint, automate the data deposit process. Future needs encompass not only technical and procedural improvements but also a strategic focus on advocacy and education to promote the value of data sharing and management.

These future needs collectively position KOSSDA not as a competitor to research institutions in data archiving, but as an integral collaborator and partner within the academic community. This perspective informs the approach to implementing KOSSDA's long-term development plans. It emphasizes the importance of KOSSDA's role beyond its current functions as an archive institution, a platform for social science research methodology education, and a social science research institute.

Consequently, the research results confirm that the long-term development directions discussed internally at KOSSDA align with those of the academic research community, including researchers and institutions. In response to these future needs, KOSSDA has actively embarked on making improvements. Firstly, KOSSDA has initiated research focused on the development of qualitative research guidelines to ensure robust and consistent data curation practices. Secondly, in September 2023, KOSSDA introduced the [online material deposit system](#). This upgrade made the entire deposit process digital, including the completion of application forms, pre-review of data through a checklist, file upload, and discussions regarding data usage conditions and disclosure methods. KOSSDA also enhanced the security and stability of data file uploads and integrated new features for viewing online deposit details and data usage statistics.

Looking forward, KOSSDA must continue to adapt and innovate to meet the evolving needs of its depositors and the broader research community, ensuring that it remains at the forefront of data archiving and management in the field of social sciences.

References

- Boehm, R. I., Calkins, H., Condon, P. B., Petters, J., & Woodbrook, R. (2023). Analysis of U.S. Federal Funding Agency Data Sharing Policies: 2020 Highlights and Key Observations. *International Journal of Digital Curation*, 17(1). <https://doi.org/10.2218/ijdc.v17i1.791>
- Cho, S.-H. (2023). Science and Technology Institutes and Government-Funded Research Institutes Indifferent to National Research Project Data Management. *Yonhap News Agency*. <https://www.yna.co.kr/view/AKR20230927093800017?input=1195m> (Accessed: August 14, 2024).
- Choi, M.-S., & Lee, S. (2020). Current Status and Issues of Data Management Plan in Korea. *The Journal of the Korea Contents Association*, 20(6). <https://doi.org/10.9728/kca.2020.20.6.220>
- Kvale, L., & Pharo, N. (2021). Understanding the Data Management Plan as a Boundary Object through a Multi-stakeholder Perspective. *International Journal of Digital Curation*, 16(1), Article 1. <https://doi.org/10.2218/ijdc.v16i1.746>
- Silvello, G. (2018). Theory and practice of data citation. *Journal of the Association for Information Science and Technology*, 69(1), 6–20. <https://doi.org/10.1002/asi.23917>
- Smale, N. A., Unsworth, K., Denyer, G., Magatova, E., & Barr, D. (2020). A Review of the History, Advocacy and Efficacy of Data Management Plans. *International Journal of Digital Curation*, 15(1), Article 1. <https://doi.org/10.2218/ijdc.v15i1.525>
- Williams, M., Bagwell, J., & Nahm Zozus, M. (2017). Data Management Plans: The Missing Perspective. *Journal of Biomedical Informatics*, 71. <https://doi.org/10.1016/j.jbi.2017.05.004>
- Whyte, A., & Tedds, J. (2011). *Making the Case for Research Data Management*.

Endnotes

¹ PhD student in Information at University of Arizona. She can be reached by email:

hyowonkim@arizona.edu

² PhD student in Information Studies at University of Maryland. She can be reached by email:

dowonkim@umd.edu

³ International Government Information Librarian and Associate Director of Social Science and Clark Library Team, University of Michigan. She can be reached by email: yangjw@umich.edu

⁴ http://likms.assembly.go.kr/bill/billDetail.do?billId=PRC_A2A3Y0Z9X0Y6G1G5F2D8E4C5D3K5J5

⁵ Please see page 40 for the data citation guidelines.

https://kossda.snu.ac.kr/download?file=user_guide.pdf