

Scientific data is not the private property of individual researchers, but a collective good. The availability of research data and their subsequent use and replication must also become a norm in sociology

Position paper of the Academy of Sociology

January 2019

Productive science utterly depends on the participation of scientists in the production of collective goods. The resulting knowledge is shared – as are the methods and data with which knowledge is created. Public goods are successfully created when intelligent institutional rules support their creation. In this context, the use of research agendas is currently being debated, and the scientific community is deciding which infrastructures and institutions should be developed and established. Making the data and other material as freely available as possible for the research community makes the research process intersubjectively comprehensible. Transparency stimulates further scientific thinking.

For some years now, the "Alliance of Science Organisations" in Germany, including the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation), has endeavoured to strengthen the standard of sustainable safeguarding and provision of research data, and to develop suitable procedures for enforcing this standard.¹ In order to facilitate a practical implementation, the DFG has called on scientific associations to participate in the discussion and to formulate subject-specific standards.

The Academy of Sociology expressly supports these initiatives to provide research data. Public access to scientific products is an integral part of open societies and ethically responsible research. Open, reusable data enable the correction of errors, they form a significant potential for inter- and multidisciplinary, as well as international, research cooperation, and thus particularly good opportunities for cumulative knowledge gain. In many cases, the potential of rich, usually

¹ Principles for handling research data, adopted in 2010: https://www.Wissenschaftsrat.de/download/archiv/Allianz_Grundsätze_Forschungsdaten.pdf. It states, for example: "The sustainable safeguarding and provision of research data serves not only to examine past results but also, to a large extent, to achieve future results. It is a strategic task to which science, politics and other parts of society must jointly contribute".

publicly funded data is far from exhausted, and individual researchers are not able to exploit the data fully. It follows from this that the provision and re-use of research data must also be the rule in sociology.

The Academy of Sociology considers it incomprehensible that this basic position is not generally shared within the subject of sociology. The German Society for Sociology (DGS) recently announced in a statement "that (...) the secondary use of data must not be accepted as a 'normal model' of research (...)". Above all, serious problems with non-standardised data are used to justify not passing on research data under normal circumstances.²

The standard procedure however should be defined exactly the other way around, because the provision of non-standardised data does not involve insurmountable hurdles. Solutions are possible and are often already implemented in other subjects³:

- 1) Data safety: where anonymization of research data would be too costly or even impossible, the data can be made available to a "secure centre" for subsequent use. There, secondary users can evaluate the data, taking only the evaluation results, but no sensitive information from the secure centre.
- 2) Rights of primary research: it is common practice for third-party funding bodies to grant primary researchers an "embargo period" (up to four years). Rules for the correct citation of data sets also already exist.
- 3) Lack of resources: third-party funding sources such as the DFG already make it possible to apply for resources for data archiving. Funding can be applied for within the framework of individual projects (e.g. personnel funding), but also within the framework of researcher's own infrastructure platforms. The time required for this does not create a career disadvantage for primary researchers if the provision of data is a general rule and not just a voluntary service.
- 4) Unresolved technical and organisational challenges and support needs: at the University of Bremen, for example, the data service centre "Qualiservice" provides support for non-standardised data (www.qualiservice.org). Qualiservice offers various services for the provision and subsequent use of non-standardized data, including their archiving, anonymization and secure use in the "Secure Centre". Thus, Germany already has a comprehensive infrastructure for the provision and re-use of non-standardized data.
- 5) Contextual information is only accessible to primary researchers and can be interpreted by them: particularly in order to increase transparency, relevant contextual information represents an integral part of research data and scientific publications. If primary researchers do not provide all relevant context information, the intersubjective traceability is undermined and research results devalued.

² "Provision and re-use of research data in sociology" of 8 January 2019; <https://www.soziologie.de/nc/aktuell/stellungnahmen/single-view/archive/2019/01/09/article/bereitstellung-und-nachnutzung-von-forschungsdaten-in-der-soziologie/>. For example: "Primary researchers have in any case the right to restrict the conditions for subsequent use [...]. Demands, for example, from research-funding organisations or publication bodies, must not relativise this right".

³ Disciplines with heterogeneous research approaches can serve as models, for example the differentiated statement of the Review Board Linguistics: http://www.dfg.de/download/pdf/foerderung/antragstellung/forschungsdaten/guidelines_review_board_linguistics_corpora.pdf

The Academy of Sociology recognises the principles of FAIR Data - (Findable, Accessible, Interoperable, and Reusable Data), as established by the European Commission, as useful standards for sociology. Deviations from these principles should be justified in all decision-making processes of the science system and community, from research funding to publication.