

A Modern Statistical Infrastructure for Excellent Research and Policy Advice –

Report on the German Council for Social and Economic Data during its First Period in Office (2004 - 2006)

by Heike Solga and Gert G. Wagner

In this article, the two chairpersons of the German Council for Social and Economic Data (RatSWD: Rat fuer Sozial- und Wirtschaftsdaten) give a summary of the council's work in the years 2004 to 2006 and its future plans for improving the statistical infrastructure for economic and social research. The full report on the work of the RatSWD is available in German at <http://www.ratswd.de>.¹

A Brief History

Given that the starting point for founding the Council was about a decade ago a brief review of its history will certainly be welcome. The idea for the RatSWD emerged in 1999 out of the Commission to Improve the Statistical Infrastructure in Cooperation with the Scientific Community and Official Statistics (*KVI: Kommission zur Verbesserung der informationellen Infrastruktur zwischen Wissenschaft und Statistik*), which had been established by the Federal Ministry of Education and Research (*BMBF: Bundesministerium für Bildung und Forschung*). The findings outlined in its 2001 final report "Towards an Improved Statistical Infrastructure"² formed the basis on which the Founding Committee of the RatSWD (*Gründungsausschuss RatSWD*) was convened. In its three years of existence—from 2001 to 2004—the founding committee recommended four Research Data Centers and two Data Service Centers as well as the RatSWD itself. The RatSWD with its 12 members was officially constituted on 1 November 2004 by the BMBF.

The first election of the RatSWD's six individual members, who must be eminent social scientists, was held at the Second Conference on Social and Economic Data (*KSWD*) in 2004. These members include Steffen Kühnel (Universität Göttingen, Sociology), Heiner Meulemann (Universität zu Köln, Sociology), Regina T. Riphahn (Universität Erlangen, Economics), Heike Solga, (Universität Göttingen, Sociology), Joachim Wagner (Universität Lüneburg, Economics), and Bettina Westle (Universität Marburg, Political science).

¹ Cf. Eine moderne Dateninfrastruktur für eine exzellente Forschung und Politikberatung - Bericht über die Arbeit des Rats für Sozial- und Wirtschaftsdaten (RatSWD) in seiner ersten Berufungsperiode, eds. Gabriele Rolf-Engel, Heike Trappe and Denis Huschka, Berlin 2007, mimeo.

² Towards an Improved Statistical Infrastructure - Summary Report of the Commission set up by the Federal Ministry of Education and Research (Germany) to Improve the Statistical Infrastructure in Cooperation with the Scientific Community and Official Statistics, in: Schmollers Jahrbuch 121 (3), 443-468. The summary is also available in German under <http://www.ratswd.de>

The six representatives of the German research infrastructures in the social sciences (Statistical Agencies, Research and Service Data Centers as well as GESIS and SOEP) include Jutta Allmendinger (IAB: Institute for Employment Research), Johann Hahlen (Statistisches Bundesamt), Eckart Hohmann (Hessisches Statistisches Landesamt), Peter Ph. Mohler (GESIS: German Social Science Infrastructure Services), Uwe G. Rehfeld (FDZ Deutsche Rentenversicherung Bund), and Gert G. Wagner (SOEP: German Socio-Economic Panel Study).

A New Infrastructure

Clear progress has been made since 2004 in labor market research and empirical social sciences in general through the establishment of the Research Data Centers which get funding by the *BMMF*. Researchers, especially in economics and sociology, today can hardly imagine a research scene without the Centers, although even five years ago their existence was virtually inconceivable.

Now we have Research Data Centers at:

- the Institute for Employment Research (*IAB: Institut für Arbeitsmarkt- und Berufsforschung*) within the Federal Employment Agency (*BA: Bundesagentur für Arbeit*)
- the German Pension Fund (*Deutsche Rentenversicherung Bund*)
- the Federal Statistical Office (*Statistisches Bundesamt*) and the Statistical Offices of the German Länder (*Statistische Landesämter*)

and Data Service Centers at

- the Center for Survey Research and Methodology (*ZUMA: Zentrum für Umfragen, Methoden und Analysen*), which is part of the German Social Science Infrastructure Services (GESIS: *Gesellschaft Sozialwissenschaftlicher Infrastruktureinrichtungen*) and
- the Institute for the Study of Labor (*IZA: Institut zur Zukunft der Arbeit*).

Maintaining this modern data infrastructure will be one of the most important tasks for the Council in the next few years.

At present, the data centers are funded as pilot projects of the BMBF and have received start-up finance. To ensure their sustainability within the research infrastructure, however, it will be crucial to find ways to anchor them in existing institutions and to secure financing.

Internal and External Lobbying

In this summary, we would like to discuss in greater detail a few selected activities carried out by the RatSWD to improve data access in the last two years. The activities presented here are less visible than the new centers and have been carried out in very diverse fields.

In the field of *educational data*, the RatSWD has been in part successful, and sadly, in part unsuccessful. Its recommendations on the University Information System (HIS) resulted in a pilot project in which a scientific use file was created by the 1997 HIS Graduate Panel. It can be accessed through the Central Archive for Empirical Social Research in Cologne (*ZA: Zentralarchiv für Empirische Sozialforschung*, which is part of GESIS as well). Scientific use files for other HIS data on the university sector will follow.

The RatSWD has also been successful in the area of vocational training data. Federal and state research data centers now provide scientific use files of data on continuing vocational training from the European surveys. Furthermore, the German Science Council's evaluation (*Wissenschaftsrat*), which included the two chairmen of the Council, of the Federal Institute for Vocational Training (*BIBB: Bundesinstitut für Berufsbildung*) has resulted in plans for the establishment of a research data center there. This will significantly improve access for external researchers to the BIBB's data on vocational training.

In the field of *health data*, a promising relationship has been established to the Robert Koch Institute (RKI). The ongoing dialogue of RatSWD with RKI has revealed that the problem here is not so much one of limited access to data as one of limited use of the available health data by researchers. In fact, the data situation in the health sector is much better than originally assumed, but what is still lacking is an exchange of information between health researchers on the one hand and empirical social and economic scientists on the other. Future work of the RatSWD will focus on disseminating information and lobbying in this area.

To explore possibilities for improving access to *corporate and financial market data*, the RatSWD held talks with representatives of the Deutsche Bundesbank and Kreditanstalt für Wiederaufbau (KfW). Although not impossible in the future, at present there are no concrete plans for offering the kind of user-friendly data access provided by research data and data service centers. However, external researchers can work with the data as guests of these institutions on project-related visits. The RatSWD therefore recommends that these opportunities be publicized, and in the case of KfW, that specific data be handed over to the Central Archive in Cologne.

In the fields of *media data* and the *statistics on crime and justice*, two thematic workshops were held to gain an overview of the data stocks available and their analytical potential, and to ensure data access and identify further data needs. Participants in these workshops included experts from important research, data production and policy areas.

Challenges

The RatSWD's central goals for the short and medium term fall into three main areas: providing a strong and permanent foundation for the data infrastructure in Germany, further improving the data situation in the economic and social sciences, and bringing data production in the social sciences to the top level internationally, which also entails participation in national and transnational planning processes.

Of immediate relevance, firstly, is the need to find creative solutions to institutionalizing and financing these modern data centers. The evaluation of research data centers and data service centers has now been completed and the BMBF project funding is coming to an end. The first measurable success in providing a permanent funding for the data infrastructure was the transformation of the Federal Labour Agency's Research Data Center, located at the Institute for Employment Research (IAB), into a regular department of this agency.

To further improve the data situation for the social sciences and economics, the RatSWD has activities planned for various areas of data production. Efforts will be continued to

open up data sources that possess great potential but have not been systematized thus far, particularly in the areas of education and health. Moreover, the RatSWD is backing plans for the establishment of a regular national election study (*Nationale Wahlstudie*) – currently missing in Germany – that has the potential to offer important data for political science research.

Furthermore, the RatSWD will continue cultivating dialogue with professional scientific organizations (*Fachgesellschaften*). The exchange of experiences and ideas with scientific fields which are not represented on the Council and with representatives of other data-producing facilities (e.g. the Fraunhofer Institute for Systems and Innovation Research in Karlsruhe) has proven immensely fruitful. Future plans also include devoting more attention to the improved access to data from federal ministerial research facilities (*Ressortforschungseinrichtungen*), an issue currently being discussed by the German Science Council.

Overall, however, the RatSWD aims to shift its main focus from increasing data access to improving data production and data quality. This goal set the agenda for a workshop held in March 2007 entitled “The Influence of the Science Community on the Compilation of the Official Statistics with the 2010/2011 Census as Example”. Here, participants discussed whether and how the research community can legitimately and systematically influence the compilation of data for official statistics. Their conclusion was that work would be required within each of the individual research disciplines to achieve a common understanding and efficient census approach. Monitoring of this process is to be carried out by scientific organizations and initiated by the RatSWD. With regard to the need to ensure high-quality data reflecting complex social interrelations, the workshop dealt in depth with possibilities for combining various data sources and the resulting methodological, legal and ethical issues.

However, an improved statistical infrastructure is needed not only on the national level. As the European research landscape evolves, it produces increased demands on the data infrastructure in order for the social sciences and economics to develop their full potential in the area of social comparisons as well. By actively participating in important developments at both the national and international levels, the RatSWD intends to work even more intensively in this important field in the future. It already provides a platform for a fundamental discussion and planning process that is almost one of a kind both in Europe and beyond. If international and interdisciplinary strategic planning is to be successful in fostering empirical research and improving the research infrastructure, however, greater involvement of the professional scientific organizations representing the social sciences and economics will be urgently needed. The RatSWD will endeavor to promote this involvement.

European research infrastructures have played an important role in the natural sciences for some time. The CERN nuclear research center may be the most prominent example, having been founded as early as 1954. In the humanities and other disciplines such as cultural studies, however, this tradition is almost entirely lacking. The same cannot be said for the social sciences—including economics—but the few infrastructure facilities that do exist (especially ESS: European Social Survey and SHARE) in these fields of research on both the national and international level do not display a unified, systematic approach.

To remedy this problem, the European Union (EU) has now accelerated the systematic planning process with the European Roadmap for Research Infrastructures (ftp://ftp.cordis.europa.eu/pub/esfri/docs/esfri-roadmap-report-26092006_en.pdf). The RatSWD was set up at an opportune moment to take advantage of this development, but most of the process of discussion and planning still remains to be done by the RatSWD, the statistical offices, and the scientific community within the social sciences and economics. The first important step in this direction could, for example, be the organization of an international conference offering prominent representatives of the social and economic research community a forum for dialogue on the strategic direction of empirical research and the data infrastructure necessary.

To better understand what would be required for this strategic planning process to succeed in Germany, a commitment of support from the professional scientific organizations (*Fachgesellschaften*) will also be needed. In this context it will be important to recall that since 2006, the German Science Council has not only had the task of evaluating “large-scale scientific instruments” but also of evaluating and if necessary making recommendations for “research infrastructures” within the hard sciences in all disciplines. As in the two areas outlined above, it is evident that the RatSWD is performing an important function at the nexus of research and data production here as well. To ensure that these efforts bear fruit, the strategic involvement of professional scientific organizations representing the entire spectrum of the social sciences will be essential.