



NATIONAL DATA SHARING POLICY IN THE ERA OF OPEN SCIENCE: A SYSTEMATIC LITERATURE REVIEW

Cahyo Trianggoro* & Tupan**

*Graduate School Universitas Padjajaran

**National Research and Innovation Agency

Email: cahyo.trianggoro@brin.go.id, tupan712190@yahoo.com

(Submitted: 05-08-2021, Revised: 01-11-2021, Accepted: 24-11-2021)

DOI: [10.24252/v9i2a2](https://doi.org/10.24252/v9i2a2)

ABSTRAK: Kegiatan berbagi data penelitian memberikan banyak manfaat bagi ekosistem penelitian. Namun dalam konteks Indonesia, kurangnya kebijakan dalam mengatur mekanisme berbagi data penelitian membuat peneliti enggan melakukannya. Pendana penelitian dan lembaga penelitian memainkan peran penting dalam mengembangkan kebijakan berbagi data. Penelitian terkait kebijakan berbagi data penelitian menjadi penting dalam rangka merancang kebijakan untuk mendorong praktik berbagi data penelitian. Dengan kajian literatur tersistematis artikel ini melihat bagaimana kebijakan berbagi data dirumuskan dan diimplementasikan di berbagai lembaga penelitian. Data diambil dari pengindeks Scopus dan Dimension menggunakan kosakata kontrol. Peran lembaga penelitian dan penyandang dana serta instrumen kebijakan dianalisis untuk melihat pola yang terjadi di antara para pihak. Kami memeriksa 23 artikel yang berisi kebijakan berbagi data. Ditemukan bahwa penyandang dana memiliki peran terbesar dalam menentukan desain kebijakan berbagi data. Penyandang dana memandang bahwa data penelitian merupakan aset dalam penelitian yang didanai oleh dana publik sehingga manfaatnya harus dikembalikan kepada masyarakat. Lembaga penelitian berperan sebagai penyedia infrastruktur penelitian yang berkontribusi dalam pembuatan data. Sementara itu, peneliti sebagai pelaku penelitian perlu memberikan masukan dalam mengembangkan mekanisme berbagi data dan mengatur aspek sensitivitas data dan aspek hukum dalam berbagi data penelitian.

Kata kunci: Pengetahuan terbuka; manajemen data penelitian

ABSTRACT: Research data sharing activities provide many benefits to the research ecosystem. However, in the Indonesian context, there is a lack of policy in regulating research data sharing mechanisms which makes researchers reluctant to undertake the practice of data sharing. Research funders and research institutions play a critical role in developing data-sharing policies. Research related to the policy of research data sharing is important in order to design policies to encourage the practice of research data sharing. With systematic literature review, this article is to see how data-sharing policies were formulated and implemented in various research institutions. The data were taken from Scopus and Dimension indexers using controlled vocabulary. The roles of research institutions and funders as well as policy instruments were analysed to see patterns that occur between the parties. We examine 23 articles containing data sharing policies. It was found that the funders have the greatest role in determining the design of the data sharing policy. Funders view that research data is an asset in research funded by public funding so that the benefits must be returned to the community. Research institutes play a role as a provider of research infrastructure that contributes to data creation. Meanwhile, researchers as research actors need to provide input in developing data sharing mechanisms and regulating data sensitivity aspects and legal aspects in research data sharing.

Keywords: Open science; research data management



1. INTRODUCTION

The new regulation on the national science and technology system (SINAS IPTEK) Number 11 of 2019 which was recently issued by the government becoming a legal basis that strengthens the role and position of research and development to achieve the national vision and to increase the competitiveness of the nation. The new regulations related to science and technology are trying to respond to changes in the open research ecosystem that in line with the open science movement. The regulation stipulates the obligation to deposit data in as mandated in article 40 for funding agency, researchers, as well as science and technology institutions. However, since the the regulation is mandated, the implementation of research data management within the research institution are relatively low.

The age of open science is a period that encourages every single stakeholder in the research and development ecosystem to share their research resources: Open Infrastructure, Open Access for Published Material, Open Data, and Open Code. In the era of Open Science, data sharing has become a common practice and ultimately encourages collaboration between researchers within their scientific discipline or across disciplines. Many benefits can be achieved from the practice of data sharing, especially for developing countries like Indonesia where the research budget is less than 1% of GDP, such as Efficiency of research activities (Higman & Pinfield, 2015); Encouraging innovation & collaboration between stakeholders (Zenk-Möltgen, Akdeniz, Katsanidou, Naßhoven, & Balaban, 2018); Minimize redundancy of research activities, The openness of data encourages researchers to produce credible research. Open research data allows for more public peer review of the data generated by researchers (Higman & Pinfield, 2015).

In previous research, several factors that encourage researchers to share research data have been investigated (Zenk-Möltgen et al., 2018). As a result, many researchers have awareness and desire to share their data. on the other hand, they feel confused whether this kind of practice is a legal practice and will avoid the process of scooping research activities (Hedlund et al., 2020; Higman, R., & Pinfield, 2015). Despite of several benefits and desire of researchers who have the desire to share data, the lack of a precise mechanism related to data sharing policies and protocols for researchers to share research data is a barrier for researchers to implement data sharing (Couture et al., 2018; Higman, R., & Pinfield, 2015; Neylon, 2017)). There are several obstacles for researchers that motivate researchers' hesitation to share data as conveyed in research conducted by Pinfield. There were related intellectual property rights aspect, commercial interests, and ethical issues. Furthermore, concerns about data misuse also become an obstacle to data sharing, when research data is made for the public (Fusi, Manzella, Louafi, & Welch, 2018). Therefore, it is urgent for Indonesia to immediately design a policy related to a research data-sharing mechanism for academics and researchers and to develop a protocol that can be used as a guide for researchers who wish to share research data to be able to answer these problems.

Therefore, it is very urgent for Indonesia to design a policy related to a research data-sharing mechanism for academics and researchers and to develop a protocol that can be used as a guide for researchers who had motivations to share research data immediately (Cheah, P. Y., Day, N. P. J., Parker, M., & Bull, 2017). Data sharing policies should be designed to find out a balance between guaranteeing the protection of the intellectual property of data owners and the public's rights to funds that have been issued by the government in research activities through data sharing policies. The research questions are focusing on answering What is the role of the stakeholders in encouraging data sharing practices in the light of data sharing policy design?

This research is significant and urgently to considering that the open science movement has developed throughout the world including Indonesia and Indonesia needs to take advantage of this momentum to improve the quality of research through policies that are in line with existing developments. The result of this study contributes to the formulation of policies related to data

sharing practices, and the clarity of policies and protocols could encourage researchers to share research data can contribute to fostering sustainable innovation in the end.

2. PREVIOUS FINDINGS

Many previous studies have conducted research related to data sharing policies. However, the research still focuses on policy in the context of data sharing in scientific journals publishers setting (Bates, 2017; Read et al., 2021; B. Suhr et al., 2020). Availability of research data that underlying the publication manuscript allowing the publishers to review the dataset as well as the manuscript. Therefore, those policy play important roles to ensure the rigorous of research through evaluate the methodology by using the dataset (Read et al., 2021). Research related to data sharing in the context of institutional policies through involving all stakeholders within the research ecosystem needs to be carried out as recommended from previous research in order to find the most optimum for data sharing governance (Mason et al., 2020; Zenk-Möltgen et al., 2018). The success in implementing the data sharing policy is a collective work among the stakeholders, both funding agency, researchers as the main actors in research activities, and research institutions that provide infrastructure and data services in the practice of sharing research data. Each party plays an important role and contributes to the success of the policy. Policy designs that regulate the roles and contributions of each stakeholder as well as policy instruments, are the focus of this research.

3. RESEARCH METHODOLOGY

Qualitative approach with systematic Literature Review is carried out to conduct this research to investigate data sharing policy that has been implemented by several research institutions. The study approach was involving five respectively steps that adopted from studied has been used by (Zuiderwijk et al., 2020): 1. Studies Identification; 2. Article Selection; 3. Data Extraction; 4. Synthesis;

We collected data from Scopus during 03 August 2021 until 05 August 2021, the coverage article published in the period from January 2017 to August 2021 and from dimension database from January 2017 to August 2021. We sort manually by analysing the article from abstract to find out the similar works. The exclude criteria. Some key words employee for search strategy are:

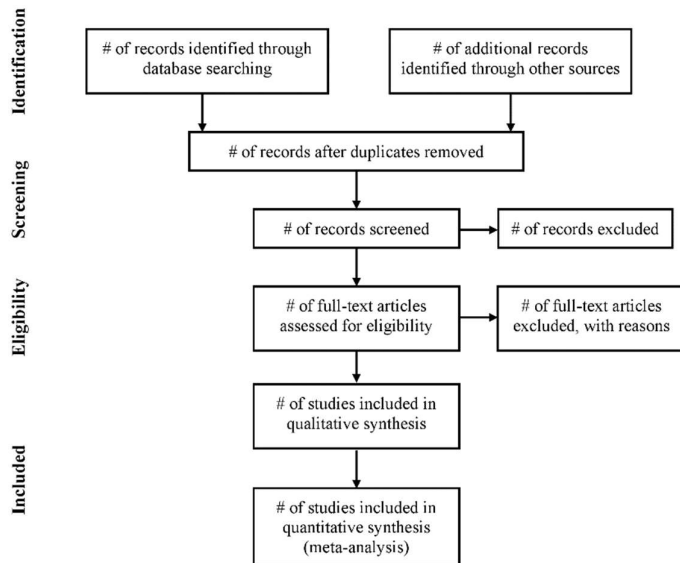
1. Scopus Database

Keyword	Field
Data Sharing	Title
Policy	Title

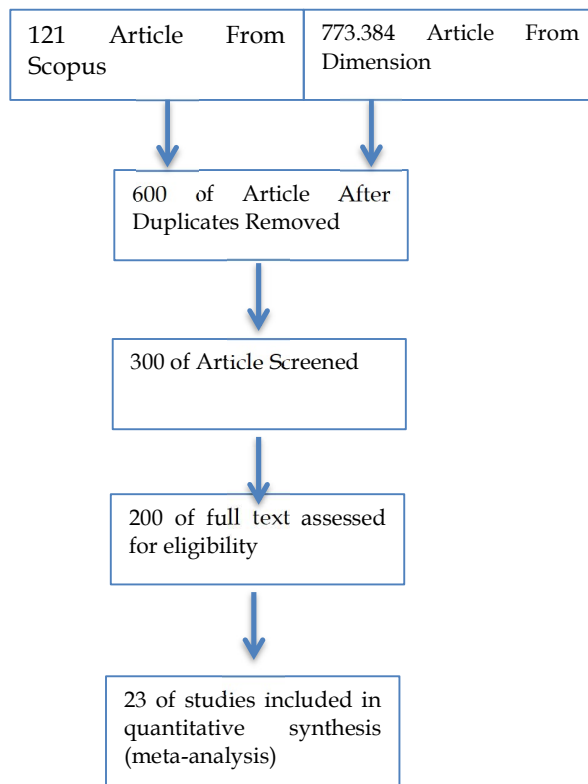
2. Dimension Database

Keyword	Field
Data Sharing	Title
Policy	Title

From the retrieval strategy, we collected 126 articles from Scopus database and 4561 articles and dataset from dimension database. After retrieval processing, we use Prisma protocol for reporting systematic reviews and meta-analyses studies as described in the diagram below.



Flow of information through the different phases of a systematic review/ doi: <https://doi.org/10.1371/journal.pmed.1000100.g001>



From every single data source, we screening and sorted carefully by reading the article from abstract and the results. We found that only 200 articles are eligible for the studies and have relevant topic in data sharing and policy design. From 200 article, we synthesis 23 manuscript that investigate the stakeholders' roles.

4. RESULTS AND FINDINGS

The policy of data sharing is an important step in the development of science and technology. The utilization of research data by the scientific community encourages research activities to be more efficient. Data on policy formulation is carried out by mapping the interests of stakeholders in the research ecosystem. A systematic literature review is carried out to see how the roles of each stakeholder and the relationship between them.

From the literature that deserves to be synthesized, it is found that it is important that stakeholders who have a role in encouraging the sharing of research data are listed in Table 1.1

Tabel 1. roles of stakeholders in data sharing policy

Stakeholders	Roles	Authors
Research Funders	<ul style="list-style-type: none"> Develop Policy (data sharing requirement) Incentivize researchers, data curator Monitoring and evaluation mechanism with capable staff 	(Waithira et al., 2019)(Blasimme, Fadda, Schneider, & Vayena, 2018; Gaba, Siebert, Dupuy, Moher, & Naudet, 2020;)
Research Institution/Agency	<ul style="list-style-type: none"> Develop data sharing Infrastructure: data repository, hardware, Internet network; Incentivize researchers who make data accessible; Provide Data curators Define career path for Data curators Access control Train data supporters and researchers Enable data sharing take into account for performance appraisal Development Guideline for data sharing 	(Blasimme et al., 2018; Cheah, P. Y., Day, N. P. J., Parker, M., & Bull, 2017; Koščík, 2018; Parida & Tripathi, 2018; Sá-Sousa et al., 2018; Schmidt et al., 2016; Slavnic, 2017; Waithira et al., 2019; Yost, 2016)
Researchers	<ul style="list-style-type: none"> Generate Data Curating Data (deidentification) Protect Interest (property rights) Access control 	(Karafili & Lupu, 2017; Robinson, Slashinski, Chiao, & McGuire, 2016; Waithira et al., 2019;)
Journal Publishers	<ul style="list-style-type: none"> Validating publication Develop policy (data sharing requirement) 	(Waithira et al., 2019;)

Research Funders

Research funders whether it is government or private are one of the key factors that can encourage or driving force researchers to share their research data. In previous research, it was revealed that research funders have a central role in determining policy direction (Arias et al., 2016; Bates, 2017; Gaba et al., 2020; Schmidt et al., 2016). Research funders can design policies

that require researchers who wish to gain access to research funding to implement data management plan at the early stage of research cycle. During research, researchers are enforce to store and share research data through a data repository designated by the research funder and whenever their research project is completed, researchers are encourage to share their data through public data repository (B. Suhr et al., 2020). Moreover, Funding agency could define a mechanism for incentivizes for the researchers who has an effort to make their research data available in public data repository for reusable purpose. Therefore, researchers have a massive motivation to make their data are open as possible.

In addition to the obligation to share research data, funders can also design incentive and disincentive mechanisms for researchers and research supporters who can make research data open to the public. This incentive is highly recommended by previous research as a form of appreciation for efforts to make research data open to the public (Blasimme et al., 2018)

Research Institution

The research ecosystem certainly requires an organization that manages research activities, providing research resources in physical infrastructure and human resources. Systems and procedures in research management are shaded in a research institute. Therefore, research institutions play a key role in the formulation and implementation of research data sharing policies.

As one of the research ecosystem stakeholders, research institutions have a role in facilitating data sharing practices by adding to the data repository infrastructure as a platform for archiving and sharing research data (Schmidt et al., 2016). As a provider of research data management infrastructure, research institutions also play a role in regulating access to research data or providing recommendations to researchers regarding the level of access to research data. In managing the data repository, research institutions perform the data curation process to ensure data quality. Therefore, it requires supporting human resources, specifically managing research data, including curating data, namely data curator.

Considering that the data curator's role is relatively new, research institutions need to invest in human resource development by investing in human resource development programs through training activities. The training is not limited only to data curators/research supporters who support sharing data and the researchers themselves as the parties producing the research data.

Furthermore, data curator's role also needs special attention from research institutions by providing a definite career path, considering that data curators' role is significant in maintaining the quality of research data that will be disclosed to the public. Data quality assurance is one of the data curator's duties that needs to be done so that the public and the scientific community have access to quality research data. Moreover, some literature reveals that the motivation behind data sharing is recognition from research institution that enable data sharing take into account into their performance appraisal (Schmidt et al., 2016).

Moreover, barriers for sharing research data are related to confidentiality and security of their research object. To overcome this issue, research institution could develop guideline for sharing research data therefore researchers would be comfortable whenever data sharing is mandated by funding agency(Ceah, P. Y., Day, N. P. J., Parker, M., & Bull, 2017).

As we can found through the literature, the most important goals of policy is to encourage behavior changes for its object. Data sharing policy aiming to build data sharing culture for the advancement of science development. Therefore, support from research institutions and staff who provide data sharing services is absolutely necessary.

Researchers

As the main actor in research activities, the role of the researcher cannot be ruled out. Researchers play an important role in ensuring the validity of the data they produce. In the data curation process, data supporters communicate intensely with researchers as data producers and, at the same time, also have the role of an owner of research data. As the data owner, the researcher has the right to regulate access to research data as long as it does not conflict with the regulations that apply to the territory where the researcher is located. Furthermore, researchers also have a role in safeguarding intellectual property rights by determining licenses for research data that will be made public. Licensing is important so that researchers' interests as owners of intellectual property over research data are well preserved.

Publishers

As one of the research ecosystem stakeholders, scientific journal publishers have a role as a means to disseminate research results in the form of scientific papers. To ensure the credibility, transparency, and validity of articles, nowadays, many journals adopt policies that require researchers to include research data that has been processed for public access as a supplementary file. Journal publishers could play important roles through their policy by mandating deposit research data that underlying the manuscript. For the early stage of adoption, publishers could make it optional or encourage researchers to make data available for public domain (M. Suhr et al., 2020). Therefore, rigorously of research through it is transparency could be driven by epistemic scientific community.

5. DISCUSSION

Due to implementation data sharing policy, research funders such as LPDP, Ristek-BRIN and others private sectors have the significant roles by driving the data sharing mechanism. Research funders could play important roles to enforce data sharing culture through data sharing policy for each grantee that seeking their funding. On the other hand, research institution could consider to recognize research data as an output that can take into account for performance appraisal therefore researchers have a motivational driver for sharing research data. Moreover, as the concern of researchers to protect confidentiality of their research object, research institution could define a guideline how to sharing the research data therefore it can overcome the barriers for sharing research data.

The development of organization by expanding of monitoring and evaluation that evaluate research data as output could be implemented. Furthermore, through expanding roles of staff to stewarding grantee to share their data through institutional data repository. Transition policy should be adopted and the socialization to every stakeholder should be designed therefore the implementation could running smoothly. Policy in technical level has the critical point therefore staff that implement the policy in daily basis has clear direction.

The formulation and implementation of research data policies need to be carried out by all stakeholders in research activities. Without good collaboration between stakeholders, the ideals of making science more transparent and developing rapidly through data sharing policies will be quite difficult to materialize. Comprehensive policies and support from various parties that encourage researchers to share data are key to success in implementing data sharing policies. In the light of the limitation of this work, there is a gap that we need to full fill to gather data from the stakeholders in Indonesia such as Kemenristek-BRIN as the government representative that leading the research and innovation activity, LPDP as the research funders, and several research institutions that operates in Indonesia. The investigation would give us more comprehensive picture regarding the policy design for data sharing in Indonesia.

REFERENCES

- Arias, J. J., Pham-Kanter, G., & Campbell, E. G. (2016). The growth and gaps of genetic data sharing policies in the United States. *Journal of Law and the Biosciences*, 2(3), 56–68. <https://doi.org/10.1093/jlb/lsv032>
- Bates, J. (2017). The politics of data friction. *Journal of Documentation*, 74(2), 412–429. <https://doi.org/10.1108/JD-05-2017-0080>
- Blasimme, A., Fadda, M., Schneider, M., & Vayena, E. (2018). Data sharing for precision medicine: Policy lessons and future directions. *Health Affairs*, 37(5), 702–709. <https://doi.org/10.1377/hlthaff.2017.1558>
- Cheah, P. Y., Day, N. P. J., Parker, M., & Bull, S. (2017). Sharing individual-level health research data: Experiences, challenges and a research agenda. *Asian Bioethics Review*, 9(4), 393–400. <https://doi.org/10.1007/s41649-017-0029-5>
- Couture, J. L., Blake, R. E., McDonald, G., & Ward, C. L. (2018). A funder-imposed data publication requirement seldom inspired data sharing. *PLoS ONE*, 13(7), 1–13. <https://doi.org/10.1371/journal.pone.0199789>
- Gaba, J. F., Siebert, M., Dupuy, A., Moher, D., & Naudet, F. (2020). Funders' data-sharing policies in therapeutic research: A survey of commercial and noncommercial funders. *PLoS ONE*, 15(8 August), 1–14. <https://doi.org/10.1371/journal.pone.0237464>
- Hedlund, J., Eklund, A., & Lundström, C. (2020). Key insights in the AIDA community policy on sharing of clinical imaging data for research in Sweden. *Scientific Data*, 7(1). <https://doi.org/10.1038/s41597-020-00674-0>
- Higman, R., & Pinfield, S. (2015). Research data management and openness: The role of data sharing in developing institutional policies and practices. *Program*, 49(4), 364–381. <https://doi.org/10.1108/PROG-01-2015-005>
- Karafili, E., & Lupu, E. C. (2017). *Enabling Data Sharing in Contextual Environments*. 231–238. <https://doi.org/10.1145/3078861.3078876>
- Koščík, M. (2018). Institutional rules and policies for sharing and storing research data. *Grey Journal*, 14(Special Winter Issue), 50–54.
- Mason, C. M., Box, P. J., & Burns, S. M. (2020). Research data sharing in the Australian national science agency: Understanding the relative importance of organisational, disciplinary and domain-specific influences. *PLoS ONE*, 15(8 August), 6–8. <https://doi.org/10.1371/journal.pone.0238071>
- Neylon, C. (2017). Building a culture of data sharing: policy design and implementation for research data management in development research. *Research Ideas and Outcomes*, 3, 25–45. <https://doi.org/10.3897/rio.3.e21773>
- Parida, P. K., & Tripathi, S. (2018). Odisha spatial data infrastructure (OSDI) -Its data model, meta data and sharing policy. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 4(5), 21–28. <https://doi.org/10.5194/isprs-annals-IV-5-21-2018>
- Read, K., Campbell, A., Kitchin, V., MacDonald, H., & McKeown, S. (2021). Embracing the value of research data: introducing the JCHLA/JABSC Data Sharing Policy. *Journal of the Canadian Health Libraries Association / Journal de l'Association Des Bibliothèques de La Santé Du Canada*, 42(1). <https://doi.org/10.29173/jchla29536>
- Robinson, J. O., Slashinski, M. J., Chiao, E., & McGuire, A. L. (2016). It depends whose data are being shared: Considerations for genomic data sharing policies. *Journal of Law and the Biosciences*, 2(3), 697–704. <https://doi.org/10.1093/jlb/lsv030>
- Sá-Sousa, A., Fonseca, J. A., Pereira, A. M., Ferreira, A., Arrobas, A., Mendes, A., Drummond, M., Videira, W., Costa, T., Farinha, P., Soares, J., Rocha, P., Todo-Bom, A., Sokolova, A., Costa, A., Fernandes, B., Chaves Loureiro, C., Longo, C., Pardal, C., ... Calvo, T. (2018). The Portuguese Severe Asthma Registry: Development, Features, and Data Sharing Policies. *BioMed Research International*, 2018. <https://doi.org/10.1155/2018/1495039>
- Schmidt, B., Gemeinholzer, B., & Treloar, A. (2016). Open data in global environmental research: The Belmont Forum's open data survey. *PLoS ONE*, 11(1), 1–29.

- <https://doi.org/10.1371/journal.pone.0146695>
- Slavnic, Z. (2017). Research and data-sharing policy in Sweden—neoliberal courses, forces and discourses*. *Prometheus (United Kingdom)*, 35(4), 249–266. <https://doi.org/10.1080/08109028.2018.1499542>
- Suhr, B., Dungal, J., & Stocker, A. (2020). Search, reuse and sharing of research data in materials science and engineering—A qualitative interview study. *PLoS ONE*, 15(9 September), 1–26. <https://doi.org/10.1371/journal.pone.0239216>
- Suhr, M., Lehmann, C., Bauer, C. R., Bender, T., Knopp, C., Freckmann, L., Öst Hansen, B., Henke, C., Aschenbrandt, G., Kühlborn, L. K., Rheinländer, S., Weber, L., Marzec, B., Hellkamp, M., Wieder, P., Sax, U., Kusch, H., & Nussbeck, S. Y. (2020). Menoci: lightweight extensible web portal enhancing data management for biomedical research projects. *BMC Bioinformatics*, 21(1). <https://doi.org/10.1186/s12859-020-03928-1>
- Waithira, N., Mutinda, B., & Cheah, P. Y. (2019). Data management and sharing policy: The first step towards promoting data sharing. *BMC Medicine*, 17(1), 1–5. <https://doi.org/10.1186/s12916-019-1315-8>
- Yost, F. R. (2016). Sharing the data: The information policies of NOAA and EUMETSAT. *IFLA Journal*, 42(1), 5–15. <https://doi.org/10.1177/0340035215611135>
- Zenk-Möltgen, W., Akdeniz, E., Katsanidou, A., Naßhoven, V., & Balaban, E. (2018). Factors influencing the data sharing behavior of researchers in sociology and political science. *Journal of Documentation*. <https://doi.org/10.1108/JD-09-2017-0126>
- Zhu, Y. (2020). Open-access policy and data-sharing practice in UK academia. *Journal of Information Science*, 46(1), 41–52. <https://doi.org/10.1177/0165551518823174>
- Zuiderwijk, A., Shinde, R., & Jeng, W. (2020). What drives and inhibits researchers to share and use open research data? A systematic literature review to analyze factors influencing open research data adoption. In *PLoS ONE* (Vol. 15, Issue 9 September). <https://doi.org/10.1371/journal.pone.0239283>