

Hahlen, Johann

**Working Paper**

## Proposals for a permanent institutional guarantee of the official informational infrastructure in Germany

RatSWD Working Paper, No. 100

**Provided in Cooperation with:**

German Data Forum (RatSWD)

*Suggested Citation:* Hahlen, Johann (2009) : Proposals for a permanent institutional guarantee of the official informational infrastructure in Germany, RatSWD Working Paper, No. 100, Rat für Sozial- und Wirtschaftsdaten (RatSWD), Berlin

This Version is available at:

<https://hdl.handle.net/10419/186207>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*



German Council for Social  
and Economic Data (RatSWD)

[www.ratswd.de](http://www.ratswd.de)

# RatSWD

## *Working Paper Series*

Working Paper

No. 100

Proposals for a permanent institutional  
guarantee of the official informational  
infrastructure in Germany

---

Johann Hahlen

---

July 2009

---



Federal Ministry  
of Education  
and Research

## Working Paper Series of the Council for Social and Economic Data (RatSWD)

---

The *RatSWD Working Papers* series was launched at the end of 2007. Since 2009, the series has been publishing exclusively conceptual and historical works dealing with the organization of the German statistical infrastructure and research infrastructure in the social, behavioral, and economic sciences. Papers that have appeared in the series deal primarily with the organization of Germany's official statistical system, government agency research, and academic research infrastructure, as well as directly with the work of the RatSWD. Papers addressing the aforementioned topics in other countries as well as supranational aspects are particularly welcome.

*RatSWD Working Papers* are non-exclusive, which means that there is nothing to prevent you from publishing your work in another venue as well: all papers can and should also appear in professionally, institutionally, and locally specialized journals. The *RatSWD Working Papers* are not available in bookstores but can be ordered online through the RatSWD.

In order to make the series more accessible to readers not fluent in German, the English section of the *RatSWD Working Papers* website presents only those papers published in English, while the German section lists the complete contents of all issues in the series in chronological order.

Starting in 2009, some of the empirical research papers that originally appeared in the *RatSWD Working Papers* series will be published in the series *RatSWD Research Notes*.

The views expressed in the *RatSWD Working Papers* are exclusively the opinions of their authors and not those of the RatSWD.

The RatSWD Working Paper Series is edited by:

Chair of the RatSWD (2007/ 2008 Heike Solga; 2009 Gert G. Wagner)

Managing Director of the RatSWD (Denis Huschka)

# Proposals for a permanent institutional guarantee of the official informational infrastructure in Germany

**Johann Hahlen**

*(johann.hahlen[at]web.de)*

1. A permanent informational infrastructure must be based on the situation in Germany. This means it must – to the extent possible – take account of the variety of data, of the multitude of data producers and, especially, of the domestic and foreign potential users and ways of use they envisage and it must be open for future-oriented topics and new questions.

When setting up a permanent informational infrastructure, it must be taken into account that there is a network of interaction between data and their users, which in Germany is determined by a number of legal and actual framework conditions. It is realistic to say that those framework conditions cannot be changed and, consequently, it is reasonable to treat them as given. The reflections in terms of scientific policy to set up a permanent informational infrastructure should take account of the following conditions:

- For natural persons, the German *Grundgesetz* (GG – constitution) grants the right to informational self-determination protecting individuals from unlimited collection, storage, use and transmission of their personal data and safeguarding the individuals' right to decide by themselves about disclosing and using their personal data. Although, where data collected for statistical purposes are concerned, the German constitution does not require their use to be strictly and concretely linked to a specific purpose, it does require relevant limits within the information system. Transmitting statistical data for scientific purposes is in line with the constitution if this is kept within the limits of what is necessary for scientific purposes, if direct reference to persons is avoided (no names or addresses), and if the recipient regularly does not have any additional knowledge that – through re-identification – may result in a violation of the relevant persons' right to informational self-determination. This was laid down by the Federal Constitutional Court in its fundamental population census judgment of 15 December 1983 (BVerfGE 65 1 et seqq.). This requirement is met by the clause relating to the scientific community in Art. 16 Para. 6 *Bundesstatistikgesetz* (BStatG – Federal Statistics Law). Local units, enterprises and legal persons engaged in economic activity cannot claim the right to informational self-determination. However, they are protected by the right regarding any business activity actually set up and performed, which is also granted by the constitution.
- The scientific use of personal data and of data on economic entities must comply with those constitutional rights, the numerous legal provisions on the collection and use of statistical data and the regulations protecting local units and enterprises with regard to

their economic activity (e.g. the protection of business secrets or fair and open competition on the market).

- Data producers and data holders – to the extent that they are part of the public administration, e.g. authorities or institutions – are bound by the principle of the rule of law according to Art. 20 Para. 3 GG. No informational infrastructure of any kind whatsoever and no scientific demand can exempt such data producers and data holders from complying with the above regulations.
- This remains unaffected by the freedom of science, research and teaching guaranteed by Art. 5 Para. 3 GG. It is true that the Federal Constitutional Court has derived from that basic right of the freedom of science (Art. 5 Para. 3 GG) the obligation for the government to provide efficient institutions to maintain free science and the relevant teaching. However, this does not mean that an individual scientist might have a claim to access specific data stocks and, even more so, it does not mean that the scientist's research might have priority over the protected legal rights of persons or enterprises.
- Germany is a federation (Art. 20 Para. 1 GG) in which the exercise of state powers generally is a matter of the Länder (federal states) (Art. 30 GG). The Länder are generally responsible for executing federal laws (Art. 83 GG). The Federation, which – according to Art. 73 Para. 1 sentence 11 GG – has the sole legislative power for “statistics for federal purposes”, was allowed by Art. 87 Para. 3 GG to establish the Federal Statistical Office as an independent superior federal authority. However, the federal legal provisions on official statistics are implemented by the Länder through their own administration (Art. 84 Para. 1 GG). At the same time, Germany has opened up to European integration and has transferred sovereign powers to the European Union (Art. 23 Para. 1 GG), so that EU Regulations and Directives are directly applicable in Germany or have to be transformed into German law. Therefore, EU Regulation No. 322/97 on Community statistics and EU Regulation No. 831/2002 concerning access to confidential data (of the EU) for scientific purposes are directly applicable in Germany. Also, the new EU Regulation on statistics – which is currently being prepared by EU bodies within the scope of the statistics article 285 of the Treaty establishing the European Community – will become applicable law in Germany.
- Finally, the principle of democracy, which is explicitly referred to by the German constitution (Art. 20 Para. 1 and 2, 21 Para. 1, 28 Para. 1 and 38 Para. 1 GG), requires a free, open, transparent and discursive process of forming opinions, which needs both the

knowledge of the facts relevant for the decision-making concerned, especially the data available, and the scientific examination and processing of those facts.

2. So there are many reasons indicating a need for a permanent informational infrastructure in Germany which goes beyond the existence of a free, non-government press and radio landscape and a free, self-determined scientific system and can use the data of official statistical institutions and – if possible – any other data stocks collected for government purposes, while safeguarding the protective rights of the entities to which the data refer (persons, local units, enterprises). However, that need – on which consensus can easily be reached in abstract terms – must cope with a number of very real weak points:
  - On the one hand, the Federal Chancellor Dr. Angela Merkel is said to have stated that “any policy starts with the facts”. On the other hand, we might just as well quote the former Saxonian Minister President Prof. Dr. Biedenkopf, having talked about a “resistance to facts” being widespread among politicians. Keynes is said to have said the following: “There is nothing a government hates more than to be well-informed; for it makes the process of arriving at decisions much more complicated and difficult.” Probably, however, the impression of there being a “resistance to facts” is merely due to the fact that the acting persons think they are sufficiently informed anyway, while frustrated statisticians/social scientists overestimate the importance of their findings.
  - In any case, it is obvious that empirical social and economic research in Germany has clearly been underfinanced for a long time already when compared with other branches of science also working empirically such as medicine and other natural sciences.
  - What is more, there has been – in part – an obvious reluctance to work empirically in German economic sciences.
  - And finally, for some scientists – by no means only in Germany – a certain reluctance is observed to scrutinise their own work for reproducibility and falsification. If this is combined with an – albeit human – attitude of competition and isolation, government agencies will not really be encouraged to invest in costly infrastructures for such scientific actors.
3. Considering those conditions and structures, the permanent informational infrastructure to be guaranteed can be defined in both negative and positive terms.

### 3.1 This is what a permanent informational infrastructure in Germany should not be:

- The public data producers and data holders belong to different levels of the state structure and, in many cases, are structured by Länder or other regional units. For example, in addition to the Federal Statistical Office there are 14 Land offices producing and storing statistical data. The Federal Employment Agency and the German Federal Pension Insurance are part of the indirect federal administration; education data are stored by the competent Land ministries; the Central Register of Foreigners belongs to the Federal Office for Migration and Refugees, which is a superior federal authority; the population registers are available at the towns and municipalities or at central Land population registers. Health data come from some 100 quite different sources.
- That patchwork is anything but comfortable for anyone interested in data for scientific reasons; it is rather confusing and at best labour-intensive. Therefore, the idea of an institution that is comprehensive in regional and subject-matter terms seems obvious, but it turns out to be an unachievable vision.

As experience shows, the various data producers and data holders in Germany are not willing to transmit their data to third parties or even to grant third parties the right of use. At the most, they are willing to be represented by a regional partner (one Land for several or all other Länder). A highly structured informational infrastructure is certainly not comfortable but the cross-reference options of modern IT can make things clear and allow orientation within the patchwork.

- A register comprising all data, such as a large central archive where all the data producers and data holders store duplicates of their data, would theoretically be a solution to the dilemma described above – but it fails due to the legal situation. This is because, in Germany, data are strictly linked to a specific purpose to protect the persons, local units or enterprises referred to by the data. This means that, already during data collection, it must be defined and communicated to the respondents for what purposes their data are collected and to whom they will be made accessible. Transmission of originals or duplicates to a “central scientific register” has so far not been covered by law and could be regulated only for the future. Consequently, no stock data could be stored in such a register unless all respondents gave their consent, which makes the whole matter unrealistic. Exceptions in this respect are not permitted by the clause relating to the scientific community as stated, for example, in Article 4a Para. 2 of the Federal Data Protection Act.



However, statistical data which are processed and used only in an anonymised form do not need to be linked to a specific purpose, so that they can be used for scientific purposes if anonymity (even de facto anonymity) is safeguarded. This does not yet allow to set up a comprehensive central scientific register because what could be stored there would be only aggregated data and microdata only in a de facto anonymised form. Although the latter is possible – with sometimes considerable efforts – for specific data stocks such as the microcensus, it is not possible for all official statistical data. Therefore, a central register limited to statistical data would be highly incomplete. The health monitoring system operated by the Robert Koch Institute and the Federal Statistical Office is not an example to the contrary because it uses only aggregated data from the various sources.

- What should not be envisaged to guarantee the informational infrastructure is the creation of a new federal authority or – either in addition or alternatively – of new Land offices. First, for the reasons shown above, they could not represent a central register. Second, this would involve considerable bureaucratic efforts; they would have to be integrated into existing responsibilities and hierarchies, they would have to acquire the required wide range of special knowledge on the various data stocks and would be limited to co-ordinating activities, while scientific data users would still have to deal with the relevant data producers and data holders.
- There are the same reasons against the attempt of putting the informational infrastructure on a permanent basis through a university institute of some kind or through one or several professors. The existence of GESIS and its practical success at the same time show the limits of such institutions. A university institute or a team of scientists would not be able to cope with those requirements.
- Also, it is not promising to use private-law institutions funded externally to permanently guarantee the informational infrastructure. As experience in Germany shows, the financial means of potential users (from the scientific community) would not be sufficient to pay the considerable staff required for such institutions to offer services meeting the wide range of requirements. It is not expected that the empirical social and economic research community will soon get considerably more funds from the relevant organisations to be in a position to set up or keep up such institutions by demanding their services.

3.2 What should a permanent informational infrastructure in Germany be like if the models rejected here are not considered and if maximum use for data users, especially

from the scientific community, is to be achieved?

- Considering the conditions of modern IT, the infrastructure must be available online 7 hours a day on 365 days a year. Where online use is not possible because of data protection and statistical confidentiality, local workstations should be kept available for use at hours common at universities.
- The infrastructure should ensure that it is equally open to anyone interested and that it runs neutrally, i.e. that it does not assess or censor user requests. It should be independent in its methodical work and be based only on accepted scientific standards. The openness, neutrality and methodical independence should each be supervised by a committee comprising representatives of data producers and scientific data users as well as the competent data protection commissioner.
- The infrastructure should be sufficiently equipped with staff and material to fulfil its tasks. At the same time, it should be lean and economical, so that it can be used without insurmountable financial obstacles. Its work should be rationalised through permanent evaluation of its processes and through wide-ranging use of IT.
- Considering the manifold subject-matter and regional breakdowns of data production and storage in Germany and the fact that centralisation is unachievable, the infrastructure should be structured in terms of subject-matter, it should cover all of Germany and it should be broken down into regions only to the extent absolutely required (e.g. by Länder).
- Although the infrastructure should be organised in a permanent manner, it should also be able – for example through revision clauses – to react flexibly to changes in the data offered and in the demand from the scientific community.
- In all this, it is necessary in practical work that the infrastructure institutions achieve an optimal reconciliation between, on the one hand, the legitimate interests of data producers and data holders as well as the rights – protected by provisions on data protection and statistical confidentiality – of the persons, local units and enterprises referred to by the data and, on the other hand, the interests of the scientific users. Keeping this constantly in mind will be one of the main tasks of the committee set up for the relevant infrastructure institution, in addition to the tasks mentioned above.

4. The institutions set up in Germany on the basis of the recommendations of the *Kommission zur Verbesserung der informationellen Infrastruktur zwischen Wissenschaft und Statistik* (KVI – Commission on Improving the Informational Infrastructure between Science and Statistics) of 13 March 2001 and with considerable support by the *Bundesministerium für Bildung und Forschung* (BMBF – Federal Ministry of Education and Research) have basically proved successful:

4.1 The *Rat für Sozial- und Wirtschaftsdaten* (RatSWD – Council for Social and Economic Data), where data producers and data users work together, has developed into an institution enhancing in a creative manner the informational infrastructure in Germany. In that council, representatives of the Federation and the Länder co-operate with persons elected in a “grass-roots” manner from the scientific community. Therefore, its proposals are practical and are welcomed. Apart from its internal work, such as exchanging opinions with the major institution funding research (BMBF) or evaluating things towards institutions of official statistics, the RatSWD is engaged in many external activities which have become important elements of the informational infrastructure in Germany and should be continued.

What should be mentioned first of all here is the Conference for Social and Economic Data which is held at regular intervals and where – apart from electing the council members from the scientific community – research results are presented that have been obtained through the data stocks made available and where gaps in the informational infrastructure are discussed.

Major suggestions on how to improve the informational infrastructure are given by the expertise contests organised by the RatSWD and the working papers and newsletters issued by the RatSWD.

4.2 The most important progress that has been made since the KVI gave its recommendations in 2001 has been the setting up of the four Research Data Centres (RDCs) and the two Data Service Centres (DSCs).

- The RDC at the Federal Statistical Office was founded in 2001 – as the first RDC in Germany – and was positively assessed in 2004. It allows empirical social and economic researchers to access official statistical microdata, while safeguarding statistical confidentiality. For that purpose, the RDC makes Public Use Files, Scientific Use Files and Campus Files available for off-site use by the research and

teaching community. Guest scientists can use less strongly anonymised data on the premises of the Federal Statistical Office in Wiesbaden, Bonn and Berlin. Also, scientists can use data stocks of the Federal Statistical Office by means of controlled teleprocessing (on-site use).

- The decentralised RDC of the statistical offices of the Länder was set up in April 2002, was positively assessed in late 2006 and offers scientists the same access to official statistical data as shown above for the RDC of the Federation. Subsequent to an amendment of the *Bundesstatistikgesetz* (BStatG – Federal Statistics Law), the statistical offices of the Länder established for that purpose a system of centralised data storage for the whole of Germany, with a breakdown by subject-matter.
- The RDC of the Federal Employment Agency was established in April 2004 at the Agency's *Institut für Arbeitsmarkt-und Berufsforschung* (IAB – Institute for Employment Research) in Nuremberg and has also been assessed positively. It makes the large data stocks of the Federal Employment Agency available for scientific analyses within the scope of Art. 75 of Volume X of the Social Code.
- The RDC of the German Pension Insurance was also established in 2004 with its two locations in Berlin and Würzburg. The Scientific Use Files produced there with regard to the statistics of new and existing pensions and the statistics of persons insured allow for the first time scientific evaluation of the vast data treasures of the German Pension Insurance.
- The two Data Service Centres – also based on the KVI recommendations of 2001 – were established in 2003 at GESIS in Mannheim and at the *Forschungsinstitut zur Zukunft der Arbeit* (IZA – Institute for the Study of Labor) in Bonn. The GESIS DSC works under the name of German Microdata Lab (GML) and offers a service and research infrastructure for official microdata.

The *Internationales Datenservicezentrum für arbeitsmarktrelevante Daten* (IdZA – International Data Service Centre for Labour Market relevant Data) at the IZA supports labour market researchers especially through a metadata portal for existing data; it has developed a special web-based tool (JoSuA) for data access via controlled teleprocessing.

All RDCs and DSCs have very much been welcomed by the scientific community and are intensively used for research and teaching, with the two RDCs of official statistics having

observed a marked recent shift in the demand for ways of access to their data stocks: While the – initially very big – demand for Scientific Use Files has been stagnating, demand is increasing for individual data sets, with which guest scientists can work at safe scientific workstations at the RDCs, and for controlled teleprocessing.

The encouraging practical efficiency of the RDCs has two major causes:

- Thanks to the start-up financing by the BMBF, the RDCs have made a wealth of official statistical data stocks available to the research and teaching community by producing Public Use Files, Scientific Use Files and Campus Files, by offering safe scientific workstations for guest scientists, and by offering controlled teleprocessing.
- The financial obstacles existing in the 1990s, which in part were insurmountable for social scientists working empirically, regarding the use of data of official statistics have been removed, which is also thanks to the start-up financing of the RDCs by the BMBF. For example, in the mid-1990s the statistical offices had to charge some DM 30,000 for making available a Scientific Use File of the microcensus to justify the considerable costs required from their budgets for its production. Since there have been RDCs, a social scientist can get there such a Scientific Use File for a “charge” of EUR 90 covering the CD and its forwarding.

4.3 The informational infrastructure developed since the KVI recommendations of 2001 also includes many larger and smaller projects and initiatives of the most different institutions, such as:

Every year since 1999, the Federal Statistical Office has been granting the Gerhard Fürst Award for dissertations and diploma/master theses dealing with empirical questions and using data of official statistics.

The statistical offices of the Länder have set up branches of its RDC at the *Deutsches Institut für Wirtschaftsforschung* (DIW – German Institute for Economic Research) in Berlin and at Dresden Technical University.

Within the scope of their conferences, the German Statistical Society organises workshops for junior scientists to introduce them to empirical work with the various data stocks.

5. Despite all the progress made so far, there still is much to improve and numerous problems remain to be solved. In Germany we have not yet succeeded in permanently

guaranteeing an adequate informational infrastructure institutionally. Financial and content-related problems need to be solved.

5.1 Financial problems appear to be most urgent at the present time and, although they are not at all excessive as to their volume (the RDC of the statistical offices of the Länder, for instance, reckons with total costs of only about EUR 3.7 million for the 2 1/2 years from 1 July 2007 to 31 December 2009), they are difficult but can be solved.

- The structures created on the basis of the KVI recommendations of 2001 (especially the RatSWD with its secretariat in Berlin and the four RDCs) owe their establishment to the support provided by the BMBF. This was temporary project support in the form of start-up financing that requires the relevant institution to contribute funds of its own, considering the benefit it draws from the project.
  - The German Socio-Economic Panel (SOEP), which meanwhile is 25 years “old“, has been financed institutionally since 2004. Thus an important recommendation of the KVI has been implemented and its work can be regarded as permanently guaranteed. In contrast, such institutional support seems out of reach for the RDCs but it is not necessary after all.
  - The financial situation of the RDCs varies considerably at the present time.
  - At the beginning, the RDC of the Federal Statistical Office was financed mainly by the BMBF. Meanwhile its core business, answering and handling user requests from the scientific community, is funded completely from its own budget. The RDC receives BMBF funds only for research projects to extend the data supply it offers, for instance by anonymising panel data of economic statistics.
  - Most of the funds required for the RDC of the statistical offices of the Länder will be provided by the BMBF until the end of 2009.
  - The RDC of the Federal Employment Agency at the IZA was partly financed by the BMBF and since the beginning of 2007 is funded entirely by the Federal Employment Agency.
  - The RDC of the German Federal Pension Insurance will be supported by the BMBF until the end of 2008.
- Consolidation and a uniform financing line for the RDCs are therefore urgently required. On the one hand, they would have to guarantee the ongoing existence of the RDCs and the

further development of their data supply. On the other, there must not be prices again for using the RDCs that users cannot afford. It is thanks to the KVI recommendations and the project support by the BMBF that this – harmful – situation no longer exists in Germany. After all, the scientific community should be able to use the respective data stocks for research and teaching purposes. At the same time one will have to accept that the BMBF generally confines itself to temporary start-up financing and regards the respective data holders and interested scientists as responsible.

- Therefore the organisations supporting the RDCs, the empirical social and economic research institutions and the BMBF should agree on the following model which should entail sustainable financing of the RDCs at affordable prices for their users:
  - The respective organisations supporting the RDCs, for example the statistical offices of the Federation and the Länder, will take over the basic financing of their RDCs.
  - The further development of methodology and special research projects of the RDCs will continue to receive project funding on a temporary basis, provided that these are important for an expansion of the informational infrastructure.
  - The RDCs will charge users to cover the expenses incurred in each case, but there will be far-reaching possibilities to reduce prices for financially “weak“ users such as Ph.D. candidates or university institutes, while “well equipped“ users, for instance economic research institutes, which can pass on their expenses to their clients, will have to pay prices fully covering the expenses.
- To accompany this solution at the statistical offices of the Federation and the Länder, it would be advisable to supplement the Federal Statistics Law, making it clear that the mandate of official statistics includes also the provision of data (both aggregated data and microdata) to the scientific community. The inclusion of such a provision into one of the next bills on statistical issues should be supported at the Federal Ministry of the Interior. After the co-operation of the statistical offices in a RDC with jointly held data was enabled by the Federal Statistics Law in 2005 (through its Art. 3a Para. 2 and Art. 16 Para. 2), the RDCs of official statistics would thus eventually be enshrined in law and their funding be indirectly guaranteed.

5.2 As regards its contents, the informational infrastructure which has emerged in Germany since 2001 provides numerous starting points for expansion and

consolidation. Depending on the perspective, priority is given to one point or another. Priorities and posteriorities should be discussed in the RatSWD and a medium-term consolidation and extension programme should be set up, focusing not only on what would be desirable but also on what chances there are to implement it. The order of the following presentation is therefore not meant as an order of preference.

- The existing four RDCs are far from opening up all data stocks which are of interest to empirical social and economic research. This is why there should be RDCs for instance also for health, education and media data. Crime control, the administration of justice and penal administration, for example with the criminal statistics of the police and judicial statistics, also are large subject fields awaiting further exploration. The situation is similar with the Central Register of Foreigners kept at the Federal Office for Migration and Refugees, the business register of the Federal Statistical Office and the population registers of the municipalities and the Länder. Finally, provisions have to be made in time for the scientific use of the data to be collected in the EU-wide population census scheduled for the year 2011. It would have to be analysed for all these areas whether RDCs should be set up and, if so, their establishment should be furthered.
- As there are different RDCs, each of them restricted to specific data stocks, it is demanded that a “special“ RDC be set up which combines the data stocks of various data producers or makes it possible to work with the data of different producers. A similar goal is pursued by the proposal to create a kind of “data trust” keeping data stocks from various subject fields and making them accessible to the scientific community via the channels known from the RDCs. Advantageous as both ideas may be from the viewpoint of empirical social and economic research, the obstacles of data protection legislation appear insurmountable so that one should not “fight a losing battle“ here.
- Such a solution might be considered, if at all, for statistical data whose collection does not have to be strictly linked to a specific purpose. But then the data kept there would have to be at least de facto anonymised. This, however, would probably not be worthwhile. Also, it must not be disregarded that combining de facto anonymised personal data from different statistics increases the chances of reidentification, which is exactly what must be prevented.



- Non-statistical data, however, have to be strictly linked to a specific purpose. This means the following: First the data - and also the microdata - of the various producers would have to be transferred to the “special“ RDC or the “data trust“. So far this transaction would generally not be covered by the respective data collection purpose and therefore be illegal. The clauses relating to the scientific community as contained in the Federal Data Protection Act (e.g. Article 14 Para. 5 No. 2) do not permit such data transmission and storage because the research purposes can actually be achieved with reasonable efforts even without a “special” RDC or without a “data trust“. The proposal to appoint the data protection commissioner in charge as trustee does not solve the problem. Apart from the fact that the Federal Commissioner for Data Protection has already dismissed such ideas for his institution, the unsolvable problem of having to alter the purpose would persist. If - despite all practical obstacles - the consent of all concerned to such a purpose-altering transfer could be obtained, reservations would remain because contrary to the order of the Federal Constitutional Court, the data would not be de facto anonymised at the earliest possible time.
- In view of this legal situation it would rather be advisable to invite a scientist, for example from the Federal Employment Agency or its Institute for Employment Research (IAB) to the RCD of the Federal Statistical Office and to entrust him or her with “data processing by order“, with the evaluation of statistical data in combination with data of the Federal Employment Agency in relation to a specific issue. The RDC of the Federal Statistical Office plans to do that with regard to the data of the Federal Employment Agency.
- As the view prevails that the clause relating to the scientific community in Art. 16 Para. 6 of the Federal Statistics Law does neither include foreign universities nor foreign scientists, the informational infrastructures created in Germany to date have not furthered the scientific co-operation with foreign countries, and this also holds for the EU. It is true that there meanwhile is a “Safe Centre“ at Eurostat in Luxembourg whose establishment was made possible by Regulation (EC) No 831 / 2002 (concerning access to confidential data for scientific purposes). However, German statistical microdata would be available there only if they had been submitted also to Eurostat, which is an exception. Therefore the establishment of such an “EU Safe Centre“ in Wiesbaden, which is planned by Eurostat together with the Federal Statistical Office, will not bring any improvements for foreign

scientists. To enable cross-border scientific work, empirical social and economic researchers should call for an extension of Art. 16 Para. 6 of the Federal Statistics Law to cover foreign scientists or they should support the inclusion of a comprehensive clause relating to the scientific community - to cover at least all scientists from the EU - into the revised EU Regulation on Community Statistics (No 322/1997), which is currently being deliberated.

- There has not been any progress in the last few years regarding the KVI recommendation of 2001 to introduce research data or scientific secrets. The restraint shown with regard to this suggestion may be due to the fact that such a research data secret has to be connected with a privilege of the scientist to decline to answer questions, and that seizure must be prohibited. However, this recommendation still deserves to be studied in detail. Because of the complexity of the matter, the RatSWD should set up a working party for the purpose. After the recent cases of data abuse at a large telecommunication provider and in call centres, voices to be taken seriously call for a codification of the right to informational self-determination and for the codification of a right to privacy of information technology records. If these attempts should materialise, the scientific community should have to be able to put its interests forward with an elaborate proposal to introduce research data or scientific secrets. Maybe it would be easier to make some progress in this difficult matter if a code of conduct existed for scientists interested in using the data stocks related with the possibility to impose sanctions, which had also been recommended by the KVI in 2001. The RatSWD should take steps also in that direction together with the other scientific institutions.
- Finally, the KVI recommendations of 2001 deserve further efforts as far as they aim at an expansion of empirical social and economic research (including university education preparing for the subject). Beyond the establishment of “empirical economic research“ as a university subject there is a sufficient number of current problems justifying for instance the creation of special research areas (for example on questions of health and education policies) or of professorships for empirical work (co-)financed by trusts.
- When the informational infrastructure is expanded on a permanent basis, continuous checks for “proliferation“, overlaps, duplication of labour and the like

must not be neglected in the course of and apart from the meanwhile common and rather strict periodical evaluation of the facilities created. Experience shows that these have to be expected especially with new developments, while the readiness to carry out necessary adjustments is generally not so well developed. In particular the informational structures resulting from federalism should be analysed in this respect.