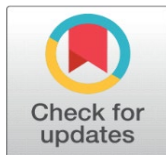
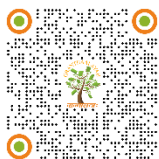


OPEN DATA CULTURE IN SOCIAL SCIENCE RESEARCH

Dr. Rituparna Basak ¹  

¹ Assistant Professor, Department of Psychology, Muralidhar Girls' College, Kolkata, India



Received 13 April 2023

Accepted 15 May 2023

Published 31 May 2023

Corresponding Author

Dr. Rituparna Basak,
rituparna.basak@muralidhargirlscollege.ac.in

DOI
[10.29121/granthaalayah.v11.i5.2023.5167](https://doi.org/10.29121/granthaalayah.v11.i5.2023.5167)

Funding: This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Copyright: © 2023 The Author(s). This work is licensed under a [Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/).

With the license CC-BY, authors retain the copyright, allowing anyone to download, reuse, re-print, modify, distribute, and/or copy their contribution. The work must be properly attributed to its author.



ABSTRACT

Open data are records that are available for anybody to access, reuse, and distribute without restriction, with the exception of sharing and attribution. Many national governments have created websites to make some of the data they gather accessible to public, joining private companies in doing so. The creation of such massive archives is expected to promote and speed scientific progress by allowing multiple uses of datasets and minimizing duplication of effort [Hand et al. \(2001\)](#). When a repository is fully made up of datasets supplied by researchers and made accessible for use by other researchers, future uses of such datasets are referred to as reuse. When it comes to reusing research data in the social sciences, it has been discovered that quantitative data reuse is more widespread than qualitative data reuse since the number of free quantitative datasets is bigger [Curty et al. \(2017\)](#) and quantitative data metadata is easier to develop. Nonetheless, there is various research on the reuse of qualitative data in social sciences [Late & Kekäläinen \(2020\)](#). Despite several issues that must be addressed, it is obvious that open access data has the potential to transform research processes in a variety of sectors.

Keywords: Open Data, Social Science, Data Sharing, Data Use and Reuse

1. INTRODUCTION

The term "open science" describes initiatives to make the findings of publicly financed research more broadly accessible in digital form to the scientific community, the business world, and the general public. Open science is the fusion of science's historical tradition of openness with the information and communications technologies (ICTs) tools, which have changed the scientific enterprise and require a critical evaluation from policymakers aiming to support long-term research and innovation. Economist [David et al. \(2003\)](#) invented the phrase "open science" in an effort to describe the characteristics of scientific output produced by the public sector and in response to the perceived expansion of intellectual property rights

into the world of digital goods. The case for open science's influence on research, business, innovation, and society at large is becoming more and more compelling.

Data is becoming increasingly important in science, and recent technical advancements, computing capabilities, and new digital settings are putting data at the core of scientific discovery. Data sharing by researchers has become a critical component of scientific advancement, as data has become a central component of modern scientific discoveries. Control can be exercised through the use of licencing, copyright, patents, access fees, and other legal restrictions. Open data is information that is available for everyone to see, use, or exchange. As a result, open data refers to information that is freely accessible to the public for any reason.

Data can be classified by kind (for example, facts, observations, and lab notes) or context (such as information obtained using public funds). Research data can occasionally be separated from data from other sources, such official statistics or industry reports [Open Knowledge Foundation \(2015\)](#). *Open Definition: Defining Open in Open Data, Open Content and Open Knowledge.*, n.d.) Data, according to certain academics, refers to objects used as evidence of occurrences for research or scholarly endeavours [Borgman \(2015\)](#). As a result, any observation, record, or other piece of information can be referred to as data.

Greater multidisciplinary scientific collaboration, data accessibility, and scientific work that is more reproducible and transparent are among the goals of Open Science. Data sharing is becoming more widely recognised as a critical engine of science's global and cross-disciplinary advancement [Tenopir et al. \(2011\)](#), [Tenopir et al. \(2015\)](#), [Schmidt et al. \(2016\)](#). In accordance with the best data management practices, scientists must make their datasets available by depositing them in trustworthy institutional, governmental, or subject repositories, providing metadata that enables data discovery, and citing or acknowledging data reuse. Understanding how scientists behave is crucial to decide how to support the scientific community by implementing best practices [Tenopir et al. \(2020\)](#).

2. OBJECTIVES AND RESEARCH

The goal of this research is to better comprehend the notion of open science and to assess the state of open data in many domains, particularly social sciences. This paper intends to explain the benefits, difficulties, and limitations of open data sharing in-depth, as well as offer potential solutions based on all of the variables. It also aims to address data management approaches in various disciplines, with a focus on FAIR principles and data handling. Finally, it goes over the ramifications, documentation, and data architecture of using and reusing both qualitative and quantitative data.

In order to achieve such goals, a supplementary method is used. The current study's data and information were gathered from a variety of national and international websites, published books, magazines, journals, reports, and materials linked to open data for the study.

This study deduced open data sharing in detail from this information, as well as data management methodologies and data processing. It explored the implications of using and reusing data in social science research.

3. OPEN DATA

Anyone may use, reuse, and share open data without restriction, with the exception of sharing and credit. Various open data initiatives are listed by

researchers [Pasquetto et al. \(2017\)](#), including virtual observatories (e.g., Digital Earth), federated data networks (e.g., World Data Centers, NASA Distributed Active Archive Centers), repositories and archives (e.g., GenBank, Protein Data Bank), domain repositories (e.g., PubMedCentral).

Openness is different in many ways. It's possible for contributors to public data repositories to maintain ownership and control over the data they've deposited. Although the data is open, only proprietary tools may be used to interpret it. Data can be created using open source software, but data use requires licensing. While some open data repositories have long-term plans, many rely on temporary funding or the success of business models. In order to respond to changes in the user population, long-term data preservation frequently requires continuing investments in curation [Baker et al. \(2015\)](#).

The acceptance of open science is expanding. Data sharing, making underlying materials and protocols available, and preregistering studies and analytic plans are all becoming more common across fields. Open Data is already a reality for many scientists, particularly those working in the social sciences. This is evidenced by the increasing number of (interested) researchers that make their data widely available and use data from others in their research. In the social sciences, this is not the case, but data still exists: primary source texts, archives, and artistic representations are all examples of data. Unlike scientists, social scientists have typically considered digital tools as an add-on to their study.

All modern social science research is concerned with complex, fast-changing "human-dominated and human-influenced systems." In this regard, the common knowledge of participants and practitioners, which can be accessed through open publication, is extremely valuable in terms of contextualizing and advancing scholarship. Effective dissemination also guarantees that new work is subjected to the broadest possible academic criticism, enhancing problem-solving, fact-checking, research accuracy, and the emergence of multiple scholarly perspectives. Finally, integrating impact-conscious distribution with open access publishing can encourage researchers to look into other open scientific efforts [Dunleavy \(2022\)](#).

The Open Science movement is closely tied to the rebirth of psychology [Nelson et al. \(2018\)](#). Open Science and enhanced transparency are frequently promoted as answers to issues like poor reproducibility and shady research techniques. For instance, open data and materials let other researchers do experiments again, examine how much results are influenced by decisions made during the analysis phase, and aid the field in determining the findings' replicability and generalizability. P-hacking, harking, and publishing bias can all be mitigated using the open methodology, and preregistration in particular. As a result, Open Science approaches can aid in the reduction of numerous biases in research.

There's no reason why such collaborative and open research endeavours in psychology can't be accomplished. Psychologists, for example, can exchange norms, databases, and other data instead of DNA sequences or sunspots. To accept Open Psychological Science, researchers do not need a "reproducibility" or "integrity issue." Indeed, current tools and emerging technology allow sharing data and content easier than ever before, with potentially significant benefits. In situations where resources are scarce, Open Psychological Science, for instance, can assist in preventing duplication of research efforts. Because of data sharing, later combining, and integration, they can have larger and more varied data sets in the end.

4. DATA SHARING

Data sharing is the practice of making data available in a manner that may be used by others. Sharing can take several forms, including confidential discussions between researchers, uploading datasets on personal or lab websites, depositing them in archives, repositories, subject-specific collections, or library collections, and including data as additional materials in journal papers [Wallis et al. \(2013\)](#).

The process of sharing de-identified individual patient data underpinning the conclusions given in scientific journals with other researchers is known as data sharing [Barbui et al. \(2016\)](#). The primary motivation to share data is accountability. By gaining access to the raw data underlying the results presented in an article, other researchers can repeat the analyses the study authors presented, plan new analyses to address the same research question, and confirm (replicate) the main findings or raise questions about their robustness and validity under various analytical or statistical hypotheses. Another incentive to share the data is that it may be used by other researchers to answer various research questions. A third reason is that systematic reviews, meta-analyses, and meta-analyses of individual-patient data can benefit from sharing datasets of similarly obtained data.

4.1. BENEFITS

Given that science is a human endeavour, psychology research can contribute to discussions on how to create scientific research methods and artefacts that address privacy issues and even encourage data altruism among participants. The dialogue will bridge disciplines more deeply if open science discussions are combined with data presented by psychologists who are both researchers and practitioners [Hesse \(2018\)](#). For the aims of openness, reproducibility, and greater rigour, the open science makes psychological researchers' data, code, and resources widely accessible to other psychologists. When combined with survey findings that directly measure open scientific behaviours, the evidence points to a dramatic movement in psychology norms toward open science [Nosek \(2019\)](#).

Sharing data benefits the greater scientific community by encouraging various perspectives, helping to uncover errors, inhibiting fraud, is valuable for teaching new researchers, and avoids repeat data collecting, resulting in more efficient use of money and patient population resources [Piwowar & Vision \(2013\)](#). Archived data allows for comparisons throughout time and the investigation of a wide range of research challenges. Data sharing can also help researchers improve their methodologies. Developing methodology, software, and technology, in turn, provide new study opportunities.

4.2. ISSUES AND CHALLENGES

Despite progress in many areas in adopting open scientific ideas, numerous barriers remain before the technique is widely adopted in the social sciences. Practicality is one factor. Other disciplines have described grand and ambitious initiatives to build inclusive repositories for sharing data sets, but after the skeletal repository was established and the call for donated data was made, none were sent [Nelson \(2009\)](#). There were several useful factors to take into account. Finding the data sets that accompanied old studies, getting them ready for distribution, and then finishing the operational process of uploading the data sets into the right repositories were typical challenges for researchers.

When exchanging data, legal and ethical problems must be considered. Participants in the empirical study must provide informed consent for data reuse; data may contain sensitive information; therefore, anonymization is required. Commercial interests, proprietary rights, and copyrights are frequently implicated. Data management necessitates forethought and the use of proper metadata [Darch et al. \(2017\)](#). The context of data generation and/or collection, the goal of data creation/collection, storage format, and access permissions are all critical pieces of information for data reuse [Jones & Alexander \(2018\)](#), [Shrout & Rodgers \(2018\)](#).

Maintaining privacy and confidentiality, giving credit to those who carried out the study and collected the data as stated by Longo & Drazen [Longo & Drazen \(2016\)](#), the expense of creating trustworthy data repositories, and the additional work and expense for researchers who may be required to develop datasets suitable for use by others and possibly pay for hosting the data in a repository are some of the additional challenges associated with sharing data [Gewin \(2016\)](#). Since new analyses cannot be pre-planned and may therefore be restricted by data guidance, conducting new studies on shared datasets may pose possible scientific difficulties.

4.3. SOLUTIONS

Researchers must pave the way for our communities to achieve the objective of an open, transparent science environment, while also using scientific evidence for the mission of catalysing desired change across all disciplines.

However, several strategies can be used to solve these problems. A data sharing plan should be included in the study protocol as a starting point for researchers to consider data sharing when they begin their research projects. It may involve a process for gaining informed consent that covers a data sharing clause, protecting patient privacy and confidentiality when data is shared, and providing information on what is intended to be shared, how it will be shared, and a schedule [Sarpawari et al. \(2014\)](#), [El Emam et al. \(2015\)](#). This is significant because, depending on the type of data collected, different methods for sharing raw data may be appropriate (for example, raw data may occasionally be published in the primary publication or in supplemental supporting files, and occasionally a web-based repository is required [Barbui et al. \(2016\)](#)).

Finally, while the social sciences share the normative expectation that research data must be shared to facilitate replication and reanalysis, there is scant evidence that this is a widespread practice. With implicit and explicit sharing rules, federal institutions and professional groups reinforce these normative expectations. The benefits of sharing data with other researchers are numerous and cumulative. As previously stated, there are significant institutional, financial, and career hurdles to data sharing. The amount of data sharing among social scientific disciplines, as well as its benefit to the social sciences, remains an open empirical subject [Pienta et al. \(n.d.\)](#).

5. DATA MANAGEMENT

Although it is not a goal in and of itself, effective data management is a vital conduit for knowledge discovery, innovation, data and knowledge integration, and community reuse after the data is public [Roche et al. \(2015\)](#).

5.1. DIFFERENT DATA IN DIFFERENT FIELDS

Data management infrastructure is required to share research data. The research data varies widely by discipline, yet in the digital age, there is no clear distinction between quantitative and qualitative sciences. However, various data types demand various storage and access methods. Data ownership and usage rights are also a concern, in addition to data formats like text or numeric data. Discipline-specific data practices and research methods have an impact on management solutions.

The handling of data by researchers is outlined in a data-management plan, which covers everything from text, spreadsheets, photographs, recordings, models, algorithms, and software to producing, disseminating, and safeguarding research data of any form. Whether the information originates from sophisticated scientific apparatuses like particle accelerators or imaging technologies, or from straightforward field observations, makes little difference [Schiermeier \(2018\)](#).

In the open-science era, data management will unavoidably become a necessary skill. High-quality digital publications are created as a result of sound data management and stewardship, which aid in streamlining and facilitating the ongoing process of discovery, assessment, and reuse in subsequent studies. On the other hand, the definition of "good data management" is mostly ambiguous and left to the owner of the data or repository. As a result, it would be very helpful to develop fundamental guidelines to instruct those who publish and/or preserve scholarly data, as well as to clarify the objectives and requirements of successful data administration and stewardship [Wilkinson et al. \(2016\)](#).

5.2. FAIR PRINCIPLES

A set of guiding principles for making data Findable, Accessible, Interoperable, and Reusable are well-described and explained in The FAIR Guiding Principles for Scientific Data Management and Stewardship [Wilkinson et al. \(2016\)](#). These FAIR Guiding Principles come before implementation decisions and don't recommend any specific technology, standard, or implementation solution; they're also not a standard or specification in and of itself. To encourage the most efficient use of research data, they directly address data publishers and producers. In addition to proper collection, annotation, and storage, data stewardship encompasses the idea of "long-term care" for significant digital assets with the intention of their discovery and reuse for study in the future, either separately or in conjunction with newly generated data [Wilkinson et al. \(2016\)](#). The Association of European Research Libraries (LIBER), which held its annual conference in 2017, and the G20 Summit in 2016 both supported FAIR [Murphy \(2018\)](#).

In terms of data deposition, exploration, sharing, and reuse—both manually and automatically—the FAIR Guiding Principles address a number of challenges for contemporary data publication situations. The elements of the FAIR Principles are related but separate. In order to facilitate third-party discovery and reuse, the Principles specify characteristics that contemporary data resources, tools, vocabularies, and infrastructures should have. The entry hurdle for data producers, publishers, and stewards who want to make their data holdings FAIR is kept as low as possible by outlining each guiding principle as simple as is practical. The Principles can be applied in any order as the "FAIRness" of data providers' publication settings increases. Additionally, the modularity of the Principles and the

distinction between data and metadata expressly permit a wide range of peculiar circumstances.

5.3. SUPPORTING DISCOVERY THROUGH DATA HANDLING

Distinct research communities have different conventions and practices when it comes to data handling. In comparison to smaller research projects, collaborative larger research endeavours demand quite different approaches since powerful accelerator facilities generate vast amounts of experimental data. A researcher should save any information that can be utilized to support their claims and conclusions. As is the case with purely theoretical scientific or conceptual work, a data-management plan may not be required if a project does not generate or reuse any data.

Research data that has been stored needs to be properly accompanied by metadata that explains its origin and purpose so that people can access, read, and grasp it. Researchers should get in touch with the library services at their host institute if they have any issues with the metadata requirements or which protocols and digital archives to use for their data [Schiermeier \(2018\)](#). Researchers everywhere have the right to reach their own conclusions regarding published science thanks to open access to research data. In the event that their findings cannot be replicated by other researchers or if moral or legal questions arise after a study has been published, scientists should cling onto their data.

Data and related information are created, archived, and organized in such a way that data remains accessible and reliable, and data safety and security are maintained throughout the data life cycle, according to research data management. Not all sorts of data and records can be openly shared. Patient information and medical records, for example, are typically anonymized. The same is true of numerous interview recordings used in empirical social research, including polls on politics and studies of individual behaviour. Any restrictions on confidentiality or copyright, for example, must be stated in data-management strategies. These could be related to collaborations between university scientists and researchers in the private sector. When creating their strategy, researchers should think about data privacy and ethical issues and any ethical, legal, or other limits.

6. USE AND REUSE OF DATA

The term "reuse" refers to the use of data that was obtained by someone else for a different purpose. Reusers of data frequently employ research approaches that differ from those employed by the original data providers. Secondary analysis is another term for data reuse. It's important to understand the difference between using and reusing research data. The former refers to using primary source data for the project or purpose for which it was collected; the latter, to using secondary source data or data collected for objectives unrelated to the current one [Late & Kekäläinen \(2020\)](#).

An archive of social science research data was examined in terms of usage and users, and it was discovered that the archive is actively used, particularly for educational purposes. Users come from many main fields, even while the research data archive investigated centred on social sciences. Because sharing of research data, reuse, and citation in the social sciences are continuously expanding, more research is needed [Late & Kekäläinen \(2020\)](#).

One of the most significant but elusive elements of data reuse is the ability to trust information obtained from others. Scientific practice is predicated on the capacity to believe the claims and outputs of other people's knowledge, or "epistemic trust". Epistemic trust is relational, has several dimensions as opposed to being inherent in a dataset. One component is interpersonal trust, such as trust in the team that created a dataset [Prieto \(2009\)](#). Other aspects of trust include the ability to assess data quality, the reputations of archives that house relevant datasets, and the organizations in charge of data curation [Borgman et al. \(2019\)](#), [Faniel & Jacobsen \(2010\)](#).

6.1. MAKING QUANTITATIVE AND QUALITATIVE DATA OPEN AND REUSE

All supporting data for published quantitative research papers and books should be easily accessible, along with details on their definitions, coding, and analyses, so that any of them can be checked if questions are raised. Most public and non-profit grant-funders are increasingly demanding that research data be made available in open archives when investigations are concluded (with clear instructions and information in formats that are accessible).

The propensity for multi-team and multi-national research has facilitated the development of reusable and comparable datasets covering several nations or datasets that have pre-defined common questions that may be 'pooled' into more complete or authoritative resources. New research that better considers the best/strongest existing research in an area may be necessary to make more advancements (employing systematic reviews, for instance). Less silo-bound disciplinary methodologies would be encouraged and made possible by the development of an open data culture in quantitative social research.

Additional researchers should be able to replicate the findings of one team using the same datasets and analysis methodologies. The Replication archive (RA) is required by many quantitative journals before they may publish quantitative research. A culture of openness has progressed even further in academic publications that have appointed Data Editors for evaluating the standard, breadth, and use of RAs. In this situation, it might be beneficial to use journaling, lab notebooks, and more systematically record searches, coding analyses, and other duties [Dunleavy \(2022\)](#).

Qualitative researchers have made far less progress in integrating open science methodologies creatively in several social science disciplines. Making knowledge, information, and data derived from archives, text sources, documentation, interviews, images, and other sources easily accessible for scrutiny or re-use presents undeniably more logistical challenges for a number of reasons. Restrictions imposed by the owner or museums; major documentary materials not available in digital form; information made available "off the record" or "non-attributably" protecting the anonymity of study respondents, patients, or anyone affected by problems being examined are some examples [Dunleavy \(2022\)](#).

Different teams of social scientists should produce consistent conclusions when they employ various analysis approaches to examine the same types of datasets or bodies of evidence. Greater triangulation of multiple research methodologies focusing on the same phenomenon might lead to higher levels of assurance that the results are not fictitious. In order to calibrate and modify study findings with confidence to anticipate outcomes in a number of situations, more social science

information may be applied across a greater range of social contexts and historical eras.

6.2. IMPLICATIONS OF DATA REUSE

Re-analyzing one or more datasets to respond to new research questions is an example of data reuse. Other examples include going back to one's own data for future comparisons, getting datasets from public or private sources to compare to freshly collected data, looking through available datasets as background research for a new project, and so forth. The implications of these activities for scientific practice, data archive architecture, public policy, and data science are all very different [Pasquetto et al. \(2019\)](#).

Researchers receive academic credit for archiving research data. Many multinational science publishers have policies that make data access a requirement for publishing. Citations to historical data can also result in a lot of credit. Articles based on data that is archived obtain more citations than publications based on data that is not shared. This is because, in addition to the study data itself, when academics use archived data for new research, they frequently mention the original data authors' papers.

Duplication of data collecting is avoided by reusing data. It can also reduce the amount of data collected from hard-to-reach or vulnerable individuals. When the only people who know anything about the data are the people who created it, valuable research data is useless to the scientific community and future studies. All information will be lost if they move to different organizations or tasks, or if they retire.

6.3. DATA AND INFRASTRUCTURE

Data for research isn't just found in nature; it's carefully generated for specific research aims. As a result, like all data, research data is a collection of local and historical artefacts. Researchers need context information such as equipment manuals, protocols, data collection and processing techniques, and experimental or laboratory settings of data handling to repurpose data obtained by others [Culina et al. \(2018\)](#).

Metadata schemas and ontologies are tools for formalizing and transferring this type of data [Mayernik \(2015\)](#), [Mayernik & Acker \(2017\)](#). Ontologies are classified by [Leonelli \(2010\)](#) as "relevance labels" while metadata is classified as "reliability labels" that contain numerous "small facts," according to [Pasquetto et al. \(2019\)](#). Together, these mechanisms enable the linking of datasets to particular research objects (such as the biological entity being studied) and the provision of information about the quality of the data, including the data format, the organisms utilized in experiments, the tools and techniques employed, and the lab settings in which the data were gathered.

Data creators typically have the most in-depth understanding of a dataset, having earned this knowledge through the design, collection, processing, analysis, and interpretation of the data. Because several people may be involved in data collection, information may be disseminated throughout time among multiple stakeholders. Subsequently, data reusers search for the necessary information through various channels, including metadata, documentation, conversations with data creators, and other techniques.

Prospective data reusers frequently bridge knowledge gaps by soliciting data creators' assistance in reusing the data and in return, they may acknowledge them as co-authors [Pasquetto \(2018\)](#), [Wallis et al. \(2013\)](#). Questions involving the effect of knowledge gaps on decision making come up in a range of fields, including statistics, psychology, economics, information science, and many more.

Expertise and data don't exist in a vacuum. The best place to start learning about data practices is with knowledge infrastructures, which are strong networks of individuals, objects, and organizations that produce, disseminate, and sustain specialized information about the natural and human worlds as stated by [Pasquetto, et al. \(2019\)](#). Beyond the actual research, exchanging data between individuals and laboratories frequently necessitates additional effort, knowledge, and price. In addition to scientific expertise, infrastructure for finding, retrieving, understanding, and utilizing datasets is necessary for the creation, processing, and distribution of datasets [Karasti & Blomberg \(2017\)](#), [Borgman \(2015\)](#).

Data sharing and reuse are also impacted by disagreements over data ownership and governance, according to [Pasquetto et al. \(2019\)](#). Domain, technique, and data types all affect community norms. To prevent "free riders" from undermining community standards, openness necessitates governance, whether for shared grazing lands or data repositories [Hess & Ostrom \(2007\)](#). Some individuals are referred to be "data parasites" for using other people's data without giving proper acknowledgement to the original creators [Longo & Drazen \(2016\)](#).

7. CONCLUSION

The endeavour known as "Open Science," which aims to make scientific data and research accessible to everybody, is based on data sharing. At every stage of the research lifecycle, there are numerous services that encourage open scientific behaviours. It is expected that this behaviour will continue among scientists, as even those who have never made data publically available are now considering it.

While typical science levels of agreement on the preeminent scientific paradigm are attained inside or behind the frontier, making new discoveries carries the greatest scientific prestige there. On the other hand, social scientific research has a bad reputation due to strong theoretical/ideological disagreements about what constitutes fundamentals and a large turnover of ephemeral findings or analytical styles [Dunleavy \(2022\)](#). A trend known as "developing open social science" has the potential to refocus social science research toward more reliable breakthroughs in knowledge, while simultaneously fostering a greater degree of social science consensus on the comprehension of important societal processes. Since sharing, reusing, and citing of research data in the social sciences are always increasing, more research is required.

CONFLICT OF INTERESTS

None.

ACKNOWLEDGMENTS

None.

REFERENCES

- Baker, K. S., Duerr, R. E., & Parsons, M. A. (2015). Scientific Knowledge Mobilization : Co-evolution of Data Products and Designated Communities. *International Journal of Digital Curation*, 10(2), Article 2. <https://doi.org/10.2218/ijdc.v10i2.346>.
- Barbui, C., Gureje, O., Puschner, B., Patten, S., & Thornicroft, G. (2016). Implementing a Data Sharing Culture. *Epidemiology and Psychiatric Sciences*, 25(4), 289–290. <https://doi.org/10.1017/S2045796016000330>.
- Borgman, C. L. (2015). *Big Data, Little Data, No Data : Scholarship in the Networked World*. <https://doi.org/10.7551/mitpress/9963.001.0001>.
- Borgman, C. L., Scharnhorst, A., & Golshan, M. S. (2019). Digital Data Archives as Knowledge Infrastructures : Mediating Data Sharing and Reuse. *Journal of the Association for Information Science and Technology*, 70(8), 888–904. <https://doi.org/10.1002/asi.24172>.
- Culina, A., Crowther, T. W., Ramakers, J. J. C., Gienapp, P., & Visser, M. E. (2018). How to do Meta-Analysis of Open Datasets. *Nature Ecology & Evolution*, 2(7), 1053–1056. <https://doi.org/10.1038/s41559-018-0579-2>.
- Curty, R. G., Crowston, K., Specht, A., Grant, B. W., & Dalton, E. D. (2017). Attitudes and Norms Affecting Scientists' Data Reuse. *PLOS ONE*, 12(12). <https://doi.org/10.1371/journal.pone.0189288>.
- Darch, P. T., & Knox, E. J. M. (2017). Ethical Perspectives on Data and Software Sharing in the Sciences : A Research Agenda. *Library & Information Science Research*, 39(4), 295–302. <https://doi.org/10.1016/j.lisr.2017.11.008>.
- David, P., Esanu, J. M., & Uhlir, P. F. (2003). The Economic Logic of “Open Science” and the Balance Between Private Property Rights and the Public Domain in Scientific Data and Information : A Primer. In *the Role of Scientific and Technical Data and Information in the Public Domain : Proceedings of a Symposium*. National Academies Press (US).
- Dunleavy, P. (2022, January 11). Eight Components for ‘Open Social Science’ – An Agenda for Cultural Change. *Impact of Social Sciences*.
- El Emam, K., Rodgers, S., & Malin, B. (2015). Anonymising and Sharing Individual Patient Data. *BMJ*, 350(mar2001), h1139–h1139. <https://doi.org/10.1136/bmj.h1139>.
- Faniel, I. M., & Jacobsen, T. E. (2010). Reusing Scientific Data : How Earthquake Engineering Researchers Assess the Reusability of Colleagues' Data. *Computer Supported Cooperative Work (CSCW)*, 19(3–4), 355–375. <https://doi.org/10.1007/s10606-010-9117-8>.
- Gewin, V. (2016). Data Sharing : An Open Mind on Open Data. *Nature*, 529(7584), 117–119. <https://doi.org/10.1038/nj7584-117a>.
- Hand, D. J., Mannila, H., & Smyth, P. (2001). *Principles of Data Mining*. MIT Press.
- Hess, C., & Ostrom, E. (2007). *Understanding Knowledge as a Commons : From Theory to Practice*. MIT Press.
- Hesse, B. W. (2018). Can Psychology Walk the Walk of Open Science ? *American Psychologist*, 73, 126–137. <https://doi.org/10.1037/amp0000197>.
- Jones, K., & Alexander, S. M. (2018). Qualitative Data Sharing and Re-Use for Socio-Environmental Systems Research : A Synthesis of Opportunities, Challenges, Resources and Approaches. <https://doi.org/10.13016/M2WH2DG59>.
- Karasti, H., & Blomberg, J. (2017). Studying Infrastructuring Ethnographically. *Computer Supported Cooperative Work (CSCW)*, 27(2), 233–265. <https://doi.org/10.1007/s10606-017-9296-7>.

- Late, E., & Kekäläinen, J. (2020). Use and Users of a Social Science Research Data Archive. *PLOS ONE*, 15(8). <https://doi.org/10.1371/journal.pone.0233455>.
- Leonelli, S. (2010). Packaging Small Facts for Re-Use : Databases in Model Organism Biology. In M. S. Morgan & P. Howlett (Eds.), *How Well Do Facts Travel ? The Dissemination of Reliable Knowledge* –348). Cambridge University Press. <https://doi.org/10.1017/CBO9780511762154.017>.
- Longo, D. L., & Drazen, J. M. (2016). Data Sharing. *New England Journal of Medicine*, 374(3), 276–277. <https://doi.org/10.1056/NEJMe1516564>.
- Mayernik, M. S. (2015). Research Data and Metadata Curation as Institutional Issues. *Journal of the Association for Information Science and Technology*, 67(4), 973–993. <https://doi.org/10.1002/asi.23425>.
- Mayernik, M. S., & Acker, A. (2017). Tracing the Traces : The Critical Role of Metadata Within Networked Communications. *Journal of the Association for Information Science and Technology*, 69(1), 177–180. <https://doi.org/10.1002/asi.23927>.
- Murphy, F. (2018). Open Access, Open Data, FAIR Data and their Implications for Life Sciences Researchers. *Emerging Topics in Life Sciences*, 2(6), 759–762. <https://doi.org/10.1042/ETLS20180163>.
- Nelson, B. (2009). Data Sharing : Empty Archives. *Nature*, 461(7261), 160–163. <https://doi.org/10.1038/461160a>.
- Nelson, L. D., Simmons, J., & Simonsohn, U. (2018). Psychology's Renaissance. *Annual Review of Psychology*, 69(1), 511–534. <https://doi.org/10.1146/annurev-psych-122216-011836>.
- Nosek, B. (2019). The Rise of Open Science in Psychology, A Preliminary Report. Open Knowledge Foundation (2015). Open Definition : Defining Open in Open Data, Open Content and Open Knowledge. (n.d.). Retrieved from 2023, May 13.
- Pasquetto, I. (2018). PhD Dissertation - From Open Data to Knowledge Production : Biomedical Data Sharing and Unpredictable Data Reuses.
- Pasquetto, I. V., Borgman, C. L., & Wofford, M. F. (2019). Use and Reuses of Scientific Data : The Data Creators' Advantage : Supplementary Materials. *Harvard Data Science Review*, 1(2). <https://doi.org/10.1162/99608f92.af83430a>.
- Pasquetto, I. V., Randles, B. M., & Borgman, C. L. (2017). On the Reuse of Scientific Data. *Data Science Journal*, 16(8), 1–9. <https://doi.org/10.5334/dsj-2017-008>.
- Pienta, A., Alter, G., & Lyle, J. (n.d.). The Enduring Value of Social Science Research : The Use and Reuse of Primary Research Data : Paper - iPRES 2011 - Singapore. Retrieved from 2023, May 13.
- Piwovar, H. A., & Vision, T. J. (2013). Data Reuse and the Open Data Citation Advantage. *PeerJ*, 1, e175. <https://doi.org/10.7717/peerj.175>.
- Prieto, A. G. (2009). From Conceptual to Perceptual Reality : Trust in Digital Repositories. *Library Review*, 58(8), 593–606. <https://doi.org/10.1108/00242530910987082>.
- Roche, D. G., Kruuk, L. E. B., Lanfear, R., & Binning, S. A. (2015). Public Data Archiving in Ecology and Evolution : How Well Are We Doing ? *PLOS Biology*, 13(11). <https://doi.org/10.1371/journal.pbio.1002295>.
- Sarpatwari, A., Kesselheim, A. S., Malin, B. A., Gagne, J. J., & Schneeweiss, S. (2014). Ensuring Patient Privacy in Data Sharing for Postapproval Research. *New England Journal of Medicine*, 371(17), 1644–1649. <https://doi.org/10.1056/NEJMs1405487>.
- Schiermeier, Q. (2018). Data Management Made Simple. *Nature*, 555(7696), 403–405. <https://doi.org/10.1038/d41586-018-03071-1>.

- Schmidt, B., Gemeinholzer, B., & Treloar, A. (2016). Open Data in Global Environmental Research : The Belmont Forum's Open Data Survey. *PLOS ONE*, 11(1). <https://doi.org/10.1371/journal.pone.0146695>.
- Shrout, P. E., & Rodgers, J. L. (2018). Psychology, Science, and Knowledge Construction : Broadening Perspectives from the Replication Crisis. *Annual Review of Psychology*, 69(1), 487–510. <https://doi.org/10.1146/annurev-psych-122216-011845>.
- Tenopir, C., Allard, S., Douglass, K., Aydinoglu, A. U., Wu, L., Read, E., Manoff, M., & Frame, M. (2011). Data Sharing by Scientists : Practices and Perceptions. *PLoS ONE*, 6(6). <https://doi.org/10.1371/journal.pone.0021101>.
- Tenopir, C., Dalton, E. D., Allard, S., Frame, M., Pjesivac, I., Birch, B., Pollock, D., & Dorsett, K. (2015). Changes in Data Sharing and Data Reuse Practices and Perceptions among Scientists Worldwide. *PLOS ONE*, 10(8). <https://doi.org/10.1371/journal.pone.0134826>.
- Tenopir, C., Rice, N. M., Allard, S., Baird, L., Borycz, J., Christian, L., Grant, B., Olendorf, R., & Sandusky, R. J. (2020). Data Sharing, Management, Use, and Reuse : Practices and Perceptions of Scientists Worldwide. *PLOS ONE*, 15(3). <https://doi.org/10.1371/journal.pone.0229003>.
- Wallis, J. C., Rolando, E., & Borgman, C. L. (2013). If We Share Data, Will Anyone Use Them ? Data Sharing and Reuse in the Long Tail of Science and Technology. *PLoS ONE*, 8(7). <https://doi.org/10.1371/journal.pone.0067332>.
- Wilkinson, M. D., Dumontier, M., Aalbersberg, Ij. J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.-W., Da Silva Santos, L. B., Bourne, P. E., Bouwman, J., Brookes, A. J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C. T., Finkers, R., ... Mons, B. (2016). The FAIR Guiding Principles for Scientific Data Management and Stewardship. *Scientific Data*, 3(1), 160018. <https://doi.org/10.1038/sdata.2016.18>.