



NASSDA as a Data- Collaborative Network to Enhance Social Science Research in East Asia

RESEARCH PAPER

PEI-SHAN LIAO

SHUAI WANG

HIROSHI ISHIDA

HEARAN KOO

WON-HO PARK

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

][ubiquity press

ABSTRACT

International organizations that promote data sharing and collaboration among data depositories have emerged globally and regionally. Such international collaboration, however, is few in regions other than Europe and North America. It is also important to share social science data, given the contributions of social science research. Although the advantages of sharing social science data has been recognized, challenges persist, particularly for data repositories focusing on curation and dissemination. In this article, we introduce the establishment of the Network of Asian Social Science Data Archives (NASSDA), a collaborative network among social science data depositories in East Asia. Its member archives include Chinese National Survey Data Archive (CNSDA), Korean Social Science Data Archive (KOSSDA), Survey Research Data Archive (SRDA), and Social Science Japan Data Archive (SSJDA). Each data archive describes its purposes for establishment, data usage and infrastructure, research and promotion activities, and future developments.

Since the establishment of NASSDA in 2016, numerous challenges have impeded progress, including differences in infrastructure and systems for survey data loading, financial and personnel situations, national policies, and domestic circumstances. The current focus of NASSDA is to collaborate within these existing constraints by establishing a website and organizing annual meetings. These initiatives serve as a platform for information exchange and idea sharing, aiming to align the future direction with those of other institutions. This fosters a mutual understanding and paves the way for further collaboration.

CORRESPONDING AUTHOR:

Pei-Shan Liao

Academia Sinica, TW

psliao@gate.sinica.edu.tw

KEYWORDS:

Network of Asian Social Science Data Archives; CNSDA; KOSSDA; SRDA; SSJDA

TO CITE THIS ARTICLE:

Liao, P-S, Wang, S, Ishida, H, Koo, H and Park, W-h 2025 NASSDA as a Data-Collaborative Network to Enhance Social Science Research in East Asia. *Data Science Journal*, 24: 3, pp. 1-14. DOI: <https://doi.org/10.5334/dsj-2025-003>

Support for data sharing is growing in both social and natural sciences and data depositories flourish in order to respond to open data initiatives. International organizations that promote data sharing and collaboration among data depositories have emerged globally and regionally, such as the World Data System (WDS), the Committee on Data of the International Science Council (CODATA), Consortium of European Social Science Data Archives (CESSDA), International Federation of Data Organizations (IFDO), and the Inter-university Consortium for Political and Social Research (ICPSR), just to name a few. Even with all these large-scale data organizations or repositories and open data policies implemented in the member states of the Organization for Economic Cooperation and Development (OECD), the sharing of research data has not been a common practice across scientific disciplines (Klump, 2017; Tedersoo et al., 2021).

In social sciences, which often involve human subjects, the sharing of research data is a complex task, due concern for participant protection and legal issues (Kirilova and Karcher, 2017; Tedersoo et al., 2021), as well as the various forms of data, such as survey data, qualitative interviews, videos, and social media data (Akdeniz et al., 2023). Although the advantages of sharing social science data has been recognized for over thirty years (Hedrick, 1988; Sieber, 1991), challenges persist, particularly for data repositories focusing on curation and dissemination. These challenges are often greater for repositories with limited collaborative support. A networked approach is expected to provide substantial benefits, while it may also present different challenges.

Most of the international collaboration mentioned above focused on Europe or global partnership, with limited engagement in other regions. While the emergence of internet has enabled cross-border data sharing, a wide variety in data-sharing culture and policies remains across countries in social sciences and humanities (Kvalheim and Kvamme, 2014). In this article, we introduce the establishment of the Network of Asian Social Science Data Archives (NASSDA), a collaborative network among social science data depositories in East Asia, and its member archives. We begin by discussing the importance of and concerns for sharing scientific data in social sciences, then describe collaboration among member archives, highlighting NASSDA's commitment to extending the network and advancing social science research in Asia and Pacific. Challenges and obstacles faced by NASSDA in sharing social science data are also addressed.

THE IMPORTANCE OF AND CONCERNS FOR SHARING SOCIAL SCIENCE DATA

Social science research deeply explores human behavior, societal structures, and cultural dynamics, examining interaction patterns, value systems, and behavioral models across diverse cultures. It also investigates organizational structures, power dynamics, communication patterns, and decision-making processes to understand internal operations. Using various types of data, social science research analyzes the impact of social systems, political trends, and economic influences on societies, focusing on social structures, policy implementations, economic developments, and cultural shifts. These aspects collectively form the core domains of social science research, and the research data facilitates a comprehensive understanding of human behavior, organizational operations, and societal development. Research data in social sciences form a cornerstone of scientific knowledge (Akdeniz et al., 2023), and the findings drawn from social science data help provide insights into human behaviors and aid policy-making process. The sharing of social science data has been encouraged to maximize its contribution to societies (Levenstein and Lyle, 2018; Sieber, 1991).

Data sharing advances scientific progress by enabling the replication of previous analyses and maximizing the return on investments in data collection (Levenstein and Lyle, 2018). It enhances openness, accountability, and scientific progress. More importantly, it allows researchers with limited resources to explore scientific questions, facilitates the replication of findings, and enables new analyses with advanced techniques. Additional benefits of data sharing include fostering diverse perspectives, reducing costs through data reuse, encouraging new research, increasing the number of publications, and preventing data fabrication, falsification, or distortion.

In addition to the above benefits, sharing social science data encourages interdisciplinary research and problem-solving by offering opportunities for understanding human societies from various aspects. Accessible social science data helps policymakers develop evidence-based policies, address societal challenges, and foster growth in one's society. It also encourages transparency and collaborative research efforts, and enhances the credibility of findings within the scientific community.

Despite the benefits, challenges, and obstacles of sharing research data in social sciences persist. Tedersoo *et al.* (2021) evaluated data availability in research articles and found that data availability was low in the field of social sciences was lower when compared to other scientific disciplines. The main reasons of declining data sharing included loss of data and concerns for privacy or legal issues. Social science data often include identifying information at individual or organizational level. Earlier research on the sharing of social science data has highlighted protection of the confidentiality and safety of research participants as a major concern among social scientist (Akdeniz *et al.*, 2023; Kirilova and Karcher, 2017). These concerns include privacy, ethics, and safety regarding the informed consent procedure and data access. Other individual concerns include loss of publication priority, perceive effort versus benefits, and potential data misuse (Kim and Adler, 2015).

Regardless of these individual barriers to sharing social science data, data archives and repositories have worked actively to provide necessary resources and support for researchers willing to share their data, aided by advancements in technology and the principles of open science. Data archives have been established in most Western countries, serving as crucial resources for empirical research and social sciences education. Since the 1960s, several social science data repositories have existed, for example the UK Data Archive, the Harvard-MIT Data Center, and the ICPSR at the University of Michigan. The US-based ICPSR has been an international consortium of more than 750 academic institutions and research organizations. Inter-university Consortium for Political and Social Research maintains a data archive of research files in the social and behavioral sciences and provides training in data access, curation, and methods of analysis for the research community.¹

Inter-university Consortium for Political and Social Research's leadership in the data science community is unparalleled. However, the data management and sharing landscapes vary significantly across regions and countries. In particular, the sharing of social science data in Asia has yet to become a firmly established research norm, and resources for data management and sharing are limited. Therefore, collaboration among regional data archives, grounded in the specific contexts and challenges of each region, is encouraged to foster open dialogue, share experiences and ideas, and stimulate future developments. To achieve this goal, the Network of Asian Social Science Data Archives (NASSDA) has been established. The member archives of NASSDA share a unified mission: promoting the secondary use of social science data within and across their respective regions. Like other cross-archive collaborations, this effort comes with both benefits and challenges, requiring strategic planning to achieve the shared goal.

NASSDA: BRIEF HISTORY AND DATA COLLABORATION

In contrast to the well-established data archives in the West, those in East Asia have historically lagged behind. However, starting in the mid-1990s, initiatives were launched to establish social science data archives in Taiwan and Japan. Subsequently, in the 2000s, South Korea and China also made efforts to develop data archives, leading to an increased momentum for social survey data archives in East Asia. With the ongoing progress in the development of data archives in East Asia, Japan (Social Science Japan Data Archive, SSJDA), South Korea (Korean Social Science Data Archive, KOSSDA), Taiwan (Survey Research Data Archive, SRDA), and China (Chinese National Survey Data Archive, CNSDA) launched the Network of Asian Social Science Data Archives (NASSDA) in 2016. Its purpose is to design an open discussion platform for active social scientists to share ideas regarding the future development of cooperative networks for data sharing and coordinated data production. These data archives also come together through annual international conferences to further enhance their collaborative efforts.

¹ Mission of ICPSR can be found on its website: <https://www.icpsr.umich.edu/web/pages/about/>.

The history leading to the establishment of NASSDA

The information sharing among data archives in East Asia began in the late 2000s. On March 14, 2008, a workshop titled ‘Data Archives and their Prospects in East Asia: Collaborative Data Collection, Dissemination, and Data Sharing’ was organized by the Institute of Social Science at the University of Tokyo. This workshop aimed to facilitate the sharing of information about the resources available for social science data archiving in East Asia, as well as ongoing collaborative data collection activities. Following the introduction of the data archiving activities in Europe, especially GESIS² initiatives, researchers from South Korea, Taiwan, and China engaged in enthusiastic discussions. This workshop marked the initiation of SSJDA’s intensified collaboration with the social science data archives of other East Asian countries and served as a catalyst for fostering these exchanges.

Since then, international workshops have been held with experts from Europe and the United States, allowing SSJDA, KOSSDA, and SRDA to regularly share information. For example, in December 2012, the Institute of Social Science at the University of Tokyo organized another workshop titled ‘Social Science Data Preservation and Dissemination in East Asia.’ Two years later, on March 7–8, 2014, the Asian Census team of the Asia Center at Seoul National University hosted the conference titled ‘Asian Social Survey and Data Sharing,’ with the aim of facilitating the exchange of ideas and experiences among Asian scholars and specialists in social surveys and data archiving. Furthermore, on February 13, 2015, the Center for Social Research and Data Archives (CSRDA) at the Institute of Social Science, University of Tokyo, invited experts from the Finnish Social Science Data Archive (FSD) to host an international conference titled ‘International Conference on Data Preservation and Dissemination in Tokyo, 2015.’

By organizing such conferences, data archives in East Asia were able to foster collaboration, share information, and establish platforms for discussions. They occasionally invited experts from Europe and North America, who contributed to the discussion. This continuous effort ultimately led to the establishment of NASSDA.

Establishment and activities of NASSDA

The Network of Asian Social Science Data Archive was initiated in January 2016, and by 2024, a total of eight meetings had been held. The Network of Asian Social Science Data Archive originated from the ‘Conference on Data Sharing and Beyond among East Asian Countries’, organized by the Korea Social Science Data Archive at Seoul National University Asia Center in January 2016. During a business meeting, KOSSDA, SSJDA, CNSDA, and SRDA agreed to form NASSDA and decided to hold annual meetings in Japan, South Korea, Taiwan, and China in rotation.

As per the agreement, the second NASSDA meeting, hosted by SRDA, was held on September 16, 2017, at Academia Sinica in Taiwan. SSJDA, KOSSDA, and SRDA all provided activity reports during this workshop. We discussed collaboration methods and agreed to construct a shared data catalog for the participating archives, incorporating web links and concise introductions to the archives on their respective websites. Additionally, a designated contact person for each data archive will be appointed to streamline future cooperation.

Having successfully completed the initial stage of international collaboration, NASSDA has now entered a more routine phase. Each member takes turns hosting an annual meeting, providing a platform for sharing updates and activities. The host also organizes a session with lectures by inviting experts from Europe and the United States, fostering a culture of knowledge sharing and mutual growth. For instance, SSJDA introduced Social Science One and Dataverse, showcasing the potential of industry-academic partnerships in 2021. In 2022, CNSDA invited experts from CESSDA to provide an update on their recent activities. The 2023 meeting featured SRDA inviting Australian Data Archive (ADA) experts to present the Social Science Data Archiving and Research Infrastructure in Australia.

Furthermore, at NASSDA’s business meetings, members exchange opinions on the work plan, possibilities for future cooperation, and ideas for collaboration. For example, at the 2018 annual meeting, we discussed collaboration methods, which included drafting a Memorandum

² The GESIS – Leibniz Institute for the Social Sciences is the largest European infrastructure institute for the social sciences, and offers advice, expertise, and services at all stages of scientists’ research projects.

of Understanding, exchanging banners on each member archive's webpage, and addressing NASSDA's webpage design, governance, and future membership. Opinions were exchanged on other collaborative issues, such as utilizing Nesstar, standardizing data identification formats, organizing data fairs, and implementing educational programs. At the 2019 meeting, discussions expanded to include banner exchanges and public outreach, collaboration in data service standards, data certification, Dataverse, funding availability, and staff training for each archive. Also, during the 2021 meeting, SRDA shared their experience in applying for the Core Trust Seal, offering valuable insights to other members.

Collaboration and challenges of NASSDA

The NASSDA members regularly share information during annual meetings, via email to deepen their professional exchanges. Various collaborations have been discussed to increase the visibility of NASSDA, as well as its member archives. The launch of the NASSDA website³ in 2019 is an important collaboration among member archives. The NASSDA's website, hosted by SSJDA, has been developed as a platform for members to share information about events they organize. It aims to promote exchanges between data archives and researchers in East Asia. Additionally, members actively promote NASSDA through domestic and international conferences.

In addition, individual archives occasionally organized educational programs with invited guests from member archives to discuss their data collections, giving local participants a better understanding of international data resources. For example, the 2017 International Workshop on Social Science Data Archives in Taiwan introduced the development and practices of SSJDA, KOSSDA, and SRDA, while the 2018 Data Fair in Korea focused on well-being data and related research across NASSDA's member archives. These events have drawn significant interest from local researchers.

It is also important to ease the accessibility to data for researchers in other member archives, particularly to those with metadata in local languages, but not all of the planned collaborations were successfully implemented. Since most member archives use Nesstar as an online analysis tool, adopting it as the main portal for data search was considered. However, this idea was discontinued due to the end of updates for Nesstar. With regard to open data initiatives and the establishment of various platforms, data catalogue that provides information of available datasets in the NASSDA member archives has been proposed. However, technical issues have hindered collaboration, because not all system configurations between archives and platforms are compatible.

A common obstacle faced by NASSDA member archives is the hesitation among social scientists to share data. Feedback received by SRDA in 2023 highlighted reasons similar to those identified in previous research, while more challenges emerged (Akdeniz et al., 2023; Kim and Adler, 2015). Promoting open data remains essential, especially within the social sciences. We then introduce each NASSDA member archive, highlighting the unique aspects of each. For example, KOSSDA possesses unique qualitative interview data unavailable in other archives, while CNSDA, despite holding the smallest data volume, has the highest number of users. In addition, SRDA effectively leverages technology to enhance user services and was the first to obtain CoreTrustSeal certification. Meanwhile, SSJDA has cultivated numerous early-career researchers through its extensive academic resources. While each institution has its unique characteristics, which can pose challenges to collaboration, this also provides a valuable opportunity for members to learn from one another and address mutual shortcomings. This network is designed to maintain flexible collaboration while enabling each institution to leverage its unique strengths fully.

Individual member archives of NASSDA

CNSDA

The Chinese National Survey Data Archive (CNSDA)⁴ is a groundbreaking social science data repository that has become instrumental in facilitating data sharing and promoting evidence-

³ The NASSDA website: <https://nassda.org/>.

⁴ The CNSDA website: <http://cnsda.ruc.edu.cn/>.

based research in mainland China. Established in 2012, CNSDA has rapidly emerged as an essential resource for researchers, students, and policymakers seeking to understand various facets of Chinese society.

Establishment and purpose

Founded in 2012, CNSDA began providing online data access services in 2014. The archive is primarily funded by the Natural Science Foundation of China (NSFC) and operates as a flagship project of the National Survey Research Center at Renmin University of China. The core mission of CNSDA is to acquire, curate, and provide open access to survey data collected in mainland China.

The Chinese National Survey Data Archive aims to collect and preserve raw data and documents from survey projects conducted in mainland China, perform data curation and documentation to enhance the usability of the collected survey data, provide open access and related services for the stored survey datasets, and promote data sharing and collaboration within the academic community. These foundational goals have shaped CNSDA's operations and growth since its inception.

Data holdings and user base

Over the years, CNSDA's data holdings have steadily expanded. As of December 31, 2023, the archive housed 106 survey datasets, with new additions being made annually. Some of the most notable datasets include the Chinese General Social Survey (CGSS), China Education Panel Survey (CEPS), and Chinese Private Entrepreneur Survey. These datasets cover a wide range of topics and provide valuable insights into various aspects of Chinese society.

The Chinese National Survey Data Archive caters to a diverse user base, including students, academic researchers, commercial users, and governmental and nonprofit organizations. As of December 31, 2023, the archive had an impressive 71,828 registered users, with a significant portion (11.3%) coming from overseas (Figure 1). The archive's international reach extends to 34 countries. This global user base underscores the international relevance and impact of CNSDA's data holdings.

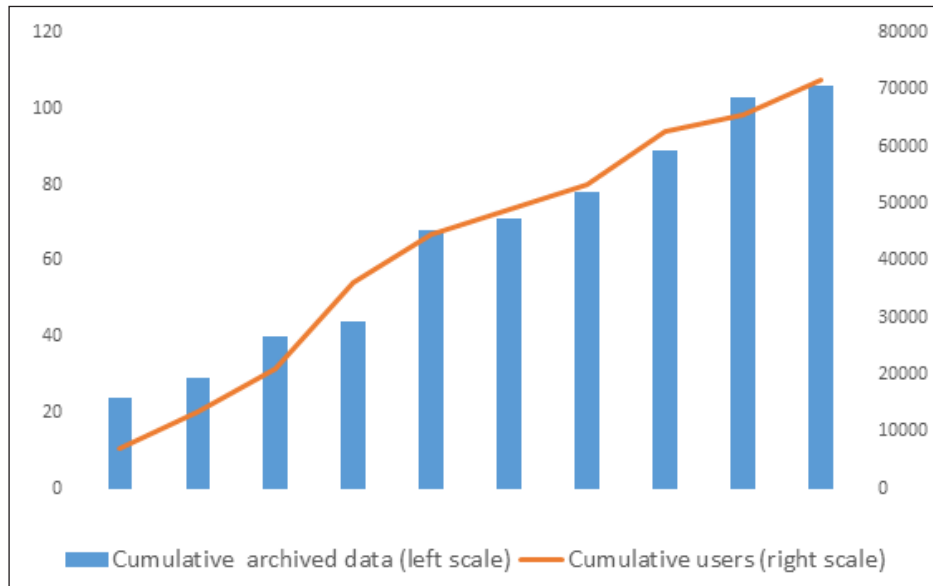


Figure 1 Archived data and registered users of CNSDA.

Research impact and knowledge creation

The datasets provided by CNSDA have been extensively used in research publications, including journal papers, theses, dissertations, and books. As of 2023, there were over 6,500 Chinese journal papers, 1,200 English journal papers, 2,500 Chinese theses and dissertations, and 150 English doctoral dissertations that utilized CNSDA's data. These numbers demonstrate the significant impact CNSDA has had on research output and knowledge creation, both within China and internationally. The archive's contribution to evidence-based research continues to grow each year.

Expanding services and collaborations

The Chinese National Survey Data Archive actively seeks to expand its services and collaborate with other institutions to enhance its offerings. In 2023, the archive collaborated with the

Department of Anthropology at Renmin University to establish a new sub-archive focused on curating and preserving academic ethnographic films related to China. This collaboration broadens the scope of CNSDA's data holdings and provides researchers with access to valuable visual anthropological resources.

Furthermore, CNSDA organizes data events such as the CGSS Data Analysis and Visualization Competition to showcase the potential of its datasets and encourage innovative research. These events foster a vibrant research community and promote the use of CNSDA's data in creative and impactful ways.

Future developments and importance

Looking ahead, CNSDA is developing a next-generation system platform to improve its services and user experience. This new system aims to enhance data access, management, and analysis capabilities, empowering researchers to fully leverage the archive's extensive data holdings. As CNSDA continues to grow and evolve, it is poised to play an increasingly important role in shaping the future of social science research in China and beyond.

The Chinese National Survey Data Archive has firmly established itself as a vital resource for social science research in China and worldwide. Through its commitment to data sharing, curation, and preservation, CNSDA has enabled researchers to conduct evidence-based studies and contribute to a deeper understanding of Chinese society. With ongoing enhancements and collaborations, CNSDA is set to further solidify its position as a cornerstone of social science research infrastructure, driving knowledge creation and policy development for years to come.

KOSSDA

Establishment and purpose

The Korea Social Science Data Archive (KOSSDA),⁵ a leading data repositories in Korea, originated as a social science library established by a non-profit foundation in 1983. In 2006, recognizing the growing importance of data in social science research and the need to address the limited accessibility of high-quality datasets produced by major research institutes in Korea, it evolved into a comprehensive data archive, housing both quantitative and qualitative research data. This shift was driven by the understanding that open access to reliable and well-documented data is crucial for advancing social science research. At that time, the idea of sharing or reusing research data was not widely recognized, so data was often lost in researchers' drawers or computers after being used once in research. Researchers and students, who had limited resources to produce data for their own, had difficulty finding and accessing the data they needed. Recognizing this gap, KOSSDA redefined its mission and aimed to maximize the potential value of data and advance social science research by ensuring data integrity and facilitating universal access for the public.

Nine years into its successful operation, the KOSSDA foundation entered into an agreement with Seoul National University to transfer its operations, aiming to foster collaboration between education, research, and data archiving in 2015. Currently based at Seoul National University, KOSSDA continues to provide an integrated platform offering diverse quantitative and qualitative research data and outputs, ensuring free and unified access across various social science fields. Its user base encompasses academic researchers, students, educators and lecturers, government analysts, and independent researchers.

Main activities

The Korean Social Science Data Archive's core activities revolve around three pillars: data service, education, and research.

Data service. The Korean Social Science Data Archive acquires data through deposit agreements with individual researchers and data sharing agreements with organizations such as university research institutes, government-funded research organizations, project teams supported by the National Research Foundation of Korea, survey firms, and corporate research institutes. Over 100 organizations share their data and provide access to their data through KOSSDA.

⁵ The KOSSDA website: <https://kossda.snu.ac.kr/>.

The Korean Social Science Data Archive specializes in curating and managing of research data, ensuring its long-term preservation and open access. Curation involves a meticulous process of data cleaning, quality control, metadata enrichment, and standardization of data formats to enhance data usability and ensure consistency across datasets. The Korean Social Science Data Archive also provides support throughout the data life cycle assisting researchers and students in depositing, documenting, and accessing data. As of 2023, KOSSDA holds over 3,100 quantitative datasets and 200 qualitative datasets. Users can search and access data at a variable level as well as at the study level, making it easier and more efficiently to find the data they want. The Korean Social Science Data Archive collects and provides citation information for data-related research outputs, including academic papers, theses, books, monographs, reports, etc., aiding users in exploring data through literature and understanding how the data in its collections has been utilized in research. Over 53,000 users visit KOSSDA’s website annually, with over 10,000 data downloads (Figure 2).

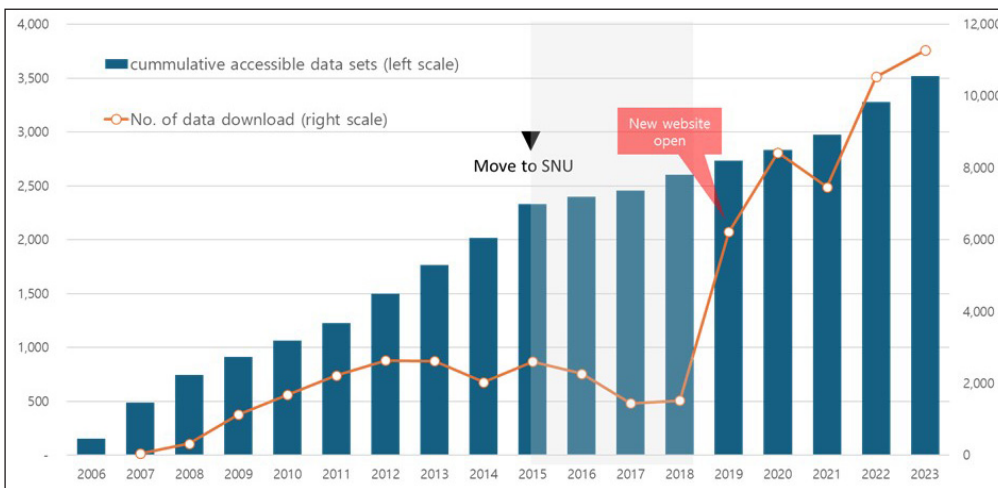


Figure 2 Deposits and usage situation of KOSSDA data.

The Korean Social Science Data Archive also collaborates with Statistics Korea to oversee access to the database housed in Statistics Korea’s Secure Data Center at Seoul National University, facilitating researchers’ secure access to administrative and sensitive data.

Education. The Korean Social Science Data Archive regularly provides online and offline educational programs on qualitative and quantitative methodologies and data analysis, ranging from basic to advanced levels. These programs include two-week or one-week sessions during summer and winter vacations, as well as four to five-day workshops held on weekends in spring and fall. Emphasizing hands-on training using selected KOSSDA datasets, these programs equip participants with practical data handling and analysis skills. In 2023, a total of 25 courses were offered and a total of 1,250 participated in the courses.

Additionally, KOSSDA hosts academic events such as the Data Fair and the KOSSDA Thesis Award, along with providing tutorial courses and seminars on research data management. The Data Fair, held twice a year, offers a platform for data producers to directly present datasets widely utilized in the social science field. It also provides an opportunity to introduce related analysis methodologies, fostering the exchange of knowledge and analysis experiences among participants. The Korean Social Science Data Archive plays a key role in organizing and connecting participants, as well as providing a space for them to interact with each other. It also advocates for open data and open science through events like the Open Social Science Data Forum and initiatives like data citation campaigns.

Research. The Korean Social Science Data Archive actively engages in research projects to promote a robust data sharing ecosystem and advance data-driven research in the social sciences. These projects encompass various areas such as assessing the utilization and impact of emerging data types such as social media and administrative data; developing comprehensive guidelines for data sharing and reuse while ensuring data privacy and security; and investigating the ethical implications of data sharing; and exploring innovative methodologies for integrating and linking diverse datasets to unlock new research possibilities. In addition, KOSSDA will collaborate with international comparative research projects such as

the World Values Survey and the Comparative Study of Electoral Systems to collect and provide high-quality data that can be widely used in the social sciences.

The Korean Social Science Data Archive also participated in a Korea Research Memory (KRM) Database Construction project funded by Korea National Research Foundation. Korea Research Memory is an integrated and open data repository and service platform that aims to preserve and reuse research data produced throughout the NRF-funded research.

Future Developments

The Korean Social Science Data Archive acknowledges the substantial demand for access to high-quality, well-documented datasets. It is committed to developing clear data and documentation standards to ensure consistency and interoperability across a variety of datasets. The Korean Social Science Data Archive also remains adaptable to new research methodologies, data types, and technological advancements that arise from the changing landscape of social science research. Moreover, collaboration between research data archives on a global scale will foster the exchange of best practices, data standards, and cross-cultural insights. By prioritizing user needs, maintaining adaptability, and fostering collaboration with various stakeholders, KOSSDA tries to be well-prepared for future changes and continues to serve as a valuable resource for the social science community.

SRDA

The Survey Research Data Archive (SRDA)⁶ was founded in 1994 by the Center of Survey Research (CSR, formerly the Office of Survey Research), and is a CSR's division of the Research Center for Humanities and Social Science at Academia Sinica in Taiwan. The Survey Research Data Archive engages in the systematic acquisition, organization, and preservation of academic survey data, and its dissemination to scholars and researchers mainly for academic purposes. Standard data management procedures are applied to ensure data are compatible with the FAIR Data Principles. By collecting and disseminating microdata on social and statistical survey research, SRDA provides valuable data resources for academic and policy-related studies both domestically and internationally.

Missions and data infrastructure

Long-term preservation of survey datasets is one of the responsibilities of SRDA. Its data stock is considered to be one of the largest in scale/quantity in East Asia. The sources of datasets mainly include government and academic surveys, as being cross-sectional, longitudinal, or cohort studies. Until the end of 2023, there are more than 2,900 datasets released by SRDA. The number of studies collected in SRDA is around 3,600 (Figure 3). While most of the datasets

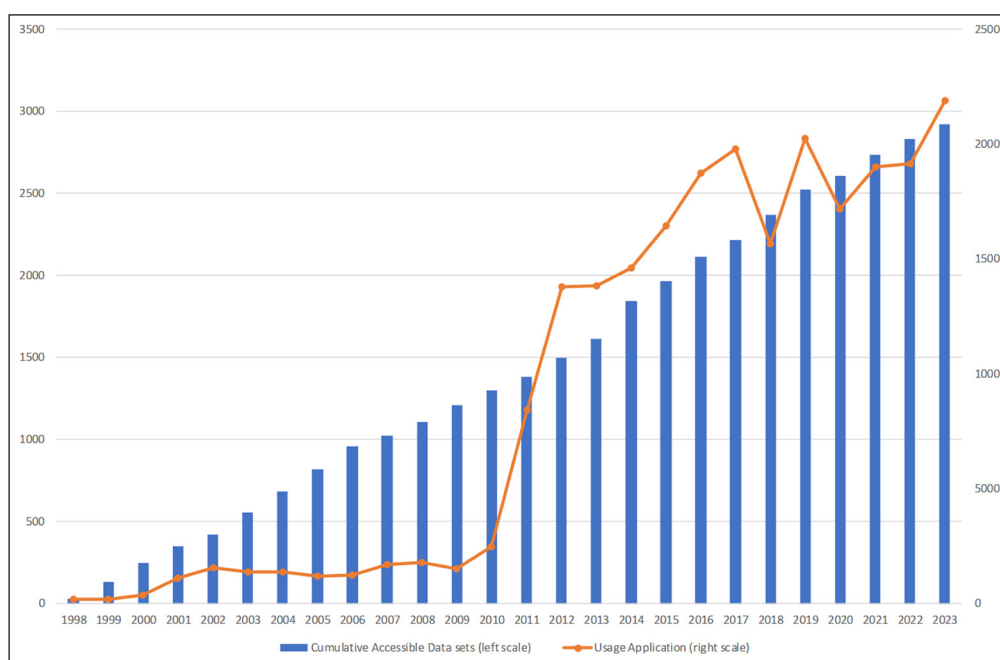


Figure 3 Deposits and usage situation of SRDA data.

6 The SRDA website: https://srda.sinica.edu.tw/index_en.php.

in SRDA are on issues in social sciences, datasets on public health, nursing, and psychiatry serve a good resource for journal articles and theses. Modules on health care and environment hazard in social surveys are utilized by multi-discipline researchers as well.

Access to raw data requires one of four SRDA memberships, each offering different data access levels. The web membership is open to the general public with a valid email address. Regular memberships, available for researchers (two-year) and students or research assistants (one-year), and AS membership, exclusively for Academia Sinica researchers. Raw data can be accessed via direct download for public or standard data, and remote access or onsite for restricted data, all free of charge. Metadata and documentation of all the data are open for browsing and downloading on the SRDA website, regardless of users' membership types.

The Survey Research Data Archive has developed a data management system for editing metadata and uploading online, so users can browse or download related files. Prior to the release of data sets, the confidentiality and sensitivity of every survey data set are evaluated. Standard data management and cleaning procedures are applied to ensure data accuracy and completeness. In addition, metadata and relevant supplement files are also edited and attached.

In order to ensure data security, SRDA has obtained ISO 27001:2005 certification for its digital data storage and usage services in 2010. The Survey Research Data Archive is dedicated to protecting not only the confidentiality, completeness and applicability of survey data, but also confidential information pertaining to the SRDA members in the process of transmission, archiving and application. In addition, SRDA has also obtained ISO 27001:2013 transition certifications in 2015 and remains valid until at least 2024.

With a large holding of datasets, SRDA has been working hard on promoting the archived data, through regular academic activities, such as webinars, workshops and lectures on campus. In addition, monthly E-digest, YouTube channel, blog, and Facebook fan page are used to promote new released data and SRDA activities as well as to communicate with the existing and potential members.

Value-added initiatives and international engagement

Bibliographies using SRDA data are continuously collected both actively and passively. Members who have used SRDA data for theses, journal articles, and book chapters can provide bibliographic information via the SRDA website platform. Additionally, emails are sent to members one year after they downloaded or analyzed data through SRDA, requesting related bibliographies. Since 2018, SRDA has employed Python algorithms to web crawl digital index systems for bibliographies of 16 commonly used research projects in Taiwan. These systems include the National Digital Library of Theses and Dissertations in Taiwan, Index to Taiwan Periodical Literature System, ProQuest, Web of Science, and EBSCOhost.

In order to add values to the existing data, SRDA has developed SRDA Data Explorer, which a platform aiming to enable better access to longitudinal data for secondary data analyses. The Taiwan Social Change Survey (TSCS) has been chosen as the first project. Great efforts have been made to align the descriptions and values of the variables of same meanings in the TSCS across survey waves. By entering keywords or selecting variable subjects and survey modules and waves, SRDA members can easily search, browse and download the TSCS data across years (1985–2014).

The Survey Research Data Archive is keeping up with global developments. Since 2016, it has been a funder and member of NASSDA. In 2019, SRDA earned the CoreTrustSeal certification, proving its status as a trustworthy data repository. Additionally, SRDA transferred DOIs registration with DataCite using the da|ra service (Helbig, Hausstein and Toepfer, 2015) in 2016 and established a data-search platform using Harvard Dataverse in 2019. To register DOIs, SRDA runs scripts weekly to create XML metadata in form, and uploads XML files through the API Client. Harvard Dataverse is requested to harvest SRDA data from DataCite to increase the discoverability of our archived data.

Future developments

For the next steps, a newly designed SRDA website is set to launch in late 2024. The new site aims to be faster, easier to navigate, and more user-friendly, featuring a clean design and intuitive, consistent navigation. Built on UX/UI principles and utilizing natural language

processing techniques in Elasticsearch technology, the website will help users easily find their data. Additionally, text mining techniques will categorize research themes, enhancing users' understanding of the relationships among datasets. With natural language algorithms and advanced features like unified queries, Boolean operations, and term association recommendations, the new system will not only optimize query speed but also transform the user experience.

In addition, advanced text mining techniques are utilized to improve data categorization for our users. Our thematic structure encompasses 19 topics, including aging, children/youth research, family, communication/social media, lecture studies, health and medical care, and more. Each topic has been curated with great care to provide users with a focused and intuitive exploration of our diverse datasets. To enhance the user experience, we offer a brief history of the datasets, data visualizations, and data guides provided by scholars. These resources assist users in navigating our thematic data collection more efficiently. The Survey Research Data Archive's commitment to thematic data collection demonstrates our dedication to making information accessible, organized, and user-friendly.

To gather insights into the experiences and preferences of users of certain large-scale surveys, SRDA received funding for a three-year project aimed at promoting the long-term development and more value-added applications of survey data in Taiwan. This project aims to achieve the following objectives:

- a. Gathering information on international empirical research trends and domestic scholars' data needs.
- b. Enhancing data management based on international and domestic survey data practices.
- c. Promoting interdisciplinary database collaboration.

It is expected that this project will provide valuable insights from data users for the future developments of SRDA.

SSJDA

The Social Science Japan Data Archive (SSJDA)⁷ was established in 1996 at the Information Center for Social Science Research in Japan, which was affiliated with the Institute of Social Science, the University of Tokyo. It collects and stores individual survey data and information related to statistical and social surveys in the social sciences in Japan. Its primary objectives include preventing data dispersion and facilitating access to secondary analyses for academic purposes. In April 2009, the Center was reorganized and renamed the Center for Social Research and Data Archives (CSRDA).

Purposes and activities

The SSJDA has been enhancing its operations to expand the number of depositors and users and to improve overall convenience. Specifically, it has focused on three key activities: (1) enhancing the quality and quantity of deposited data; (2) establishing a unified management system to streamline information from both depositors and users, enabling follow-up on data usage; and (3) exploring the introduction of new data delivery methods to complement the traditional method of providing microdata to users on CD-Rs.

To enhance deposited data quality and quantity (1), we have been actively engaging with private and public survey agencies and university researchers to encourage data deposits. Specifically, we have received data from more than 200 institutions, organizations, and researchers. The microdata deposited in the archive are organized into a format suitable for secondary use after creating metadata and performing the necessary anonymization processes. Furthermore, the data are made public after confirmation by the depositor. The entire data curation process has been conducted by specialized staff, ensuring consistent quality of data for secondary analysis. For managing information regarding depositors and users ((2) and (3)), we have been using a user database system since fiscal year 2006. This system allows us to centrally manage information from application submission to reporting via mail, thus facilitating the smooth handling of an increasing number of users. Additionally, to improve user convenience and streamline operational tasks, we began developing a web system (SSJDA Direct) in fiscal

⁷ The SSJDA website: <https://csrda.iss.u-tokyo.ac.jp/english/>.

year 2008, which has been operational since 2009. This system enables online application submission, data download, usage reporting, administrator approval, and comprehensive usage information management.

Data usage and infrastructure development

Both new data deposition and usage have been increasing annually, with a total of 2,453 cumulative datasets published as of March 2023 (Figure 4). The number of individual data users has also steadily increased, reaching 1,410 applicants between April 2023 and March 2024, with a total of 7,790 users, including those utilizing data for educational purposes. The total number of datasets provided has reached 23,800. Among them, 75 applications from 56 foreign institutions provided 369 datasets. In fiscal year 2022, 373 papers and books were published using the datasets provided by the SSJDA.⁸

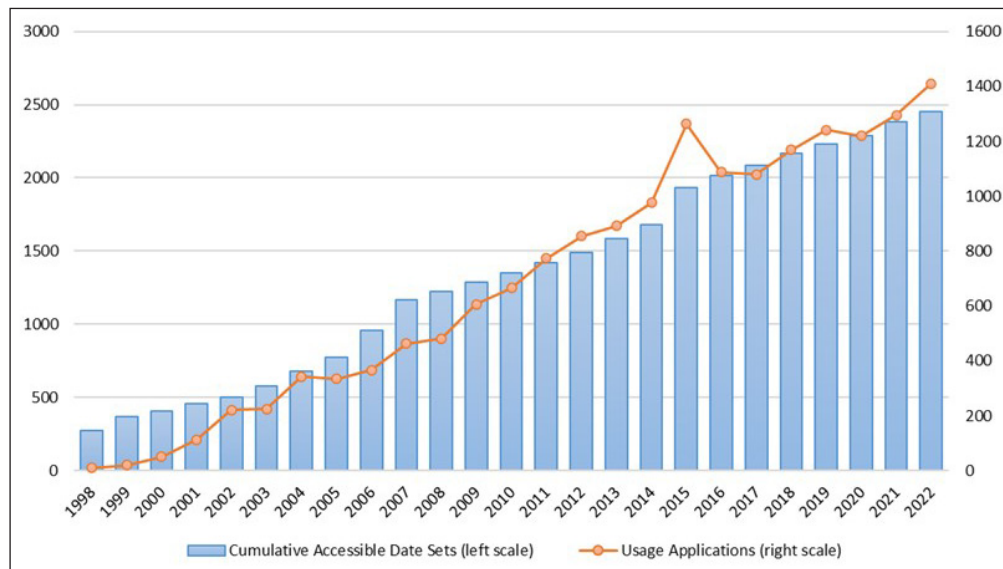


Figure 4 Deposits and usage situation of SSJDA data.

In addition to providing data for secondary use, SSJDA offers an online system for simple computation and analysis. The remote computation system, which has been in operation since October 2005, allows users to quickly generate cross-tabulation tables and other preliminary analyses on the web before conducting full-scale secondary analyses. Furthermore, in 2024, following the discontinuation of Nesstar, we introduced a new online aggregation system called SSJDA Data Analysis, which was launched in October 2022. We plan to progressively incorporate survey data into this system in anticipation of further expansion of its usage.

Promotion initiatives and collaborations

Furthermore, SSJDA conferred commendation awards to depositors and users to promote the deposition of survey data and secondary data analysis. The Depositor Commendation Award, initiated in fiscal year 2009, evaluates the deposited survey data based on three aspects—quality, quantity, and diversity—and recognizes outstanding depositors. A total of 26 institutions and individuals have been awarded. Additionally, reinstated in fiscal year 2010, the Outstanding Paper Commendation Award selects and commends outstanding papers from a list of achievements reported in that year. To date, 35 authors have been honored with this award.

The CSRDA organizes ‘Secondary Analysis Workshops’ and ‘Quantitative Analysis Seminars’ to promote the effective utilization of the survey data archived in the SSJDA. The research outcomes of these activities were compiled into a Research Paper Series (SSJDA-RPS). Additionally, the CSRDA launched the development of a new panel survey, the SSJDA Panel, which invites researchers to propose additional questions. The resulting dataset is then made publicly available through the SSJDA.

⁸ Please refer to the Annual Report of the Institute of Social Science, the University of Tokyo, p. 40. (in Japanese).

Furthermore, initiatives toward the internationalization of the data archive are being pursued, including membership in the International Federation of Data Organization, operation as a hub institution for the ICPSR Japanese National Membership, hosting international seminars and workshops, translation of survey instruments into English, and collaboration on East Asian international comparative surveys. These initiatives aim to strengthen collaboration with the data archives in Taiwan, South Korea, and China, while promoting further international collaborative research.

Future developments

Stakeholders from various sectors recognize and value the activities of the SSJDA and the efforts of the CSRDA. In 2010, the CSRDA was designated as a joint-use and joint-research facility by the Japanese Ministry of Education, Culture, Sports, Science and Technology, further enhancing its research activities. Moreover, since 2018, the CSRDA has been recognized as a base institution for the Program for Constructing Data Infrastructure for Humanities and Social Sciences by the Japan Society for the Promotion of Science (JSPS). Since 2023, it has also been certified as a core institution for the Program for Strengthening Data Infrastructure for the Humanities and Social Sciences by the JSPS. These recognitions reflect its steady progress in both infrastructure enhancement and internationalization efforts.

In the future, SSJDA aims to take the initiative in Japan and lead international collaboration, while inheriting its mission and striving to further enhance its activities.

CONCLUSION

The collaboration of data archives in East Asia and the establishment of NASSDA have greatly benefited the member archives. One of the most notable advantages is the sharing of updated information and challenges we face as well as the opportunity to learn from each other's experiences. Since their inception in 2008, exchanges among social science data archives in East Asia have continued for over a decade. Initially, these exchanges began in the form of discussions among researchers on the construction and sharing of comparable international survey data in East Asia rather than focusing on the practical aspects of data management and archive operations. While until the 2010s, digitization was primarily centered around academic journals and papers, since the early 2010s, open science, including the open sharing of research data, has emerged as a global trend that has been further accelerated by the COVID-19 pandemic. Against this backdrop, East Asian data archives have faced increasing pressure to enhance infrastructure development and improve data utilization, while also recognizing the increased need to strengthen collaboration among archives, leading to the NASSDA initiatives.

As discussed earlier, common challenges persist for all member archives. While each archive has seen a significant increase in data users, the number of data depositors has not risen at a similar rate. Regarding open data initiatives, many researchers recognize the importance and benefits of sharing research data, yet few move beyond expressing willingness, and actual data-sharing practices remain limited (Fecher et al., 2017). All NASSDA's member archives are eager to increase the volume of archived social science data, though individual researcher-related obstacles persist.

Indeed, strengthening collaboration between archives is not always straightforward. Numerous challenges can impede progress, such as differences in infrastructure and systems for survey data loading, financial and personnel situations, national policies, and domestic circumstances. The current focus of NASSDA is to collaborate within existing constraints, such as by establishing a website and organizing annual meetings.

While NASSDA has focused on fostering information exchange and collaboration among East Asian archives, it should now expand its role to support the development of data sharing culture and infrastructure across Asia. Furthermore, NASSDA can contribute to the advancement of global social science by engaging with archives in other regions. To achieve this, we must explore strategies for expanding our network, facilitating data exchange, and collaborating with archives in other parts of the world.

The authors have no competing interests to declare.

AUTHOR CONTRIBUTIONS


This paper was collaboratively developed. Each author contributes to the writing and editing and is responsible for all parts of the article.


AUTHOR AFFILIATIONS

Pei-Shan Liao  orcid.org/0000-0001-9809-7418
 Academia Sinica, TW

Shuai Wang  orcid.org/0000-0001-7981-4768
 The University of Tokyo, JP

Hiroshi Ishida  orcid.org/0009-0009-7227-8195
 The University of Tokyo, JP

Hearan Koo  orcid.org/0009-0004-6918-4992
 Seoul National University, KR

Won-ho Park  orcid.org/0000-0002-7360-9111
 Seoul National University, KR

REFERENCES

- Akdeniz, E., Borschewski, K.E., Breuer, J. and Voronin, Y.** (2023) 'Sharing social media data: The role of past experiences, attitudes, norms, and perceived behavioral control', *Frontiers in Big Data*, 5, pp. 971974. Available at: <https://doi.org/10.3389/fdata.2022.971974>
- Fecher, B., Friesike, S., Hebing, M. and Linek, S.** (2017) 'A reputation economy: how individual reward considerations trump systemic arguments for open access to data', *Palgrave Communications*, 3(1), pp. 17051. Available at: <https://doi.org/10.1057/palcomms.2017.51>
- Hedrick, T.E.** (1988) 'Justifications for the sharing of social science data', *Law and Human Behavior*, 12(2), pp. 163–171. Available at: <https://doi.org/10.1007/BF01073124>
- Helbig, K., Hausstein, B. and Toepfer, R.** (2015) 'Supporting Data Citation: Experiences and Best Practices of a DOI Allocation Agency for Social Sciences', *Journal of Librarianship and Scholarly Communication*, 3(2), eP1220. Available at: <https://doi.org/10.7710/2162-3309.1220>
- Kim, Y. and Adler, M.** (2015) 'Social scientists' data sharing behaviors: Investigating the roles of individual motivations, institutional pressures, and data repositories', *International Journal of Information Management*, 35(4), pp. 408–418. Available at: <https://doi.org/10.1016/j.ijinfomgt.2015.04.007>
- Kirilova, D. and Karcher, S.** (2017) 'Rethinking data sharing and human participant protection in social science research: Applications from the qualitative realm', *Data Science Journal*, 16(43), pp. 1–7. Available at: <https://doi.org/10.5334/dsj-2017-043>
- Klump, J.** (2017) 'Data as social capital and the gift culture in research', *Data Science Journal*, 16(14), pp. 1–8. Available at: <https://doi.org/10.5334/dsj-2017-014>
- Kvalheim, V. and Kvamme, T.** (2014) *Policies for sharing research data in social sciences and humanities: A survey about research funders' data policies*. Institutional Federation of Data Organizations for Social Science. Available at: https://www.dcc.ac.uk/sites/default/files/documents/reports/ifdo_survey_report.pdf.
- Levenstein, M.C. and Lyle, J.A.** (2018) 'Data: Sharing is caring', *Advances in Methods and Practices in Psychological Science*, 1(1), pp. 95–103. Available at: <https://doi.org/10.1177/2515245918758319>
- Sieber, J.E.** (1991) *Sharing social science data: Advantages and challenges*. SAGE Publications, Inc. Available at: <https://doi.org/10.4135/9781483325620>
- Tedersoo, L., Kungas, R., Oras, E., Koster, K., Eenmaa, H., Leijen, A., Pedaste, M., Raju, M., Astapova, A., Lukner, H., Kogermann, K. and Sepp, T.** (2021) 'Data sharing practices and data availability upon request differ across scientific disciplines', *Scientific Data*, 8(1), pp. 192. Available at: <https://doi.org/10.1038/s41597-021-00981-0>

TO CITE THIS ARTICLE:

Liao, P-S, Wang, S, Ishida, H, Koo, H and Park, W-h 2025 NASSDA as a Data-Collaborative Network to Enhance Social Science Research in East Asia. *Data Science Journal*, 24: 3, pp. 1–14. DOI: <https://doi.org/10.5334/dsj-2025-003>

Submitted: 24 July 2024

Accepted: 16 December 2024

Published: 08 January 2025

COPYRIGHT:

© 2025 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC-BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. See <http://creativecommons.org/licenses/by/4.0/>.

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