Activities Report 2017
of the Research Data Centres (RDCs)
accredited by the
German Data Forum (RatSWD)
Activities Report 2017
of the Research Data Centres (RDCs)
accredited by the
German Data Forum (RatSWD)
Contents

Preface ................................................................................................................................. 7
1 Overview of the German Data Forum’s (RatSWD) research data infrastructure ........ 8
2 Characteristics of the research data centres (RDCs) ............................................. 11
3 Data protection, archiving and quality assurance .............................................. 14
4 Data provision ............................................................................................................... 18
5 Services for users .......................................................................................................... 22
6 Current trends regarding the research data infrastructure in the social, behavioural, and economic sciences: how RDCs train and retain data users .......... 29
7 Complaints management ............................................................................................... 32

Appendix

Appendix A: Development of the German Data Forum (RatSWD) research data infrastructure .... 34
Appendix B: The monitoring commission ........................................................................ 40
Appendix C: Contributors ................................................................................................. 41
3,571 Datasets
357 new datasets

21,467 Full-time staff

69,083 Publications based on the available data

2,042 Datasets

Locations and guest researcher workstations of accredited research data centres
Most RDCs keep their German acronym FDZ

*  RDCs with guest researcher workstations
**  External RDC guest researcher workstation

Fig. 1: Key figures in 2017
The data below were provided by 31 RDCs for 2017

© RatSWD 2018
Preface

The Activities Report of the research data centres (RDCs) accredited by the German Data Forum (RatSWD) provides a comprehensive overview of the centres’ dynamic development. The many interconnections between the individual RDCs (see Fig. 3), the wide distribution of centres across Germany, and the numerous on-site work stations for guest researchers are testimony to an active and growing infrastructure.

The positive development of the research data infrastructure is evident, first and foremost, in some of the figures for 2017 (see Fig. 1). The 31 RDCs offered a total of 3,571 datasets. Of these, 357 datasets were newly added in 2017. The datasets were used by more than 21,000 active data users, generating 69,083 downloads of openly accessible data. This is impressive evidence of the relevance of the RDC data to the social, behavioural, and economic sciences in Germany. What is more, at least 2,042 scientific publications were produced with the help of RDC-held datasets in 2017 alone. RDC staff made a significant contribution to this output: the pool of 264 full-time equivalents at the RDCs produced a total of 459 publications in 2017. This reaffirms the philosophy of RDCs combining in-house research and infrastructure to provide top-quality services for data users.

A key service provided by the RDCs is to support scientists in using data. This is reflected in a total of 100 training opportunities (including online and offline courses, workshops, and conferences), which in 2017 were used by some 1,700 participants. The main emphasis of these training opportunities was on making researchers – and especially the international scientific community – more aware of the available data and their research potential (Fig. 21). This is one of many aspects underscoring our goal to strengthen the research data infrastructure on a national and international level.

While the expansion of the research data infrastructure is a clear goal, the accreditation of RDCs is based on well-defined quality guidelines established by the German Data Forum (RatSWD). The annual monitoring audit serves to ensure that previously accredited RDCs comply with these quality criteria, suggests opportunities for improvement, requires a long-term commitment, and thereby ensures high standards on an ongoing basis. To communicate this standard to international audiences as well, we plan to explain how our criteria comply with the FAIR principles, which are attracting more and more attention at the national and international level. After all, Findability, Accessibility, Interoperability, and Re-usability are tightly woven into the philosophy and accreditation criteria of the German Data Forum (RatSWD) and the network of research data centres and have long been part and parcel of our daily practice.

The German Data Forum (RatSWD) and the RDC Committee wish to express their gratitude to the members of the monitoring commission for their dedication in preparing this report. The monitoring commission’s push for refining the monitoring process over the past three years has helped to improve the infrastructure’s profile, thereby enabling data users to clearly recognise the capacity of the research data infrastructures. Because the monitoring process was optimised in each of the past three years, year-by-year comparisons are not yet possible for some figures owing to modified indicators. Lastly, we would like to thank the team of the German Data Forum (RatSWD) business office for their excellent support in the monitoring process and in putting together the publication.

Dr. Pascal Siegers
(Chair of the RDC Committee)

Dr. Jan Goebel
(Deputy Chair of the RDC Committee)
The RDC Committee is a steadily growing, decentralised network of 31 research data centres (RDCs) accredited by the German Data Forum (Rat für Sozial- und Wirtschaftsdaten, RatSWD) (as of September 2018). The RDC model and German Data Forum (RatSWD) accreditation is a desirable goal for many organisations: in the past year, the RDC for Higher Education Research and Science Studies (fdz.DZHW) was accredited; two more RDCs applied for guest status in the RDC Committee and are pursuing accreditation.

To further expand this unique RDC network in the long run, four strategic fields of action were defined for the research data infrastructure in the work programme of the sixth German Data Forum (RatSWD) appointment period:

- further development of the archiving structure
- further development of data access
- development of common guidelines
- expansion of training opportunities in the RDCs

These long-term efforts are meant to ensure that the research data infrastructure remains competitive in the international arena as well and that it pushes early on for important developments with respect to increasing digitisation and interconnection.

The research data infrastructure accredited by the German Data Forum (RatSWD) provides independent researchers in Germany and abroad with affordable, transparent, and user-friendly access to a wide range of prepared research data. In many cases, these data are sensitive microdata and hence especially worthy of protection. Unless they are anonymised, such data may be re-attributed to individual persons and households. That is why the EU General Data Protection Regulation (GDPR, see info box 1) directly affects the work in the RDCs when it comes to processing personal data. This is especially evident in the requirement to anonymise datasets and the modified policies on informed consent in data processing. Before and after the GDPR, complying with legal requirements while also preserving the high analytical potential of the research data is an ongoing mission for the RDCs. One key task in 2017 was preparing for the GDPR.

---

1 A list of all RDCs is available in the appendix on page 34-39.
2 Likewise, the possibility to re-attribute data to individual firms or legal persons may require anonymisation work at the RDCs in certain legal constellations.
The new General Data Protection Regulation (GDPR) took effect on 24 May 2016 and became directly applicable in all EU member states on 25 May 2018. Due to its strong overlap with existing German data protection legislation, the adjustments required for scientific research in Germany are manageable. The research article (Article 89) reflects the special societal relevance of science in terms of data protection law. New requirements include additional obligations regarding documentation. Compared to the German Federal Data Protection Act (old version), however, the legal definition of ‘de facto anonymised data’ was dropped. This aspect is especially relevant to the RDCs accredited by the German Data Forum (RatSWD). In order to comply with data protection requirements while also preserving the potential for scientific analysis, the data access paths offered by the RDCs are usually based on data with a limited degree of anonymisation (see Fig. 2).

The German Data Forum (RatSWD) sought to clarify the new legal situation. It turned out that, according to various experts, the concept of anonymisation, as before the GDPR, is not to be understood to mean absolute anonymisation. Rather, anonymity is an abstract concept referring to an ongoing risk assessment regarding the possibility of re-identification against the background of current technological developments and openly available data sources. Accordingly, the RDCs may continue to apply the previously introduced concept of ‘de facto anonymisation’. Although the GDPR has now become directly applicable law, the exact legal practice in science and research will only be specified in the months and years ahead through practice and example cases, jurisdiction, and statements issued by the European Data Protection Board. Furthermore, the political processes to implement the EU data protection reforms have not been completed at the national level: the proposed 2nd Act to Adapt Data Protection Law to Regulation (EU) 2016/679 and to Implement Directive (EU) 2016/680 (2. DSAnpUG-EU) is designed to implement numerous individual legal amendments at the federal level. At the state (Länder) level, the respective state data protection legislation and some additional special legislation have yet to be adopted.

Fig. 2: Anonymisation as risk assessment
How it all began:
Following an initiative by the scientific community, the ‘Commission on Improving the Informational Infrastructure’ (KVI) was established in 1999. That commission presented a comprehensive report in March 2001. One of its key recommendations was to set up research data centres at public data producers. In the same year, the German Data Forum’s founding committee was set up.

Evolution and consolidation:
After 2001, the research data infrastructure saw continuous development through the foundation of various RDCs. The reasons that led institutions to establish RDCs were diverse: some followed recommendations issued by commissions (e.g., the KVI), by the German Council of Science and Humanities, or by scientific advisory groups; other RDCs were self-commissioned by their host institutions to help promote research. What all of them had in common, however, was the desire to expand and strengthen the research data infrastructure in Germany by making research data available to the scientific community.

The German Data Forum (RatSWD) was founded in 2004. To promote productive dialogue between the RDCs, the ‘Standing Committee Research Data Infrastructure’ (RDC Committee) was established. The main responsibility of this committee is to continuously secure and improve the research data infrastructure – that is, the quality and quantity of the available data and of data access.

To assure the quality of the research data infrastructure, the German Data Forum (RatSWD) defined minimum standards and accreditation criteria in 2010. The years 2015 to 2017 saw revisions to these standards and criteria, as well as the implementation of regular monitoring and evaluation processes, such as the introduction of a monitoring commission consisting of elected members of the RDC Committee.

German Data Forum (RatSWD) accreditation is a seal of quality for all RDCs. Before accreditation is awarded, an RDC is reviewed for compliance with the mandatory criteria. Moreover, accreditation comes with additional support services: an accredited RDC can network with other accredited RDCs, receives information about best practice solutions to help guide the ongoing development of its own infrastructures, and participates in the exchange of knowledge and experience with other RDCs. The collaborative connections between all accredited RDCs beyond their shared work in the RDC Committee are illustrated in Figure 3 (p. 12).

---


5 For background information and the members of the monitoring commission, see Appendix B.
2 Characteristics of the research data centres (RDCs)

The information summarised in the following chapters is taken from the annual monitoring process, in which all accredited RDCs participate. In the reporting year 2017, 31 RDCs took part in the monitoring.

Tab. 1: Staff employed at the 31 RDCs in full-time equivalents (FTEs)
Information current on 31.12.2017

<table>
<thead>
<tr>
<th></th>
<th>Numbers in FTEs (n=31)</th>
<th>Range in FTEs in 2017</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total staff</td>
<td>264.03</td>
<td>0.4 – 34.6</td>
</tr>
<tr>
<td>Academic staff</td>
<td>154.5</td>
<td>0.25 – 23</td>
</tr>
<tr>
<td>Non-academic staff</td>
<td>67.14</td>
<td>0 – 11.6</td>
</tr>
<tr>
<td>Student assistants</td>
<td>42.39</td>
<td>0 – 6</td>
</tr>
</tbody>
</table>

The differences between the individual RDCs are illustrated by the wide range of staff employed.

On 31 December 2017, the 31 RDCs employed a total of 264.03 persons in full-time equivalents (FTE) (see Tab. 1). This means the number of FTEs increased by nearly 38 compared to 2016. Aside from staff increases in individual RDCs, this growth in FTEs is explained in particular by the fact that information on student assistants was collected in a slightly modified and more precise manner and that the staff of the RDC newly accredited in the reporting year was added to the total count. As a result, 42.39 full-time equivalents in 2017 were covered by student assistants, who are employed by 70% of RDCs. The number of employees varies strongly between the RDCs, as illustrated, for example, by the wide range in scientific staff, which ranges from 0.4 to 34.6 FTEs.
Collaboration and research activities

**Fig. 3: Collaboration between the accredited RDCs in 2017**

Collaboration within and beyond the RDC Committee create a multi-layered, interdisciplinary network of RDCs.

The 31 RDCs engage in close and continuous collaboration in and beyond the RDC Committee (see Fig. 3 on collaborations). The collaborations found for the reporting year 2017 refer to the RDCs accredited by the German Data Forum (RatSWD) (i.e. not to the home institution in its entirety). In 2017, 20 RDCs worked together in specific institutionalised research collaborations. Of the 10 RDCs not involved in such research collaborations, one reported to be preparing for a collaboration.

**Fig. 4: Independent research by scientific staff in 2017**

<table>
<thead>
<tr>
<th>Does the RDC employ scientifically trained staff, who conduct independent research?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
</tr>
<tr>
<td>No</td>
</tr>
</tbody>
</table>

In almost all RDCs, scientific staff also conducts their own research.
Twenty-nine RDCs employ scientifically trained staff who conduct independent research (see Fig. 4). These research activities refer to content-related questions, methodological issues, or technological innovations. As a general principle, it is desirable for RDCs to conduct research of their own in order to be able to provide data, services, and consulting at an appropriate scientific level and in order to be able to offer assistance to data users. At 15 RDCs, scientific staff who do research have a fixed percentage of about one-third of their working hours to be devoted to research.

Fig. 5: Number of scientific publications by RDC staff in 2017

Please indicate the number of scientific publications produced by your RDC staff, regardless of the data used and regardless of whether the publication was prepared during RDC working hours.

<table>
<thead>
<tr>
<th>Category</th>
<th>Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>459</td>
</tr>
<tr>
<td>Peer-reviewed journals</td>
<td>115</td>
</tr>
<tr>
<td>Journals incl. peer-reviewed journals</td>
<td>151</td>
</tr>
<tr>
<td>Monographs incl. edited volumes</td>
<td>9</td>
</tr>
<tr>
<td>Articles in edited volumes</td>
<td>62</td>
</tr>
<tr>
<td>Grey literature incl. technical reports</td>
<td>225</td>
</tr>
<tr>
<td>Theses submitted for a degree</td>
<td>12</td>
</tr>
</tbody>
</table>

RDC staff produces an average of almost nine publications per week.

The RDCs’ scientific staff document their research activities through a number of publications. Having publishing experience of their own is helpful to offer services from the users’ perspective. In 2017, RDC staff produced 459 scientific publications (see Fig. 5), regardless of the data used and regardless of whether the publication was prepared during RDC working hours. A total of 151 articles were published in journals; of these, 115 appeared in peer-reviewed journals, indicating the high quality of the research. Another 62 articles were published as part of edited volumes. The largest share, 225 publications, was published as grey literature, which also includes working papers and technical reports. If these figures are only tied to scientific staff, the RDCs produced almost three publications per full-time equivalent. If non-scientific staff is included, and assuming that all employees have an equal share in the production of publications, the 2017 publication output is still more than one-and-a-half publications per FTE across all RDCs.
3 Data protection, archiving, and quality assurance

Concepts for making data available on a long-term basis
Making research data available on a long-term basis is part of good scientific practice. More and more scientific journals expect that published empirical findings can be replicated. This underlines the responsibility of research data centres to store data on a long-term basis. This RDC activity is important especially because researchers are not allowed to share most of the data provided by RDCs and are often required to delete the data by the deadline defined in their contracts. This requires RDCs not only to physically store the data (in different versions, if necessary) but also to ensure that they can be read, interpreted, and used in the future. One challenge in this endeavour is to store the data in formats that will also be compatible with future technologies. Aside from archiving data, RDCs must also provide protection against data loss and against unauthorised or unintentional manipulation.

Fig. 6: Data preservation at the RDCs in 2017

Does your RDC ensure that data are available on a long-term basis (according to the rules of good scientific practice, for at least ten years after they were last used) and in standard formats?
(Multiple answers possible)

The RDCs take a very comprehensive approach to making research data available on a long-term basis.
All RDCs ensure data documentation, thereby enabling secondary use (see Fig. 6). If they have their own access to prepared data, this access is archived as well. It is only in exceptional cases that raw data are not stored even if the RDC has access to them. In centres where questionnaires or other metadata are available, these are usually archived as well. There are only six RDCs where this is not the case because they do not have access to this type of metadata. Overall, the RDCs strive to make data available to users in the broadest and deepest possible scope. The individual RDCs use different methods to store data in databanks, on servers, or in specialised data archives involving regular backups. The data are archived using different data formats (e.g. CSV, SPSS, or Stata). Specialised data archives use distributed and redundant storage on various media – that is, the archived data are duplicated and stored simultaneously at multiple physical sites and on different storage mediums (hard drives, CDs, etc.).

Twenty-eight centres store their data redundantly on multiple storage mediums. Half of these centres report storing data at different physical locations as well. Only two RDCs exclusively use their in-house servers to store data. Seven RDCs have now obtained certification for their storage solution: six have a ‘Core Trust Seal’, one has a ‘Data Seal of Approval’, and one has gone through both certification processes.6

**Ensuring data protection and research interests**

In the reporting year, four RDCs, presumably in anticipation of the new General Data Protection Regulation, changed their data protection procedures. In one case, this included restricting data access from non-EU countries. In another, output control policies were made more rigorous. But two RDCs also made it easier for users to access their data, for instance, by facilitating registration or by creating an additional guest researcher work station.

---

6 In the reporting year 2017, these were the two key certifications available for repositories. The idea of certification is to provide researchers with guidance regarding the choice of repositories for storing and curating research data. Accordingly, these certifications provide information about a repository’s compliance with transparent standards defined in the certification criteria.
Assuring the quality of the datasets

Fig. 7: Data checks at the RDCs in 2017

Is data checking (checking the quality of the data shared) part of the responsibilities of your RDC?

- Yes, only by the RDC: 5
- Yes, in parts by the RDC: 22
- No: 4

Number of RDCs in 2017 (n=31)

Data checking is among the responsibilities of many RDCs.

At RDCs where the home institution collects the data itself or has others collect them on its behalf, quality assurance starts with checking the survey instruments and each stage of data collection. Usually, the data go through various stages of preparation, during which they are checked for completeness, consistency, and plausibility. If data errors or missing data are detected, these issues are addressed, if possible. In five of the 31 centres, these quality assurance measures are performed exclusively by the RDC’s (see Fig. 7), whereas the majority of centres only perform some of these checks. Four RDCs do not perform any data checks.

Fig. 8: Types of data checks at the RDCs in 2017

Types of data checks in the RDCs in 2017 (Multiple responses possible)

- Value range checks: 21
- Label checks: 22
- Filter checks: 18
- Consistency with data documentation: 24
- Annual comparisons: 13
- Comparisons of marginal distributions: 13

Number of RDCs (n=27) where data checks are fully or partially among the responsibilities of the RDC

The RDCs perform appropriate data checks on the datasets they collect.
The data checks performed at the RDCs include checking for consistency with the data documentation, for compliance with the value ranges, and for filters and labels. About half of the centres also check marginal distributions, or perform annual comparisons (see Fig. 8). 13 of the 27 RDCs that perform data checks have a fixed set of rules to guide their data checking. Moreover, various RDCs document their quality assurance measures on a regular basis in quality reports or data reports to accompany individual data products. This is also reflected in the high number of working papers and technical reports (see Fig. 5) featuring data reports.

**Fig. 9: Data corrections at the RDCs in 2017**

<table>
<thead>
<tr>
<th>Yes, only by the RDC</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes, in parts by the RDC</td>
<td>23</td>
</tr>
<tr>
<td>No</td>
<td>6</td>
</tr>
</tbody>
</table>

Number of RDCs in 2017 (n=31)

© RatSWD 2018

At 25 RDCs, working directly on data is a key responsibility.

Only six RDCs do not generate or correct data (see Fig. 9). The key responsibility here is generating additional variables. Another frequent task is coding missing values. Correcting implausible values and harmonising longitudinal survey data is a core responsibility at 16 and 14 RCDs, respectively. Five RDCs reported they were also charged with imputing missing values. Thirteen of the 25 RDCs where data corrections are performed use a fixed set of rules when correcting data. At 17 of these 25 RDCs, data corrections can also be recognised by users at the case level.
4 Data provision

The number of available datasets has seen continuous growth in recent years.

On the cut-off date (31.12.2017), 3,571 datasets were available at the RDCs. This means that 357 new datasets were added in the reporting year – datasets that received a new digital object identifier (DOI), for example, or would in principle be eligible to receive one (see Fig. 10). The number of newly added datasets varies by RDC, ranging from a one-digit number to 66 new datasets in one RDC. A dataset may consist of multiple individual studies, meaning that the number of new studies available is significantly higher.

Twenty-seven RDCs assign persistent identifiers (PID) to the datasets available to ensure that they can be found and cited in the long term. RDCs that do not currently assign PIDs are in the process of introducing them.
The majority of RDCs allow for early data release without embargo periods.

In 18 RDCs, embargo periods do not exist, meaning that data are made available immediately after they are received and prepared (see Fig. 11). Five RDCs report using fixed embargo periods ranging from six months to no more than two years. At the other RDCs, embargo periods depend on certain conditions, such as the end of a research project or requirements by funding agencies or the data providers themselves.

The majority of RDCs do not charge any fees for scientific users.

Of the 31 RDCs, 23 report not charging any user fees (see Fig. 12). The fees mentioned by the other eight RDCs are mostly in the two- or lower three-digit euro range (sometimes per dataset, survey year, data access path, or depending on the specific amount of work needed to anonymise the data).
A notably higher number of publications was reported in 2017.

For the reporting year, 24 RDCs reported a total of 2,042 publications (see Fig. 13) based on the research data provided. Researchers often forget to notify the RDC about a publication or to send a copy. Moreover, as not every research data centre has the resources to find dataset citations (which are not yet handled consistently by authors either), it is safe to assume substantial underreporting here.

Compared to the preceding year, there was a notable increase in reported publications. Articles in peer-reviewed journals continue to be the type of publication most frequently reported by data users. A significant increase can be observed for research papers produced to earn a degree.
Tab. 2: Data use: type, purpose, and number of data users

<table>
<thead>
<tr>
<th></th>
<th>Number 2017 (n=31)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Free download</td>
<td>69,083</td>
</tr>
<tr>
<td>Scientific projects</td>
<td>5,108</td>
</tr>
<tr>
<td>Data users</td>
<td>21,467</td>
</tr>
</tbody>
</table>

The wide-ranging use of the data provided underlines the relevance of the RDC infrastructure.

One of the core responsibilities of the RDCs is to provide researchers with comprehensive and flexible data access and to keep expanding and improving that access. Researchers have the choice – partly because of varying data protection requirements and legal regulations – between various types of access. Owing to this flexibility in access paths and the differences in the way information on contracts, projects, and data users is collected at each RDC, it is difficult to determine the total volume of data use. For instance, double counting cannot be ruled out regarding contracts and projects or between the various access paths.

Overall, the datasets made available as free downloads were downloaded approximately 70,000 times in 2017 (see Tab. 2). In most cases, prior registration is required for downloading files, but some public use files may also be downloaded without registration. That is why the scope of this type of use cannot be documented by all RDCs.

The standard case, however, is that RDC-held research data are made accessible only after users have registered or signed a contract. For data protection reasons, contracts on data access and usage contain an explicit reference to a research project (purpose limitation), meaning that a contract must be made for each research project that uses such data. In 2017, more than 5,000 contracts were signed.7

In 2017, the number of active, registered data users was 21,467. Depending on the documentation methods at each RDC, different information regarding users is submitted: the RDC either knows the total number of active users in 2017 or it reports the number of new registrations in 2017. As a makeshift solution, a few RDCs count the number of active data users based on the number of new contracts signed in 2017. The total number of data users reported here should therefore be viewed as a well-founded approximate value.

---

7 The contracts signed by a RDC are not included in this figure because it is possible that multiple contracts were signed with one person.
5 Services for users

Access paths
The only way for researchers to gain access to sensitive data is by entering into data use contracts specifying both the possibilities of data use and the data handling requirements that users must comply with. The research data centres have monitoring systems in place to ensure that the data are only used for scientific research.

Fig. 14: Data access path in 2016 and 2017

What are the access paths through which users may access the data at your RDC?
(Multiple answers possible)

- Data sent by e-mail
- Storage mediums sent by regular mail
- Data download
- Remote data processing
- Guest researcher workstations

Guest researcher work stations are the primary way to access sensitive data.

Secure work stations for on-site data use are provided by 24 RDCs (see Fig. 14). Researchers may use these work stations to access sensitive data (Fig. 1 shows the distribution of guest researcher work stations in Germany). At twelve RDCs, researchers can use data via (controlled) remote data processing, for instance via a remote connection or the transmission of analytical scripts. Such off-site services have seen particularly strong growth at many RDCs in recent years. Twenty-five RDCs now enable researchers to use their data off-site by allowing them to download the datasets, typically as scientific use files (SUF). Seven RDCs send data by e-mail; ten RDCs send storage mediums by regular mail.
**Info box 2: Data access paths and data access formats**

**On-site data access**

**Guest researcher work stations**
Secure work stations at the RDCs for researchers to access sensitive data. Typical features of these guest researcher work stations include that they do not provide uncontrolled internet access and do not allow for saving data locally.

**Remote data processing**
This data access path allows researchers to work with RDC-held data without having to be on-site themselves. Access is enabled via a remote connection or via the transmission of analytical scripts. Researchers write these scripts at their own workstations (possibly with the help of a structure dataset they can use to test the program syntax) and then transmit the analytical scripts to the RDC to have them applied to the original data. Depending on the degree of data sensitivity and/or the legal basis, the output files are checked at the RDC (output control) before they are sent to the researchers. This type of access is known as controlled remote data processing.

**Off-site data access**

**Scientific use files (SUF)**
Research datasets that are more strongly anonymised compared to the data available at the RDC’s secure in-house work stations but still offer strong potential for analysis. Researchers are allowed to analyse these files at their own workplace.

**Campus files (CF)**
These files are designed for university teaching and are even more anonymised than scientific use files.

**Public use files (PUF)**
These research datasets have been anonymised to such a high degree that no use restrictions apply, making it possible to share the data outside the scientific community as well.
Fig. 15: Formats in which data were made available in 2017

In which format do you make data available at your RDC?
(Multiple answers possible)

- Sent by e-mail: 1
- Sent by regular mail: 2
- Download: 3

- Scientific use file (SUF): 4
- Public use file (PUF): 7
- Campus file (CF): 10

(n=28, some RDCs could not provide information on this question)

Most RDCs offer scientific use files for users to download.

Most RDCs make their data available as scientific use files (SUF). Public use files (PUF) are offered by eleven, campus files (CF) by 13 centres (see Fig. 15). For more information on data access paths and data access formats, see the info box on page 23.

Provision of data documentation and tools

Almost all RDCs provide users with informational and documentation materials for each dataset they make available. However, the scope of these materials varies depending on the RDC and the dataset. Metadata for surveys usually include at least the questionnaire and codebooks, sometimes supplemented by method reports, dataset descriptions, quality reports, classifications, technical papers, and variable lists.

RDCs also provide dataset-specific tools such as variable descriptions (e.g. codebooks), programme codes (e.g. do files, syntax files), technical metadata descriptions (e.g. XML files), web applications (e.g. metadata search engines), or theory-oriented documents providing background information on individual questions or variables.

These materials are typically provided online via websites as open access materials available for download without prior registration. If the data are delivered as a package, that package in most cases also includes the available tools.
Fig. 16: Information channels used by the RDCs to communicate their data services in 2017

How and via which channels do you make the scientific community aware of the data available at your RDC? (Multiple answers possible)

<table>
<thead>
<tr>
<th>Channel</th>
<th>Number of RDCs (n=31)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Social media platforms</td>
<td>13</td>
</tr>
<tr>
<td>Flyers/posters</td>
<td>26</td>
</tr>
<tr>
<td>Press releases</td>
<td>10</td>
</tr>
<tr>
<td>Publications</td>
<td>25</td>
</tr>
<tr>
<td>Mailing lists/circular mails</td>
<td>12</td>
</tr>
<tr>
<td>Metadata portals (da</td>
<td>ra, Datacite etc.)</td>
</tr>
<tr>
<td>Organisation of (international) conferences/workshops/trainings</td>
<td>24</td>
</tr>
<tr>
<td>Presentations at (international) conferences/workshops/trainings</td>
<td>27</td>
</tr>
<tr>
<td>RDC newsletter</td>
<td>18</td>
</tr>
<tr>
<td>RDC website</td>
<td>31</td>
</tr>
</tbody>
</table>

In addition to the centres’ own advertising, metadata portals are the main channel for communicating RDC services.

The RDCs use a variety of channels to inform the scientific community about their data services. Figure 16 provides an overview of the information channels used in 2017. Self-advertising, for instance via the RDC’s own website, is the predominant channel, followed by presentations at (international) conferences, trainings and workshops, flyers and posters, and publications on the research data available. Likewise, metadata portals such as da|ra and Datacite play a key role when it comes to informing researchers about data services and expanding the number of users. The widespread use of these portals is a promising basis for raising awareness (and improving the findability) of RDC services among researchers of other disciplines at the national and international levels.
Processing time from application to data transfer

Researchers receive the requested data immediately or within a few hours if the data they are requesting can be retrieved directly through online access (e.g. through simple registration or by downloading the data from the data catalogue, an operation that users can perform themselves). In the case of sensitive, less strongly anonymised data, downloading the data, and sometimes data transfer in general, is impossible. In such cases, the processing time from application to data transfer varies between the RDCs, ranging from one hour to up to eight weeks. Longer processing times must be expected when researchers ask for datasets to be prepared in a special way or when special permissions must be obtained before the data can be released.

Fig. 17: User services in the RDCs in 2017

Which services does your RDC offer for users? *(Multiple answers possible)*

<table>
<thead>
<tr>
<th>Service Description</th>
<th>Number of RDCs in 2017 (n=31)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Provision of tools (e.g. codebooks, variable descriptions, syntax)</td>
<td>28</td>
</tr>
<tr>
<td>Conferences</td>
<td>12</td>
</tr>
<tr>
<td>Workshops/seminars (e.g. summer/winter schools)</td>
<td>22</td>
</tr>
<tr>
<td>Trainings</td>
<td>11</td>
</tr>
<tr>
<td>FAQs</td>
<td>11</td>
</tr>
<tr>
<td>User support via phone, e-mail, or face to face</td>
<td>31</td>
</tr>
</tbody>
</table>

Users have the opportunity to get individual advising services at all RDCs.

All RDCs have contact persons whom external researchers may contact by phone, by e-mail, or in person at the RDC (see Fig. 17). If specialised knowledge is required, user questions are mostly forwarded to in-house specialists. In addition, the RDCs offer advice on the right choice of dataset, on a dataset’s analytical potential, or on data management, including the required measures in terms of data protection and data security.

Nearly all RDCs also offer tools. In the reporting year, this category saw the highest absolute increase in mentions compared to 2016, followed by workshops, trainings, conferences, and frequently asked questions. Some RDCs also offer various training opportunities (courses, workshops, etc.) to enable data users to make more of the analytical potential of specific datasets (see Chapter 6).
Service quality assurance

Most RDCs use process-integrated measures to assure the quality of their services. User feedback is paramount in this regard – sometimes given on users’ own initiative, sometimes collected specifically through user surveys.

Nearly half of all RDCs conduct such user surveys, but only a few do so systematically (e.g. after users have used the services). One RDC asks users for feedback once a year; the others do so less often. The RDCs normally use the results of the user surveys to improve their services. Other quality assurance measures include support from various advisory boards and committees, exchanges with external experts, and informal in-house exchanges among RDC staff.

Internationalisation

The internationalisation of the research landscape is already a daily reality in the RDCs. In response to the needs of the scientific community, research institutions and research infrastructures are becoming more and more international in their outlook. Being guided by and using internationally accepted standards of data description and data provision ensures interoperability and harmonisation across national borders. In particular, the use of persistent identifiers (PID) for datasets and the related standardisation in the description of the key characteristics and contents of research data (metadata) is now common practice at all RDCs. This helps ensure that data can be used and transferred directly to international contexts.

Fig. 18: Support for international scientists in 2016 and 2017

How do you support scientists from foreign countries?
(Multiple answers possible)

Contracts in English 22 24
Data documentation in English 22 23
Access options in English 23 25

Increasingly, RDCs offers their services in German and English.

Twenty-four RDCs help international scientists gain direct access to their data by offering contracts in English; English-language data access options and data documentation are provided by 25 and 23 RDCs, respectively (see Fig. 18). The standard service at some RDCs also includes English versions of data documentation materials that contain descriptions and information not only on technical issues but also on the methods applied.
Close research collaborations underline the international outlook of half of all RDCs.

The fact that staff at most RDCs perform independent research was already pointed out in Chapter 2. These research activities also take place as part of international research collaborations. There are various ways in which individual RDCs are integrated into international consortia. For 2017, 15 RDCs report being part of international research collaborations in which they work on specific issues together with their partners (see Fig. 19). The collaborations reported indicate content-related cooperation and a high level of commitment.

More than two-thirds of all RDCs maintain contacts with research institutions abroad

Ultimately, the best indicator of a centre’s internationalisation status is the extent to which its research data are used by researchers based in other countries. The proportion of international data users varies widely across the RDCs. Twenty-two centres have contacts to research institutions abroad or serve data users from foreign institutions (see Fig. 20), including many German scientists who work and conduct research abroad.

As a consequence, scientists abroad can work with an infrastructure for German data and hence build a better international network. At the same time, data produced in Germany increasingly inform international research.
6 Current trends regarding the research data infrastructure in the social, behavioural, and economic sciences: how RDCs train and retain data users

The research data centres offer training opportunities to enable data users to assess the available data with respect to various questions and to make better use of the data’s analytical potential. In 2017, 21 RDCs conducted a total of 100 training opportunities, including summer/winter schools, workshops, seminars, conferences, and online courses (e.g. MOOCs, webinars). Eight RDCs offered a total of 30 trainings, 17 RDCs a total of 59 workshops or seminars, six RDCs a total of ten conferences, and one RDC an online course (see Tab. 3).

Tab. 3: Number of training opportunities offered by the RDCs in 2017

<p>| Information provided by 21 RDCs (n=21, 10 of the 31 RDCs did not offer any training opportunities in 2017) |
|-------------------------------------------------|-------------------------------------------------|-------------------------------------------------|</p>
<table>
<thead>
<tr>
<th>Trainings (e.g. summer/winter schools)</th>
<th>Number of RDCs offering programmes</th>
<th>Number of programmes</th>
<th>Number of participants</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>30</td>
<td>511</td>
<td></td>
</tr>
<tr>
<td>Workshops/seminars</td>
<td>18</td>
<td>59</td>
<td>799</td>
</tr>
<tr>
<td>Conferences</td>
<td>7</td>
<td>10</td>
<td>414</td>
</tr>
<tr>
<td>Online courses (e.g. MOOCs, webinars, etc.)</td>
<td>1</td>
<td>1</td>
<td>n/a</td>
</tr>
</tbody>
</table>

The RDCs reached more than 1,700 participants through 100 training opportunities.

The attendance figures for the training opportunities offered in 2017 show that data users do take advantage of these opportunities and that there is strong demand for them. Overall, more than 1,700 interested individuals participated in these training programmes in 2017. Workshops and seminars were most popular, drawing a total of 799 participants. The number of participants per workshop or seminar varied between 15 and 130. The trainings also saw strong attendance with a total of 511 participants (range: 10-240 participants). The conferences were attended by 414 data users. Participant numbers for the online course were not collected (see Tab. 3). In-house training for RDC staff and training not organised by the RDCs themselves are not considered in this overview. This clearly suggests that numbers are underreported for the abovementioned categories.
Which topics were covered in 2017?
(Multiple answers possible)

- Statistical methods: 9
- Research data management: 9
- Research ethics: 5
- Data protection: 6
- Data collection/survey methodology: 8
- Research potential/available data: 16

Number of RDCs in 2017 (n=21, 10 of the 31 RDCs did not offer any training opportunities in 2017)

Communicating the data’s research potential is the most important topic of the training opportunities offered by the RDCs.

Sixteen RDCs used their training opportunities to spread information about their data services and the related research potential. Nearly half of all RDCs that implemented training opportunities offered them on statistical methods and research data management (nine RDCs each) and on data collection/survey methodology (eight RDCs). Training opportunities on data protection and research ethics was offered less frequently (six RDCs and five RDCs, respectively) (see Fig. 21). Six RDCs offered additional training opportunities on RDC-specific and more general topics. These programmes are not considered in the figure.
Half of all RDCs implemented their training opportunities independently.

The RDCs implemented training opportunities independently (16 RDCs) and in cooperation with German higher education institutions or non-university research institutions (14 RDCs) (see Fig. 22). Offering training opportunities in cooperation with other German Data Forum (RatSWD)-accredited RDCs or with foreign universities or non-university research institutions is less common, with only four and five RDCs having implemented their programmes through such partnerships, respectively.

Tab. 4: Language of the training opportunities

Information provided by 21 RDCs (n=21, 10 of the 31 RDCs did not offer any training opportunities in 2017)

<table>
<thead>
<tr>
<th>Number of RDCs offering programmes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Some German, some English</td>
</tr>
<tr>
<td>Only German</td>
</tr>
<tr>
<td>Only English</td>
</tr>
<tr>
<td>Other</td>
</tr>
</tbody>
</table>

Most training opportunities were also available in English.

A high share of data access paths, data documentation, and contracts is already available in English (see page 27 – Internationalisation) and hence more accessible for international scientists. This commitment to internationalisation is also reflected in the training opportunities. Five RDCs implemented their training opportunities only in German, whereas 13 RDCs were also prepared to accommodate international participants by offering a mix of English and German (see Tab. 4). Two RDCs even conducted their training opportunities entirely in English.
7 Complaints management

Safeguarding and improving the quality of the work of the RDCs is among the core responsibilities of the German Data Forum (RatSWD). In previous years, it already served as a point of contact for complaints related to the data and services provided by the RDCs. The monitoring commission, aside from conducting the annual monitoring process, of which the present activities report is one outcome, is also charged with handling complaints by data users. To professionalise complaint management and make it more transparent, a complaints office was established at the German Data Forum (RatSWD) business office. It helps ensure swift and professional reactions to complaints and derives from them new ideas for further improving the data infrastructure.

If data users notice major shortcomings in the data services of an accredited RDC, they should first approach the RDC directly and try to find a solution. If the problem cannot be resolved, users may direct their concern to the German Data Forum (RatSWD) complaints office.

The complaints office is only responsible for issues concerning an RDC’s compliance with the accreditation criteria. The German Data Forum (RatSWD) should not be contacted in case of delays in the normal procedures or with respect to staff conduct at the RDCs. Complaints of this kind should be addressed directly to the RDC in question.

For detailed information about the procedures, see German Data Forum (RatSWD) Output 8 (5) or the updated version of that output on the German Data Forum (RatSWD) website.9

Current complaints procedures in 2017

Towards the end of the reporting period, the German Data Forum (RatSWD) received a complaint, which was classified as relevant by the monitoring commission. The RDC to which the complaint was addressed subsequently issued a statement, which served as the basis for a counterstatement submitted to the monitoring commission by the complainant. The complaints procedure was carried out and completed in 2018. The commission did not find any violation of the mandatory criteria on the part of the RDC to which the complaint was addressed.

---


9 https://www.ratswd.de/en/info/complaints-office
Appendix
## Appendix A

Development of the German Data Forum’s (RatSWD) research data infrastructure

<table>
<thead>
<tr>
<th>Year</th>
<th>Event</th>
</tr>
</thead>
<tbody>
<tr>
<td>1999</td>
<td>The Commission to Improve the Informational Infrastructure between Research and Official Statistics (KVI) recommends the establishment of research data centres (RDCs).</td>
</tr>
<tr>
<td>2001</td>
<td>The German Data Forum (RatSWD) Founding Committee is set up.</td>
</tr>
<tr>
<td></td>
<td>The following RDC is founded:</td>
</tr>
<tr>
<td></td>
<td><strong>Research Data Centre of the Federal Statistical Office (FDZ-Bund)</strong>&lt;sup&gt;*&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td>Germany-wide access to official statistics microdata from the following fields: population, education, health, business, agriculture, environment, administration of justice, finance, and taxes.</td>
</tr>
<tr>
<td></td>
<td><a href="http://www.forschungsdatenzentrum.de/en">www.forschungsdatenzentrum.de/en</a></td>
</tr>
<tr>
<td>2002</td>
<td>The following RDC is founded:</td>
</tr>
<tr>
<td></td>
<td><strong>Research Data Centre of the Statistical Offices of the Länder (FDZ-Länder)</strong>&lt;sup&gt;*&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td>Germany-wide access to official statistics microdata from the following fields: population, education, health, business, agriculture, environment, administration of justice, finance, and taxes.</td>
</tr>
<tr>
<td></td>
<td><a href="http://www.forschungsdatenzentrum.de/en">www.forschungsdatenzentrum.de/en</a></td>
</tr>
<tr>
<td>2003</td>
<td>The following RDCs are founded:</td>
</tr>
<tr>
<td></td>
<td><strong>Research Data Centre German Microdata Lab at GESIS (RDC GML)</strong>&lt;sup&gt;*&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td><strong>International Data Service Centre at the Institute for the Study of Labour (FDZ IZA, IDSC)</strong>&lt;sup&gt;*&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td>National and international labour market datasets with standardised information (eddi-conferences.eu). Research with, methods and resources for using online data for labor economics and social science. Development of tools and methods for remote access (statsdirect.org) and remote processing (JoSuA).</td>
</tr>
<tr>
<td></td>
<td><a href="http://www.iza.org/organization/idsc">www.iza.org/organization/idsc</a></td>
</tr>
<tr>
<td>2004</td>
<td>The German Data Forum (RatSWD) is founded.</td>
</tr>
<tr>
<td></td>
<td>The following RDCs are founded:</td>
</tr>
<tr>
<td></td>
<td><strong>Research Data Centre of the German Federal Employment Agency at the Institute for Employment Research (FDZ BA at IAB)</strong>&lt;sup&gt;*&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td>Data on persons, households, and employers, as well as combined datasets consisting of survey data and administrative research data in the fields of social security and labour market, and employment research.</td>
</tr>
<tr>
<td></td>
<td><a href="https://fdz.iab.de/en.aspx">https://fdz.iab.de/en.aspx</a></td>
</tr>
</tbody>
</table>
Research Data Centre of the German Pension Insurance (FDZ-RV)*
Data on the insurance accounts of individuals insured in the Federal Pension Insurance. The accounts contain data on the insured persons’ insurance history and the pension and rehabilitation benefits they received.
www.fdz-rv.de/FdzPortalWeb/dispcontent.do?id=main_fdz_english

2008
The following RDCs are accredited:

Research Data Centre of the Federal Institute for Vocational Education and Training (BIBB-FDZ)
Firm-level and individual-level datasets of vocational education research dealing primarily with the attainment and use of vocational knowledge and skills.
www.bibb.de/en/53.php

Research Data Centre of the Institute for Educational Quality Improvement (FDZ IQB)
German datasets from the major national and international school performance studies and national studies measuring educational standards.
www.iqb.hu-berlin.de/fdz

2009
Establishment of the Standing Committee Research Data Infrastructure (RDC Committee) of the RatSWD

2009
The following RDCs are accredited:

Research Data Center of the Socio-Economic Panel Study at DIW Berlin (RDC SOEP)
Data from representative annual surveys of private households. The SOEP CORE sample features topics such as income, employment, education, and health. In addition, there is the longitudinal innovative sample (SOEP-IS), which enables external researchers to contribute research projects of their own.
www.diw.de/en/diw_02.c.222518.en

Research Data Centre of the Survey of Health, Ageing and Retirement in Europe (FDZ SHARE)
Data from the multidisciplinary and cross-national panel study “Survey of Health, Ageing and Retirement in Europe” (SHARE), which produces microdata on health, socio-economic conditions, and social and family networks of approximately 123,000 individuals aged 50 or older in more than 20 European countries and Israel. The seventh wave of SHARE was collected in 2017.
www.share-project.org

* The research data centres Federal Statistical Office, Statistical Offices of the Länder, GML, IZA, BA at IAB, and RV were established prior to the foundation of the German Data Forum (RatSWD) and became part of the research data infrastructure in 2004. In these cases, the year of the RDCs’ foundation is listed. All other RDCs were accredited after 2004 by the German Data Forum (RatSWD). With these RDCs, the year provided is the year of their accreditation.
Research Data Centre International Survey Programmes at GESIS (RDC International Survey Programmes)

Internationally comparative survey data from more than 70 countries on nearly all social science topics: Comparative Study of Electoral Systems (CSES), European Values Study (EVS), Eurobarometer, European Election Studies (EES), International Social Survey Programme (ISSP).


Research Data Centre Elections at GESIS (RDC Elections)

Access to German national election surveys (federal elections and state elections), Politbarometer, Forsa-Bus, ARD Deutschlandtrend and Surveys for the Federal Government. The RDC’s largest project at this point is the German Longitudinal Election Study (GLES).

www.gesis.org/en/elections-home/elections-home

Research Data Centre ALLBUS at GESIS (RDC ALLBUS)

Data from the Allgemeine Bevölkerungsumfrage der Sozialwissenschaften (ALLBUS) and German General Social Survey (GGSS) in English, on the attitudes, behaviours, and social structure of the German population.

www.gesis.org/en/allbus/allbus-home

The following RDCs are accredited:

- Research Data Centre for Business and Organizational Data (FDZ-BO)
  Quantitative and qualitative business, organizational data, linked employer and employee data, and data from employee and member surveys.
  www.uni-bielefeld.de/(en)/soz/fdzbo

- Research Data Centre of the German Centre of Gerontology (FDZ-DZA)
  Data from the long-term German Ageing Survey (DEAS) on the changing life situations and ageing processes of people in mid- and older adulthood, and from the German Survey on Volunteering (FWS), a representative survey programme with a focus on voluntary activities and civic participation in Germany.
  www.dza.de/en/fdz.html

- Research Data Centre PsychData of the Leibniz Institute for Psychology Information (FDZ PsychData at ZPID)
  Pooled quantitative datasets from both basic research and applied psychology; data archiving with a focus on longitudinal studies, large-scale survey studies, and development testing.
  www.psychdata.de/index.php?main=none&sub=none&lang=eng

- Research Data Centre of the German Family Panel (FDZ pairfam)
  Datasets from the “Panel Analysis of Intimate Relationships and Family Dynamics” (pairfam), a representative, interdisciplinary longitudinal study for the analysis of private living arrangements in Germany.
  www.pairfam.de/en
Research Data Centre Ruhr at the RWI – Leibniz Institut for Economic Research (FDZ Ruhr at RWI)

Specialisation on regional data: socioeconomic data measured by 1 square km grids. Aside from geo-referencing data on a scientific basis, the RDC provides various individual-level and employer-level data collected in RWI research projects.

http://fdz.rwi-essen.de

The following RDCs are accredited:

2011

LMU-Ifö Economics & Business Data Center (EBDC)

Datasets of German companies, including survey data collected by the ifö Institute on firms' business status, innovativeness, and investment behaviour, as well as external data on corporate financing and governance structure. Merged panels of the aforementioned two data sources are also available.

www.cesifo-group.de/ifohome/facts/EBDC.html

Research Data Centre of the Robert Koch Institute (RDC RKI)

Data on the state of health and health-related behaviour of Germany's resident population, collected on the basis of nationally representative studies.

www.rki.de/puf

Research Data Centre of the Federal Centre for Health Education (FDZ BZgA)

Data from nationally representative surveys, repeated at regular intervals, measuring the population's susceptibility to health education and prevention campaigns, as well as the knowledge, attitudes, and behaviour in the general population concerning the health issues addressed by BZgA.

www.bzga.de/home

2012

The following RDCs are accredited:

Research Data Center Wissenschaftsstatistik of the Stifterverband (RDC Wissenschaftsstatistik)

Data on the research and development activities of German companies, on the financial volume, structure, and regional distribution of research and development activities (R&D), and on R&D staff in the business sector.

www.stifterverband.org/research_data_center

Research Data Centre Education (FDZ Bildung) at the German Institute for International Educational Research (DIPF)

The hosted datasets include approaches of qualitative educational research such as video data, transcriptions, contextual materials and survey tools of quantitative educational research such as questionnaires and assessment tests. The collected datasets refer to the quality of instruction and to the quality of schools.

www.fdz-bildung.de
Research Data Center of the Leibniz Institute for Educational Trajectories at the University of Bamberg (RDC-LiFBI)

Longitudinal data from the National Educational Panel Study (NEPS), which was launched in 2010 with more than 60,000 panel participants in six starting cohorts to study skills formation, educational processes, educational decisions, and educational returns in formal, non-formal, and informal contexts across the lifespan.
www.lifbi.de/en-us/home.aspx

ZEW Research Data Centre for European Economic Research (ZEW-FDZ)

The ZEW-FDZ provides microdata from ZEW firm surveys on innovation activities, the development of young firms, the use of information and communication technologies, and further topics. Data from individual and expert surveys are also accessible – for example, the ZEW Financial Market Survey.
https://kooperationen.zew.de/en/zew-fdz

The following RDCs are accredited:

Research Data Centre of the German Youth Institute (FDZ-DJI)

Data from the surveys on children and young people growing up and the life situations of adults and families, conducted in regular intervals since 1988.
www.dji.de/en/the-dji.html

FDZ SFB 882

Qualitative and quantitative datasets from inequality research
(The RDC was discontinued in 2016. Depending on the data type and basis, the data of RDC SFB 882 were transferred to different organisations: IAB data were handed over to FDZ BA im IAB; qualitative data with an organisational connection were handed over to FDZ BO; the remaining data were handed over to the SOBI archive at the University of Bielefeld (currently under development). (Last update on 09/14/2017)
https://sfb882.uni-bielefeld.de/en/fdz-sfb882.html

Research Data Center Archive for Spoken German at the Institute for the German Language (FDZ AGD)

Data on spoken German in interactions (conversation corpora) and data on domestic and non-domestic varieties of German (variation corpora).
http://agd.idos-mannheim.de

Research Data Center Programme for the International Assessment of Adult Competencies (PIAAC) at GESIS (RDC PIAAC)

German and international data of the Programme for the Assessment of Adult Competencies (PIAAC). For Germany, additional regional data and longitudinal data are available.
www.gesis.org/en/piaac/rdc
### 2015

The following RDCs are accredited:

<table>
<thead>
<tr>
<th>RDC Name</th>
<th>Website</th>
</tr>
</thead>
<tbody>
<tr>
<td>Deutsche Bundesbank Research Data and Service Centre</td>
<td><a href="http://www.bundesbank.de/Navigation/EN/Bundesbank/Research/RDSC/rdsc.html">www.bundesbank.de/Navigation/EN/Bundesbank/Research/RDSC/rdsc.html</a></td>
</tr>
</tbody>
</table>

### 2017

The following RDC is accredited:

<table>
<thead>
<tr>
<th>RDC Name</th>
<th>Website</th>
</tr>
</thead>
</table>

### 2018

Guest status is granted to the following RDCs (Last update: August 2018):

<table>
<thead>
<tr>
<th>RDC Name</th>
<th>Website</th>
</tr>
</thead>
<tbody>
<tr>
<td>Research Data Centre eLabour (FDZ eLabour)</td>
<td><a href="http://elabour.de">http://elabour.de</a></td>
</tr>
<tr>
<td>Research Data Centre at the Federal Motor Transit Authority (FDZ im KBA)</td>
<td><a href="http://www.kba.de/EN/Home/home_node.html">www.kba.de/EN/Home/home_node.html</a></td>
</tr>
</tbody>
</table>
Appendix B

The monitoring commission
For quality assurance purposes, the German Data Forum (RatSWD) agreed to establish a monitoring commission that began its work in July 2016. Its main task is to collect and assess the regular RDC reports. Moreover, the commission monitors compliance with the obligations arising from provisional first-time accreditation. The RDC Committee elects the commission from its own membership for a three-year term concurrent with the German Data Forum (RatSWD) appointment period. Thus, the commission enjoys a special level of trust and legitimacy. It consists of four members of the RDC Committee and two deputy members (to replace elected members, if required) and the German Data Forum (RatSWD) chairpersons as guests.

Members of the monitoring commission

Maurice Brandt
Research Data Centre of the Federal Statistical Office (FDZ-Bund)

Dr. Lea Eilers (Chair)
Research Data Centre Ruhr at the RWI – Leibniz Institute for Economic Research (FDZ Ruhr at RWI)

Holger Quellenberg
Research Data Centre of the German Youth Institute (FDZ-DJI)

Dr. Karsten Stephan
Research Data Centre for Higher Education Research and Science Studies (fdz.DZHW)

Deputy members of the monitoring commission

Tatjana Mika
Research Data Centre of the German Pension Insurance (FDZ-RV)

Standing guests of the monitoring commission

Prof. Regina T. Riphahn, Ph.D.
Chair - German Data Forum (RatSWD)

Prof. Stefan Bender
Vice Chair - German Data Forum (RatSWD)
Appendix C

Contributors

Members of the monitoring commission

Maurice Brandt
Research Data Centre of the Federal Statistical Office (FDZ-Bund)

Dr. Lea Eilers (Vorsitz)
Research Data Centre Ruhr at the RWI – Leibniz Institute for Economic Research (FDZ Ruhr at RWI)

Holger Quellenberg
Research Data Centre of the German Youth Institute (FDZ-DJI)

Dr. Karsten Stephan
Research Data Centre for Higher Education Research and Science Studies (fdz.DZHW)

German Data Forum (RatSWD) Business Office

Marie Bormann

Mathias Bug
This report is the result of a project, which is funded by the Federal Ministry for Education and Research (reference number: 01UW1402). Unless otherwise stated, the responsibility for this publication lies with the German Data Forum (RatSWD).

doi: 10.17620/02671.36

Citation suggestion:
Established in 2004, the **German Data Forum** (Rat für Sozial- und Wirtschaftsdaten, RatSWD) is an independent council. It advises the German federal government and the federal states (Länder) in matters concerning the research data infrastructure for the empirical social, behavioural, and economic sciences. The German Data Forum (RatSWD) has 16 members. Membership consists of eight elected representatives of the social, behavioural, and economic sciences and eight appointed representatives of Germany’s most important data producers.

The German Data Forum (RatSWD) offers a forum for dialogue between researchers and data producers, who jointly issue recommendations and position papers. The council furthers the development of a research infrastructure that provides researchers with flexible and secure access to a broad range of data. The German Data Forum (RatSWD) has accredited 31 research data centres and fosters their interaction and collaboration.