GESIS Open Science Strategy

gesis Leibniz-Institut für Sozialwissenschaften

GESIS Papers

2020|20

GESIS Open Science Strategy

GESIS Papers 2020|**20**

GESIS Open Science Strategy

GESIS – Leibniz-Institut für Sozialwissenschaften 2020

This document represents a joint effort by many at GESIS - Leibniz Institute for the Social Sciences. A draft version was prepared by Tobias Heycke and Bernhard Miller (editors). Particular thanks for their contributions and advice go to: Arnim Bleier, Matthias Bluemke, Johannes Blumenberg, Johannes Breuer, Hannah Bucher, Fabian Flöck, Agathe Gebert, Reiner Mauer, Peter Mutschke, Klaus Pforr, Jonas Recker, Pascal Siegers, Oliver Watteler, Bernd Weiß and Christof Wolf.

This document was adopted as a GESIS Strategy Paper by the Board of Directors. GESIS reserves the right to amend this strategy at any time.

GESIS Papers

GESIS – Leibniz-Institut für Sozialwissenschaften
Präsidialbüro
Postfach 12 21 55
68072 Mannheim

Telefon: (0621) 12 46 - 257
E-Mail: bernhard.miller@gesis.org

ISSN: 2364-3781 (Online)
Herausgeber,
Druck und Vertrieb: GESIS – Leibniz-Institut für Sozialwissenschaften
Unter Sachsenhausen 6-8, 50667 Köln

1 Perspectives for Open Science

Transparency is a necessary condition for accumulating knowledge, for eliminating errors and for making scientific progress. Open Science can, thus, help achieve these aims. As Europe's largest research infrastructure for the social sciences GESIS has a long tradition in supporting openness. Recent developments in technology have created new possibilities for making research transparent and open; the topic, therefore, warrants revisiting. In this strategy document, we reflect on our current efforts in open science and expand our vision to multiple aspects of open science in the social sciences.

Open Science documents and makes the entire path to scientific knowledge publicly and globally accessible – from data collection, through data analysis, arriving at findings, to the interpretation of these findings in a reproducible manner through the internet (Grötschel 2016). It enjoys considerable support both internationally and across disciplines and is, hence, presumably the most relevant development in contemporary science.

Open Science as an umbrella term refers to many different elements of the research process. As a research infrastructure provider, we structure our Open Science strategy along the research process. The GESIS Open Science strategy focuses on four key areas: Open Methodology, Open Data, Open Access, and Open Source. *Open Methodology*, which refers to the transparent and reproducible documentation of data collection, processing and analysis is at the core of this strategy. In the social sciences, data collection and data analysis are central elements of scientific work. The second part, *Open Data,* refers to the provision of research data with the aim of the greatest possible openness and availability and can be considered a central part of a transparent documentation. Only once data and methodology are open, can results be reproduced. Additionally, Open Data can be re-used and help to address new research questions. *Open Access* refers to the immediate, cost-free and permanent access to scientific publications in sustainable infrastructures on the Internet. Finally, *Open Source* links and complements the other three elements. Open Source software technology is provided with its source code and is freely accessible and modifiable. Open Source, therefore, helps to access methods used in the research process and the data it is based on.

Open Science is not a new notion. Yet, if we look at press statements by German universities and other scientific organizations, mentions of open science are clearly increasing, pointing to increased attention to openness across many different communities. In 2015 around 15 press statements a month mentioned Open Science, in 2020 the number had risen to around 25. At the same time the number of initiatives promoting or supporting Open Science has grown as evidenced by papers of scientific organizations and professional associations on the topic (e.g., DFG 2017, AS 2019, DGS 2019). The call for open access to scientific knowledge is amplified by other, non-scientific organizations' statements and recommendations (e.g., the OECD (2015) and the European Commission (2016)). Finally, to take as witness the potentially most prominent platform for open source development on which scientific data and code are, of course, only one element: Github reported 10 million users in September 2015[1], by the end of 2019 the number had increased to 40 million.[2]

[1] https://en.wikipedia.org/wiki/Timeline_of_GitHub (accessed on 21 October 2020)

[2] https://github.blog/2019-11-06-the-state-of-the-octoverse-2019/ (accessed on 21 October 2020)

The idea of Open Science and the underlying appreciation for sharing to harvest more potential, therefore, is gaining traction. It is in this context that GESIS wishes to put forward a set of ideas and a path forward to help implement them in its domain. At GESIS, we envision that Open Science will be embraced at the level of the scientific system, at the level of institutions, and, most importantly, at the level of the individual researchers. It is individuals who will fill this Open Science vision with life. GESIS pledges its support to all researches embracing Open Science and will strive to serve as a good example for all its employees and the social sciences. We hope that with this strategy – and in concert with all its other protagonists - we can support a more open science.

In any case we pledge to evaluate the progress we make on the path to open science.

2 Path to Open Science

GESIS provides “research-based services for the social sciences”. These services are targeted at researchers who work with methods of empirical social research. With the present strategy, GESIS, as the largest social science infrastructure institution in Europe, commits itself to the desideratum of achieving the greatest possible openness in the scientific process and will extend its services accordingly. Open Science is becoming part of the institute’s culture and engrained in our services because transparency is the basis of fundamental principles of scientific inquiry, such as refutability, intersubjective reproducibility, and cumulative knowledge (Dafoe 2014, AS 2019, DGS 2019, DFG 2017, RatSWD and Allianz der Wissenschaftsorganisationen 2010).

With the present strategy, GESIS reinforces Open Science as a part of the scientific work culture both regarding the services it provides as well as its own research.

We proudly note that the spirit of this strategy is already inspiring efforts to increase openness in the context of research data provided in the social sciences (cf. the German Longitudinal Election Study; Bucher et al. 2021)

2.1 Strategy Objectives

To proceed towards a more open way of conducting science, GESIS defines the following strategy objectives:

- GESIS’ services shall aim to establish an infrastructure supporting Open Science in the social sciences.
- GESIS shall help make different pieces of research more cumulative by linking all elements necessary for understanding the contributions of this research: data, methods, and published conclusions.
- GESIS shall ensure reproducibility for all of its own data-based publications. A benchmark for this will be developed in the course of implementing this strategy.
- GESIS commits to examining the documentation of Open Science in its performance indicators.
- GESIS will acquire further expertise as a leading and competent partner in Open Science.
- GESIS will facilitate the implementation of this strategy by proving infrastructure and support to its employees.

GESIS will persistently pursue these objectives and integrate them in its established strategic process.

2.2 Premise and trade-offs

Both the alignment of the provision of services toward Open Science and the adaptation of research processes at GESIS require far-reaching changes. It is not without reason that GESIS emphasizes that Open Science must become fully ingrained in the institute’s own culture. They, therefore, result in trade-offs that need to be addressed.

To embrace the potential of Open Science, GESIS needs openness towards errors and a productive way of dealing with them. Errors in scientific work are spotted more easily when processes are more transparent. Yet, individual researchers might be reluctant to make themselves vulnerable

by working openly and transparently. GESIS acknowledges that Open Science can result in errors to be spotted more easily and, therefore, aims to embrace an open error culture.

Learning new tools and adopting new workflows will come at initial costs. Importantly, we recognize that change towards Open Science needs to happen incrementally and new tools and workflows cannot be adopted at once. On the upside however, initial setup costs will likely result in long-term benefits and time savings: better documentations will initially cost more time, however, continuing a well-documented project after some time will likely save time compared to a less well-documented project. Additionally, open science tools will take time to learn but will likely be beneficial in collaborations and save time in the long run. Beyond that, many funding agencies and journals already request open science practices (e.g., uploading data with a publication). Hence, adopting open science tools and workflows now will result in time savings later when having to comply with Open Science requirements by funders, publishers or other stakeholders. GESIS, as an organization, must provide resources, for example, for training and the preparation of best-practice guidelines. Additionally, GESIS will acknowledge contributions to Open Science by its researchers. However, these individual efforts must be visible (e.g., by means of documentation in research information systems like GRIS). Moreover, together with others, GESIS will strive to ensure that Open Science efforts are acknowledged in an appropriately reputation-enhancing way. Such efforts are, for example, under way within the National Research Data Infrastructure (NFDI).

2.3 Implementation and Incentivization: Bringing about a change in culture

Since the first version of this strategy in 2017, Open Science has made substantial progress at GESIS. The same is true for the social sciences as a whole over the last few years. Most researchers would probably agree that openness is desirable, maybe even necessary. Yet, only a small number of researchers share all of their data or document their methodology in detail for others to easily retrace their steps. It is, hence, worth addressing how Open Science can become a core part of the institute's culture.

GESIS has defined a set of measures to implement this strategy and reach the goals identified below. Towards its community GESIS will provide services and lead by example. Internally, GESIS also strives to re-shape incentives to encourage Open Science. Specifically, we aim to

- adapt our understanding of good scientific practice to the DFG codex making Open Data the new normal (DFG 2019).
- show appreciation for efforts toward Open Science (e.g., through recognition).
- support recording Open Science outputs (e.g., the number of publications published open access and the number of publications documented according to open methodology standards) in the GESIS Research Information System (GRIS).
- encourage implementation of Open Science principles (particularly sharing of data and methodology in a reproducible fashion) at the scientific journals, starting with the in-house publications.

3 Open Methodology

3.1 Status Quo

3.1.1 Open Methodology in the Social Sciences

Open Methodology refers to the transparent documentation of data collection, processing and analysis, and means of reproducing these steps. In the social sciences, the debate about Open Methodology is still at an early stage, while at the same time, some research approaches, like machine learning make reproducibility yet more cumbersome. The Open Notebook Science approach developed in the natural sciences is often referred to when discussing transparent research. This approach aims to document all relevant pathway decisions and observations (also the unsuccessful ones) during the research process, thereby maximizing methodological transparency. Such practices are, to a certain extent, already common in qualitative research (e.g., ethnography), for example, via the logbook function of the relevant software for qualitative analysis. Quantitative research is ultimately not reproducible without the complete documentation of the collection of the data and of the analysis methods used (documented by means of code, e.g., Python-, R-, SPSS-, Stata code).[3] A transparent description of the data collection procedure also allows others to independently run a replication of the study (i.e., collection of new data with the same method) and critically assess the data collection procedure.

Open Methodology can be understood as a continuum. The degree of openness will continue to differ as long as standards are still in the process of being established. Several proposals have already been made, for example by the American Psychological Association (APA), by Dafoe (2014), and by King (1995). Ultimately, it will also be profession-specific standards that define what "open" and "reproducible" may and should mean in the respective context. Practical tips can increasingly be found for how the data for reproducing statistical analyses of collected data can meet the aforementioned demands.[4]

In all cases, one desideratum of Open Methodology is that the documentation of and/or the software used for a study must allow others to reproduce and replicate the study, preferably without consultation with the author(s). Moreover, wherever possible, the documentation of the data analysis should use open software, and open data formats. The aim must be to ensure that conditions for reproduction exist for as long as possible. Accordingly, analysis code, the software used, and other relevant details must be documented. One needs to accept, however, that the transition to Open Methodology cannot happen instantly, even once there is a consensus on a (field specific) definition. Therefore, every step towards a more reproducible and transparent description of the *data collection* and *analysis* should be encouraged and embraced.

A hotly discussed topic in the debate on quality assurance in science is preregistration. That is, the submission of the study design and analysis plan to an independent registry before the data are collected or hypotheses are tested (Nosek, et al. 2017). The aim is to ensure that research questions and results are not subsequently "fitted" to the empirical data. Preregistration can, thus, be

[3] This applies equally to qualitative research data with regard to the supplementing of raw data (e.g., transcripts, corpora) with the documentation of the coding schemas within the software program used for qualitative data analysis (QDA) and, ideally, also of the underlying development processes, whereby particularly the interviews and their comparatively time-consuming anonymization must be taken into account.

[4] This is particularly challenging in the case of administrative data (Playford, et al. 2016).

understood as a part of Open Methodology and is gaining traction in the social sciences (Freese & Peterson 2017; Burlig 2018; Jacobs 2019). Preregistration with journals (often referred to as registered reports) is now possible with more than 250 journals spanning all areas of empirical sciences.[5]

3.1.2 Existing support for Open Methodology at GESIS

- GESIS Notebooks offers an integrated research environment for the analysis of data in a "software as a service" architecture. The virtual research environment enables readily re-runnable encapsulations of data analysis pipelines in the Binder format, as well as their transparent documentation. The environment is based on Project Jupyter and enables online executable analyses using R, Python, LaTeX, Markdown as well as Git. It is also open to proprietary extensions, e.g., for Stata and SPSS.
- As part of its Survey Guidelines GESIS provides open science guidelines on transparent research practices in documenting data collections and data analyses to the scientific community.
- Numerous internal and external workshops are held to assist scientists with open methodology tools (such as LaTeX, Markdown, R, Python, and Git).
- SowiDataNet|datorium is a research repository for the social sciences and economics that enables researchers to easily document, publish and share (quantitative) data and analysis code. Persistent identifiers (DOI) allow the data and code to be easily cited and accessed.
- Replikationsserver.de, which builds on the SowiDataNet|datorium data sharing repository, is a joint initiative of GESIS and the journals *Zeitschrift für Soziologie* and *Soziale Welt* for the provision of social science data for reproduction purposes. Authors who publish in these journals make the quantitative data on which their publications are based and the accompanying analysis code publicly accessible.

3.2 Objectives

In order to further foster Open Methodology at GESIS we set objectives in the areas of reproducibility and replicability, preregistration, and training.

Ensuring the *reproducibility and replicability* of research done at GESIS will provide infrastructure to assist with a transparent research process:

- GESIS will extend the Notebook infrastructure to enable a wide use of reproducible analyses and remote secure data access.

We also aim to establish the institute as a positive example within the social sciences:

- GESIS employees are asked to provide methodological details and analysis scripts with their publications to enable others to reproduce data generation and analyses. To encourage this practice GESIS will assist authors to find the resources necessary (mainly time).

[5] See https://www.cos.io/initiatives/registered-reports (accessed on 21 October 2020)

- Replication studies are encouraged to be conducted as part of research projects at GESIS.

Assisting with *preregistration*

- GESIS will examine possibilities to encourage preregistration (e.g., provide the questionnaire of a survey to the community before making the data publicly available, so data users can preregister their hypotheses before the data can be inspected). The German Longitudinal Election Study will optimize the data publication process to enable and encourage users of GLES data to pre-register their research hypotheses (Bucher, et al. 2021).
- Additionally, GESIS employees are encouraged to preregister their research with independent registries or directly with journals[6]. GESIS will examine how these efforts can be highlighted within the institute.

Firmly establishing Open Methodology in *training* courses

- Open Methodology will be a standard part of GESIS training courses.
- GESIS will encourage external trainers to use Open Methodology and Open Data in their courses.

[6] See https://cos.io/rr/ for more information on preregistration with journals (and a list of participating journals, accessed on 21 October 2020).

4 Open Data

4.1 Status Quo

4.1.1 Open Data in the Social Sciences

The social sciences generate diverse types of data; some at substantial financial costs. Therefore, particularly in the area of cross-national comparative surveys, there has been a long tradition of data sharing. Infrastructures organizing the access, exchange, and preservation of data emerged as early as the late 1950s. Yet, a truly broad consensus that data should be shared is still only emerging (DFG 2019).

4.1.2 Existing support for Open Data at GESIS

If the potential of scientific data is to be fully exploited, they must be *findable*, *accessible*, and, to the greatest extent, possible, *interoperable* and *re-usable (FAIR)*.[7] GESIS already offers infrastructure supporting these goals.

Findability: GESIS provides extensive metadata for every dataset it curates. Datasets, accompanying documents, and metadata receive a Digital Object Identifier (DOI) which refers to a landing page at GESIS. The DOI and the landing page with the data-related information are searchable, findable, and citable. GESIS data can be found in domain-specific search portals such as the DataCite Search or the CESSDA Data Catalogue.[8] Further measures are planned to improve findability (e.g., search engine optimization). However, there is still a great lack of mechanisms to ensure that GESIS data are properly findable via Web search engines.

Accessibility: GESIS strives to make data and metadata as open as possible. All metadata describing data hosted by GESIS are freely accessible and available for reuse under a CC0 license. At the same time GESIS balances data protection requirements, the desire for the greatest possible openness of research data, and the legitimate interests of the data producers. Therefore, GESIS offers different access regimes for research data. These range from entirely unrestricted access via the web to highly controlled on-site access via the institute's Secure Data Center. With its preservation services, GESIS guarantees long-term interpretability and accessibility of data in a constantly changing technological environment.

Interoperability: GESIS supports researchers through a controlled ingest process for all incoming data in order to guarantee data publication and subsequent re-use without technical or legal impediments. Standardized metadata and the use of controlled vocabularies (e.g., on the sampling and data collection process) help to make data interoperable. Furthermore, GESIS aims at standardizing and harmonizing dataset structures and file formats. [9]

Re-usability: The re-usability of data depends on technical aspects like the file's format, legal aspects like copyright and data protection, and the quality of the data documentation. For all data disseminated by GESIS at least a minimal set of metadata sufficient for use of the data is available. For important survey programs granular documentations are created (mostly variable level infor-

---

[7] https://www.go-fair.org/fair-principles/ (accessed on 21 October 2020)

[8] https://search.datacite.org/ ; https://datacatalogue.cessda.eu/

[9] The Comparative Study of Electoral Systems (CSES) has, for example, extended interoperability by facilitating linkage to other datasets through identifiers.

mation for surveys). The tools to create documentation for surveys use the Data Documentation Initiative standard[10]. Nevertheless, the documentation is mostly provided using text documents (i.e., questionnaire documentations and codebook in pdf-files) that are not machine readable.

4.2 Objectives

GESIS will further simplify access to research data. To this end, barriers to archiving data and to re-using research data shall be removed to the extent legally possible. GESIS shall organize its data services for the social sciences in accordance with the following objectives:

- *Make documentation accessible.* GESIS will support a transparent research process by making accessible all documentation and research output related to data provided through GESIS. To help achieve this, GESIS will promote sharing of code (via SowiDataNet|datorium, GESIS Notebooks or through other services), publications (via SSOAR or through other services), and of instruments (via ZIS or through other services). This will become a part of our data usage agreement.
- *Make data even more findable*: To ensure that GESIS data are properly findable via common web search engines GESIS will continuously apply search engine optimization techniques.
- *Further increase accessibility*: GESIS shall increase its efforts to encourage researchers to share their research data. In doing so, it will also foster its long-standing collaborations with similar institutions and initiatives, for example in the German Data Forum, CESSDA, or the Research Data Alliance. GESIS shall create opportunities for data depositors to make their data available as openly as possible, removing, wherever possible, existing barriers. Moreover, GESIS shall continue its commitment at various levels (e.g., in the context of the National Research Data Infrastructure - NFDI) to harmonize conditions of access to data. Importantly, GESIS shall enable remote access to its sensitive data holdings (in the Secure Data Center). Research data shall be available in formats commonly used by the social science community, preferably open data formats which can be expected to remain readable in the long term. For research data submitted in open and community-supported formats, GESIS will guarantee long-term accessibility.
- *Strengthen interoperability*: To increase the ability of GESIS (meta)data to be interpreted not only by its own community or by community-specific agents, but also by generic agents, GESIS needs (1) to make its (meta)data available in widely used format(s), such as RDF, and (2) to either annotate (meta)data by using common Web vocabularies or to ensure to make the vocabulary systems used themselves are FAIR, such as by defining them in the Web Ontology Language (OWL) and by sharing them via a publicly accessible vocabulary registry.
- *Further increase re-usability*: In order to make data more re-usable GESIS strives to employ Creative Commons Licenses wherever possible. It also requires (1) to increase the information on data preparation provided with the metadata and (2) develop metadata interfaces for machines.

[10] https://ddialliance.org/explore-documentation

GESIS will also strengthen Open Data in its own research practice. Therefore, in addition to proposals for an open documentation of methods (see 3.2) and in line with the DFG's Code for Good Scientific Practice, GESIS will make open data the default for research projects carried out by GESIS staff.

5 Open Access

5.1 Status Quo

5.1.1 Open Access in the Social Sciences

Open Access is a catalyst to Open Science. It widens access to research findings and democratizes knowledge; it increases citations of scientific work by factor 1.5 (Archambault, et. al. 2016) and enhances timely engagement with new findings. Also, it attracts researchers to Open Data, Open Methodology, and Open Source used or cited in publications. As Open Access publications are cited more often, they increase the visibility of research, and, by extension, of the data or methods used therein.

Nevertheless, Open Access publishing and archiving of full texts is not yet as widespread in the social sciences as it is in other disciplines. On the one hand, there simply is not yet a sufficient number of Open Access journals in the discipline. On the other hand, social scientists still publish a lot in collective works or monographies for which, so far, no systematic genuine open access model has been developed. At the same time, Plan S[11] – a coalition of research funders from 15 countries – pushes for changes. These also reveal themselves in a growing number of commitments to Open Access expressed through institutional policies, changes to legal parameters (e.g., the right of secondary publication, Section 38 (4) of the German Copyright Act, UrhG) as well as new forms of licensing according to the so-called Read & Publish model (e.g., DEAL negotiations). Despite a growing number of indexed, genuinely open access journals, the so-called "green road to Open Access", that is, the self-archiving of a version of the work in a digital repository is still highly relevant. It yields higher citations than golden OA (Young and Brandes 2020). Given the relative infancy of Open Access in the Social Sciences the green way is likely to be central for at least another 5 years to fulfil funders' mandates and to open access to publications in collective works or monographies.

5.1.2 Open Access at GESIS

GESIS actively supports the Open Access Transformation in the Social Sciences – for its own employees as well as for the discipline as a whole. Through its infrastructures it strengthens gold *and* green open access: In cooperation with Springer/BMC GESIS edits an interdisciplinary Open Access Journal Measurement Instruments in the Social Sciences (MISS). With the Social Science Open Access Repository (SSOAR), GESIS operates the largest disciplinary repository for the social sciences in Europe, if not worldwide. Full texts are archived with appropriate metadata and a persistent identifier.

All GESIS publication series are freely accessible immediately or after an embargo period (delayed Open Access). GESIS itself publishes the Open Access formats *mda*, HSR, and *Survey Methods: Insights from the Field* in collaboration with FORS. Furthermore, GESIS has adopted an Open Access Policy[12] to encourage and support their employees with regards to Open Access publishing.

Through SSOAR and its library GESIS is part of a greater network that strives to enhance Open Access in the discipline.

---

[11] https://www.coalition-s.org/ (accessed on 21 October 2020)

[12] http://www.gesis.org/angebot/publikationen/open-access-policy/ (accessed on 21 October 2020)

5.2 Objectives

GESIS supports Open Access to enhance permanent access to social science research and through it to other aspects of Open Science. One of GESIS' key objectives remains to anchor SSOAR firmly in the publication culture of the social sciences as an infrastructure service for archiving.

To further promote Open Access in the social sciences, GESIS has the following objectives:

- *Establishing reliable archiving infrastructure:* SSOAR will continue to develop sustainable cooperation with publishers and research institutes and offer self-archiving services to individual researchers. Along with libraries, medium-sized publishers, intermediaries (such as pledging partners) GESIS supports the Open Access Transformation through its archiving infrastructure. GESIS will continuously improve SSOAR as a service (e.g., through author disambiguation/ORCID).
- *Digitization of so called "Klassiker der Soziologie":* In cooperation with relevant publishing houses and libraries GESIS will strive to digitize and archive highly relevant publications for social science research from before the year 2000 which so far are not available in a digital format.
- *Linked Open Data:* GESIS will focus on providing links to its research data and measurement instruments archived at GESIS when they are cited in publications. The basis for that will be the research-data bibliographies published annually.[13]
- *Data import and indexing:* The time and effort required to record content shall be reduced through the reuse of already existing metadata by means of (semi-)automatic indexing and tagging.

For the further promotion of Open Access at GESIS, the following objectives apply:

- GESIS will continue to offer to its researchers the service to make publications available in SSOAR as soon as possible considering legal restrictions.
- SSOAR will systematically archive publications available that cite research data or survey instruments that have been produced or archived at GESIS.
- GESIS will test and actively promote metrics that help ensure the scientific quality which, to some extent, might be through alternative metrics (Bornman and Haunschild 2017; Mehrazar, et. al. 2018).
- In the context of the DEAL process[14], GESIS will follow a strategic licensing policy to increase access to the portfolio of relevant publishers as well as offer the best possible options to its employees to publish in Open Access Journals at reasonable costs.

Through an appropriate rights-management GESIS will assure that the full text-collection of SSOAR can be used for data-mining and other information research.

[13] https://www.gesis.org/en/allbus/contents-search/allbus-bibliography (accessed on 21 October 2020)

[14] https://www.projekt-deal.de/aktuelles/ (accessed on 21 October 2020)

6 Open Source

6.1 Status Quo

6.1.1 Open Source in the Social Sciences

Open Source refers to software whose source code is openly accessible to third parties and, therefore, reusable and often modifiable. Importantly, authors retain the intellectual ownership when publishing their software open source. The intentions behind Open Source are the complete reproducibility of the program code and the algorithms, as well as free access to software by researchers.

From a licensing perspective, open source software is generally subject to less legal obligations than software made available under commercial licenses. Open Source software is usually made available free of charge.

6.1.2 Open Source at GESIS

Open Source is not as well established at GESIS as other elements of Open Science. Yet, because of the above-mentioned advantages, GESIS strives to provide its software tools for the social sciences as Open Source. Where possible, GESIS also uses open source software in its projects or services. Furthermore, a GitLab server is available for open source projects at GESIS, and GESIS Notebooks provides browser-based reproducibility for open source data analysis projects. Additionally, trainings on open source software are provided to GESIS employees which enable them to use these tools in their daily work.

6.2 Objectives

- Using standardized publication guidelines, GESIS shall publish self-developed software as open source. To make such contributions more visible GESIS aims to accompany software publications with corresponding journal publications, for example in the *Journal of Statistical Software*.
- Computer code is often produced in the framework of complex statistical analyses (e.g., do-files for Stata or R scripts). This type of code shall be published under an open source license.
- To increase the visibility of its software projects, GESIS will make them available in a publicly accessible Git repository.
- GESIS wishes to foster the abilities of its employees to use open source software. It will, therefore, support training courses related to open source software.

7 References

Akademie für Soziologie (AS). 2019. Wissenschaftliche Daten sind kein Privateigentum einzelner Forschender, sondern ein kollektives Gut. Die Bereitstellung von Forschungsdaten zur Nachnutzung und Replikation muss auch in der Soziologie die Norm sein. Positionspapier der Akademie für Soziologie. Januar 2019. Mannheim.
https://akademie-soziologie.de/wp-content/uploads/2019/11/AS-Stellungnahme-Nachnutzung-Forschungsdaten-01-2019.pdf

Allianz der deutschen Wissenschaftsorganisationen. 2010. Grundsätze zum Umgang mit Forschungsdaten. Juni 2010.
https://www.leopoldina.org/fileadmin/redaktion/Publikationen/Allianz/100624-allianz-forschungsdaten.pdf

Archambault, É., Côté, G., Struck, B., and Voorons, M. 2016. Research impact of paywalled versus open access papers. Copyright, Fair Use, Scholarly Communication, etc.
http://digitalcommons.unl.edu/scholcom/29

Bornman, L., and Haunschild, R. 2017. Does evaluative scientometrics lose its main focus on scientific quality by the new orientation towards societal impact? Scientometrics. 2017; 110(2): 937–943. https://doi.org/10.1007/s11192-016-2200-2

Bucher, H., Stroppe, A.-K., Burger, A., Heycke, T. 2021. GLES Open Science Strategie. GESIS Papers, *forthcoming*. https://doi.org/10.21241/ssoar.68698

Burlig, F. 2018. Improving transparency in observational social science research: A pre-analysis plan approach. Economics Letters, 168, 56–60.
https://doi.org/10.1016/j.econlet.2018.03.036

Dafoe, A. 2014. Science Deserves Better: The Imperative to Share Complete Replication Files. PS: *Political Science and Politics* January 2014, 60–66.
https://doi.org/10.1017/S104909651300173X

Deutsche Forschungsgemeinschaft (DFG). 2017. Replizierbarkeit von Forschungsergebnissen. Eine Stellungnahme der Deutschen Forschungsgemeinschaft. April 2017. Bonn
https://www.dfg.de/download/pdf/dfg_im_profil/reden_stellungnahmen/2017/170425_stellungnahme_replizierbarkeit_forschungsergebnisse_de.pdf

Deutsche Forschungsgemeinschaft (DFG). 2019. Guidelines for Safeguarding Good Research Practice. Code of Conduct. September 2019 Bonn.
https://www.dfg.de/download/pdf/foerderung/rechtliche_rahmenbedingungen/gute_wissenschaftliche_praxis/kodex_gwp_en.pdf

Deutsche Gesellschaft für Soziologie (DGS). 2019. Bereitstellung und Nachnutzung von Forschungsdaten in der Soziologie. Stellungnahme des Vorstands und Konzils der DGS. Januar 2019. Essen

European Commission (2016). Digital Agenda for Europe.
http://ec.europa.eu/digital-agenda/en/open-science

Freese, J., and Peterson, D. 2017. Replication in Social Science. Annual Review of Sociology 43, 147-165. https://doi.org/10.1146/annurev-soc-060116-053450

Grötschel, M. 2016. Elektronisches Publizieren: Open Access, Open Science und ähnliche Träume. https://soziologie.de/aktuell/news/bereitstellung-und-nachnutzung-von-forschungsdaten-in-der-soziologie

OECD. 2015. Making Open Science a Reality. OECD Science, Technology and Industry Policy Papers, No. 25, OECD Publishing. https://doi.org/10.1787/23074957

Jacobs, A. 2020. Pre-registration and Results-Free Review in Observational and Qualitative Research. In *The Production of Knowledge Enhancing Progress in Social Science*. Cambridge University Press.

King, G. 1995. "Replication, Replication." PS: Political Science and Politics, 28, Pp. 444-452. Copy at https://j.mp/2oSOXJL

Mehrazar, M., Kling, C. C., Lemke, S., Mazarakis, A., and Peters, I. 2018. Can We Count on Social Media Metrics?: First Insights into the Active Scholarly Use of Social Media. WebSci '18: Proceedings of the 10th ACM Conference on Web Science. May 2018: 215-219. https://doi.org/10.1145/3201064.3201101

Nosek, B. A., Ebersole, C. R., DeHaven, A. C, and Mellor, D. T. 2017. *The Preregistration Revolution*. OSF Preprints. July 2017. https://dx.doi.org/10.17605/OSF.IO/2DXU5

Playford, C., Gayle, V., Connelly, R., & Gray, A. 2016. Administrative social science data: The challenge of reproducible research. Big Data and Society. http://dx.doi.org/10.1177/2053951716684143.

Rat für Sozial- und Wirtschaftsdaten (RatSWD). 2010. Grundsätze zum Umgang mit Forschungsdaten. Allianz der deutschen Wissenschaftsorganisationen . RatSWD Working Paper Series – Working Paper Nr. 156. Juni 2010. Berlin. https://www.ratswd.de/download/RatSWD_WP_2010/RatSWD_WP_156.pdf

Young, J. S., and Brandes, P. M. 2020. Green and gold open access citation and interdisciplinary advantage: A bibliometric study of two science journals. *The Journal of Academic Librarianship.* https://doi.org/10.1016/j.acalib.2019.102105