



German
Historical
Institute
London

German Historical Institute London Bulletin

Conference Report:

*Infrastructures for Social Data in Contemporary History:
Proposals for a Better Future*

by Kerstin Brückweh and Pascal Siegers

German Historical Institute London Bulletin
Vol. XLV, No. 2 (November 2023), 142–7

ISSN 0269-8552

Infrastructures for Social Data in Contemporary History: Proposals for a Better Future. Conference organized by the working group Social Data in Contemporary History in cooperation with the German Historical Institute London and held at the Werner Reimers Foundation, Bad Homburg, Germany, 21 November 2022. Conveners: Christina von Hodenberg (GHIL), Lutz Raphael (Trier University), and Albrecht von Kalnein (Werner Reimers Foundation).

Data from social science research and official statistics from past decades are becoming increasingly important as sources for research in history. The project 'Social Data as Sources of Contemporary History', funded by the German Research Foundation (DFG), has investigated what support contemporary historians need in order to productively use social data, whether qualitative (mainly transcripts and audio and video recordings) or quantitative (mainly surveys and statistics of various origins). The project began by publishing a needs assessment outlining how research in this area might be enabled.¹ In parallel, workshops were held to explore how working with social data could lead to new insights for contemporary history.² Finally, the project has published a position paper outlining measures for improving research infrastructures.³

The aim of the meeting in Bad Homburg was to discuss this position paper with representatives of the research community and NFDI4Memory, the nascent consortium for the historical sciences in the German National Research Data Infrastructure (NFDI).⁴ Since the

¹ Kathrin Zöller, Clemens Villinger, Pascal Siegers, Sabine Reh, Lutz Raphael, Christina von Hodenberg, and Kerstin Brückweh, 'Sozialwissenschaftliche Forschungsdaten als historische Quellen: Welche Infrastrukturbedarfe hat die zeitgeschichtliche Forschung?', *RatSWD Working Paper Series*, 277 (2022), at [<https://doi.org/10.17620/02671.66>].

² See e.g. the special issue on historical social data published in *Geschichte und Gesellschaft*, 48/1 (2022).

³ Kerstin Brückweh, Christina von Hodenberg, Lutz Raphael, Sabine Reh, Pascal Siegers, Clemens Villinger, and Kathrin Zöller, 'Positionspapier zu Infrastrukturen für historische Sozialdaten in der Zeitgeschichte', *Zenodo* (2023), at [<https://doi.org/10.5281/zenodo.7781159>].

⁴ An overview of NFDI4Memory's activities is available at [<https://4memory.de/>], accessed 26 Apr. 2023.

application for the DFG project was submitted on 1 August 2018, the environment for research data infrastructures has changed fundamentally. The establishment of the NFDI in October 2020 created a new framework for the development and integration of research data infrastructures in Germany. Further proposals for a data infrastructure for research in contemporary history must therefore at least refer to the new NFDI structures in the humanities and social sciences, or ideally be fully integrated with them.

After a warm welcome by Lutz Raphael (for the working group on Social Data and Contemporary History) and Albrecht von Kalnein (for the Werner Reimers Foundation), the conference began with a brief presentation of the project group's position paper by Christina von Hodenberg and Pascal Siegers (GESIS—Leibniz Institute for the Social Sciences). From its analysis, the project group has identified three pillars on which a supporting structure should rest. In the paper, this proposed structure is flanked by a discussion of existing policies for handling social data—specifically, the future preservation of social data by requiring researchers to offer their data to public archives.

The project's needs analysis clearly showed that a lack of skills in dealing with social data greatly increases the effort required to use it.⁵ Yet history curricula rarely teach practical skills in data management and analysis. New forms of knowledge and skills transfer are therefore needed that focus on the forms of social data analysis specific to historical research. It follows that the first pillar should be a teaching and training centre that would develop and offer courses on the secondary analysis of qualitative or quantitative social data in the historical sciences.

The second pillar would be an information portal on historical social data, with the aim of supporting contemporary historians in their search for suitable data by bringing together and preparing information from the fragmented social science research data infrastructure. This would include information on the conditions governing data access, such as the need to comply with data protection

⁵ Zöller et al., 'Sozialwissenschaftliche Forschungsdaten als historische Quellen'.

regulations. The information provided would be complemented by a forum for sharing innovative evaluation methods and discussing the challenges of using historical social data.

Finally, a third pillar—a forum on ethics and law in contemporary history—would address precisely these aspects of the present-day use of data collected long before current legal regulations were in place. The needs analysis showed that researchers currently have no point of contact within the historical sciences to turn to with their legal and ethical questions. Unlike other disciplines, history lacks research ethics guidelines for handling data and sources. This third pillar therefore aims to support the formation of standards, for example, by developing a code of conduct for working with historical social data.

One problem with using historical research data is that much of it is not preserved for the long term. Despite a growing awareness of its value, there is still no obligation in Germany to store and publish data in a repository or archive. The DFG's rules of good scientific practice only require data to be kept for ten years. At present, however, no account is taken of the fact that data from research projects at universities, universities of applied sciences, and non-university research institutes in Germany is subject to federal and state archive legislation. Under these rules, the archival value of such data is derived, among other things, from its significance for 'research into and understanding of history and the present'. Social science data is an important means of self-observation in society and therefore its archival value must be examined on a case-by-case basis. This should result in an obligation to offer social science and other research data to archives or research data repositories, which in turn can ensure its long-term preservation and usability.

Finally, von Hodenberg and Siegers emphasized the need to link new infrastructures for contemporary history with existing structures in the social sciences and to integrate them into the emerging NFDI4Memory. Infrastructures for using social data as sources can be conceived of and eventually implemented as a bridge between the social and historical sciences.

The presentation of the position paper was followed by a short introduction of the NFDI4Memory consortium by its leaders,

Johannes Paulmann and John Wood (Leibniz Institute of European History). They explained that NFDI4Memory works to promote the development of a data culture in the historical sciences, to strengthen data literacy (in the sense of basic skills in dealing with data), and to develop standards for data management and documentation. Legal and ethical issues are also part of the planned work programme. The NFDI4Memory consortium sees itself as a ‘network infrastructure’ linking actors from all periods of historical research. However, the available funds are insufficient to provide large-scale technical services for indexing, processing, and storing data. A significant part of the budget is flexible, allowing incubator projects to be integrated into the NFDI4Memory work programme from 2024 onwards through a competitive process. A major difference to consortia in the social and linguistic sciences becomes apparent here: because no (research) data centres have been established in the historical sciences that are comparable with the German Data Forum (RatSWD) or CLARIN-D, the structures for processing and curating data in historical institutes have yet to be established or permanently funded. Although the German Council for Scientific Information Infrastructures (RfII) recommended the establishment of the NFDI primarily with the aim of developing more stable funding instruments for data infrastructures, funding for all NFDI consortia will end in 2028. Complementary initiatives are therefore needed.

An important aspect of the discussion was the obvious question of what points of connection there are between the proposals of the DFG project and the 4Memory consortium. Overlaps were particularly evident in the areas of law and ethics, as well as skill building. Teaching formats developed for the evaluation of social data as historical sources fit well into the overall concept of the NFDI consortium. At the same time, there was a call for increased dialogue with representatives and committees of state archives and memory institutions, which face similar challenges to historical research in the course of digitization. Archives in particular have an important role to play in ensuring that data produced today can be made available to historians in the future. This task goes far beyond obliging researchers to make social science data available, as archives need to adapt their work to new technologies and digital object types. In this context,

critical voices were raised about the working group's position paper because it does not sufficiently address how to enable the future use of data in the context of digital humanities. Digital sources open up new possibilities for linking, extraction, and aggregation, which are not among the clearly stated goals of the social data infrastructure.

The central issue in all discussions about improved research data infrastructures is the availability of stable funding, as initiatives need a certain longevity to be effective. As the NFDI is not permanent, at least in its first phase, its prospects remain limited. The only long-term funding instrument available for infrastructure in Germany is the Strategic Special Fund (*Strategischer Sondertatbestand*), which the institutes of the Leibniz Association can apply for in a competitive process. The prerequisite for success is a close strategic fit with the goals of the institutes and their willingness to make a substantial financial contribution of their own. This path requires medium-term preparation in order to continue and extend initiatives from university research. Therefore, implementing the pillars of a research infrastructure for historical social data proposed in the position paper will not be possible without start-up funding in the form of projects. Under the current conditions, this path of piecemeal projects seems to be the only funding option.

Participants agreed that each of the pillars can be implemented independently. The proposed actions relating to capacity building and knowledge transfer are particularly important to enable researchers to properly use social data as a source in the historical sciences. The current situation is that the needs for social data in contemporary history research have been formulated, but the major projects (first and foremost NFDI4Memory) still need to be developed. In the meantime, those responsible for the DFG project do not intend to remain inactive; instead, as in previous years, they are relying on community building by organizing conferences within the framework of the 'Social Data in Contemporary History' working group based at the Reimers Foundation. This means that in the coming years the importance of analysing social data as historical sources for research in contemporary history will become clear. The first of these conferences, on social inequality, will be held in the autumn of 2023, when representatives of various institutions and disciplines will meet in the established format to

discuss the potential of social data and its analysis and to place it on the most secure possible footing.

KERSTIN BRÜCKWEH (Berliner Hochschule für Technik) and PASCAL SIEGERS (GESIS—Leibniz Institute for the Social Sciences)