

# 6

## STATE-OF-THE-ART DATA INFRASTRUCTURE FOR BETTER INFORMED DECISIONS

### I. Motivation: better data for better decisions

### II. Starting position: need for improvement in data collection, linkage and access

1. Incomplete collection of data
2. Time delay in the provision of data
3. Availability of existing data restricted
4. Data linkage at the best on a case-by-case basis
5. Data access not user-friendly

### III. Options for action: a data infrastructure for the 21st century

1. Improve the regulatory framework of the FSO/LSOs system
2. Close data gaps
3. Less time delay and greater adaptability
4. Improve the availability of data already collected
5. Better linkability across different data providers
6. Simple and user-friendly data access

### Appendix

### References

This is a translated version of the original German-language chapter "Zeitgemäße Dateninfrastruktur für fundierte Entscheidungen", which is the sole authoritative text. Please cite the original German-language chapter if any reference is made to this text.

## KEY MESSAGES

- By international comparison, Germany's research data infrastructure is outdated. Data are lacking, data linkage is hardly possible, and access is not user-friendly.
- Legislation on statistics should regulate the information requirements that official statistics are expected to fulfil and give them freedom in their concrete surveys to make them more flexible and responsive. The resources of official statistics must be expanded.
- The Research Data Act should be introduced quickly. Research should be given a lot of weight in the trade-off with data protection.

## SUMMARY

**Data-based analyses are indispensable for well-informed decisions in politics and public administration.** Evidence-based empirical research and policy advice depend on reliable data to inform policy-makers, public administration and the public about causal relationships, assess the consequences of planned measures, and evaluate the effectiveness and efficiency of implemented measures. This requires a comprehensive, easily accessible data infrastructure. It must be adaptable in order to react quickly to new information requirements. Investments that enhance the potential of official statistics to provide such a data infrastructure is an investment in better informed, more targeted and more cost-effective policy decisions.

**By international comparison, the data infrastructure in Germany is outdated.** There are deficits in the scope of available data and its timely availability. Some data are not collected at all. Furthermore, existing administrative **data** from the administration are **not accessible** for researchers and policy advisers. At the same time, **existing data from** different data producers can only be **linked in a very limited form** to new data products that enable new insights. Even in cases where data are in principle available for research, **access** to them in Germany, especially to official statistics, is **technically outdated** and not very user-friendly.

In order to improve the research data infrastructure, progress needs to be made urgently in all of the above-mentioned areas of action. As a first step, the Federal Statistics Act, as the legal basis for the network consisting of the Federal Statistical Office and the Statistical Offices of the Länder (FSO/LSOs system), should be further developed with the aim of making the **legislation on statistics truly output-oriented**. To this end, legislators should in future assess what information requirements exist and what services the FSO/LSOs system should provide. However, the latter should retain the legal freedom to itself define the surveys that are needed for this. Furthermore, additional data should be collected, for example in the fields of education and wealth, and a regular, high-frequency household survey should be established.

A **more extensive use of administrative process data** could expand the available range of official statistics in a cost-effective way. Furthermore, the **Research Data Act** should **improve data access, expand the possibilities for linking** data from different public data producers for research, and **assign research a sufficient weight in the trade-off with data protection**. In order to make access to data more user-friendly, remote access to data should be established as a standard access route for research data, and the publicly accessible data offered by the FSO/LSOs system should be brought up to the latest technical standards.

# I. MOTIVATION: BETTER DATA FOR BETTER DECISIONS

530. **Data-based analyses are an indispensable foundation for good decisions in politics and public administration.** The evidence-based design of policy measures and regulatory frameworks is not possible without a reliable data basis. If such a basis is lacking, measures may be adopted that either do not achieve the intended goals at all (ineffective measures), or only at too high a cost (inefficient measures), or are accompanied by unintended side effects.

Data are not only necessary **ex ante as a basis for evidence-based policy design** but also **ex post for evidence-based evaluation**. Evaluation results can be used both to improve measures and to design them in a more targeted and cost-efficient way in the future. In contrast, a lack of data can give information advantages to insiders and lobby organisations, which can then more easily influence policy and public administration when lack of or insufficient data creates a situation of uncertainty.

Data can be used directly by policy-makers and regulators, especially in aggregated form. Very often, however, decision-makers rely on evaluation and interpretation of data by independent research and evidence-based policy advice. **Providing the necessary research-data infrastructure** thus represents an **investment in better-informed, better-targeted and more cost-effective policy decisions**.

531. The **data infrastructure in Germany is lagging behind** other countries. This was **particularly evident during the COVID-19 pandemic and the energy crisis**. ↘ [ITEM 540](#) During the COVID-19 pandemic, for example, there was a lack of up-to-date **information on** the potential or actual impact of the pandemic on **companies** (GCEE Annual Report 2021 boxes 11 and 15). Had such information been available, COVID-19 assistance could have been better targeted, evaluated almost in real time, and improved as needed. In France, by contrast, the implementation of aid to businesses was directly accompanied by a decision to monitor and evaluate the assistance. For this purpose, data on the disbursement of assistance and the usage of short-time working measures were collected, linked with company micro-data from the national statistics office, and evaluated by an independent commission. In Germany, the data infrastructure for evaluating business assistance is only now being created, more than three years after the pandemic began (Profit, 2023).

Furthermore, there was no sufficiently large **high-frequency household survey** in Germany that could be adapted at short notice to measure how individual households were affected by the pandemic. In the United States, by contrast, the U.S. Census Bureau quickly set up the Household Pulse Survey and the Small Business Pulse Survey in 2020 to measure the pandemic's economic impact on households and small companies. In addition, the Current Population Survey (CPS) was rapidly adapted to assess the impact of the pandemic, for example on

remote work or COVID-19-related business closures and job losses (U.S. Bureau of Labor Statistics, 2020). In Norway, real-time register data were used to measure the impact on the economic situation of households within a few days of the start of the COVID-19 pandemic (Alstadsæter et al., 2020). They were then used to inform the public and as a basis for government’s decision-making.

In the **energy crisis**, a lack of or **insufficiently detailed, up-to-date data on the energy consumption and costs** of private households made it difficult to promptly assess the financial impact of the energy crisis. In this case, too, a **high-frequency household survey** could have provided information on the different ways in which various income groups and household types were affected and could have improved the design of support measures. Furthermore, there was a lack of data on the supply side of the energy market. Even now, unlike in parts of the United States for example (Palmer et al., 2022), no data are collected on which types of power plants set the price in the electricity market – and when. As a result, it has not been possible to directly estimate the extent to which the various electricity-producing companies have benefited from the crisis.

**532. Numerous legislative projects at national and European level** show that these problems have been recognised at the political level. Examples include the Register Modernisation Act, the further development of the national data strategy, which is currently under discussion, the Research Data Act, which is in the consultation process, and the European Data Governance Act. These and other initiatives have the potential to improve the data situation for research and thus the decision-making basis for policy-makers. **Progress has been slow, however.** Whether or how much these legislative projects will improve the data basis will depend decisively on how the hitherto rather crude strategic objectives, e.g., of the data strategy, are implemented in law in practice. If strengthening the regulatory framework really is to lead to the expected improvement in the research data infrastructure, practical implementation in the public administration and support at the management level will also be crucial.

**533.** The data infrastructure in Germany is largely determined by the **general regulatory framework of the Federal Statistical Office and the statistical offices of the Länder** [↪ BACKGROUND INFO 22](#). This framework can be improved in particular by shifting statistics legislation away from defining the input of statistics, i.e., the concrete surveys, towards the outputs, i.e., defining the information needs. This would make official statistics much more flexible and responsive.

[↪ ITEM 553](#)

In addition, there are **five key areas of action for improving the data infrastructure**. First, data gaps in Germany should be closed by **gathering new data**. [↪ ITEM 557](#) Second, the time lag in the provision of data should be shortened and the reaction speed of the statistics increased. [↪ ITEM 562](#) Third, the **availability** of already collected data should be improved. [↪ ITEM 564](#) Fourth, the **possibilities of linking data** should be expanded, especially between data sets from different data producers. [↪ ITEM 568](#) Fifth, **data access** should be made much simpler, more modern and more user-friendly. [↪ ITEM 572](#)



➤ BACKGROUND INFO 22

**Background: The Federal Statistical Office and the statistical offices of the Länder – regulatory framework**

The national system of public statistics consists of the network of the Federal Statistical Office and the statistical offices of the Länder (FSO/LSOs system), the other national data producers (Other National Authorities, ONAs) such as the Deutsche Bundesbank, which participate in the production of European statistics (Decker and Vorgrimler, 2021), and other public institutions that produce statistics outside of the European statistical system. The FSO/LSOs system, its tasks and the legal framework for its work are **laid down in the Federal Statistics Act (BStatG)**. The system represents the **core of the national statistical system** and assumes a coordinating role both nationally and internationally in the collection, processing and transmission of data. It **produces federal statistics** that are mandated by law (section 5 of BStatG). In addition, the Federal Statistical Office can be given the task of preparing data from administration procedures (process data) and processing them into **business statistics** (section 8 of BStatG).

## II. STARTING POSITION: NEED FOR IMPROVEMENT IN DATA COLLECTION, LINKAGE AND ACCESS

534. In Germany, the supply of and access to research data have **improved significantly since the turn of the millennium**. Before that, researchers had mainly aggregated key figures at their disposal. The German Data Forum (RatSWD), which has played a key role in the development of the data infrastructure in these areas since its founding, has particularly contributed to the improvement. ➤ [BACKGROUND INFO 23](#) The number of institutions accredited by the RatSWD through which data producers make their data available to research (research data centres, FDZs) has now risen to 42 in 2023. These **FDZs** cover data in the fields of social and economic sciences, education, health and psychology in particular, but also other data such as those of the Federal Motor Transport Authority. Their task is **to process and document data and make them available to** researchers for specific purposes, confidentially, promptly and in a quality-assured manner, **while ensuring data protection**. They are to provide several access routes for using the data, ideally including remote data access. The work of the RatSWD is **part of the National Research Data Infrastructure**, which systematically makes data accessible, networked and usable for the entire science system.
535. **By international comparison**, however, the German research data infrastructure is still **lagging behind**. In the Scandinavian countries, for example, comprehensive register data, both on individuals and companies, are available and can be linked on request. These data cover the labour market, education and healthcare, and the income and wealth situation (Ludvigsson et al., 2019). Furthermore, the data can be used via remote access. In Austria, similar possibilities

are offered by the Austrian Micro Data Center (AMDC) and in France by the Centre d'accès sécurisé aux données (CASD; Gottschalk et al., 2023). [↪ BACKGROUND INFO 24](#)

536. Despite the progress that has been made over the past 20 years, **there is therefore still a considerable need for improvement**. In the light of the RatSWD's guiding principles, [↪ BACKGROUND INFO 23](#) which are shared by the GCEE, there are five key problems and thus **five areas for action**. First, some important data are not collected (**data collection**). Second, there is a significant time delay lag between when the data are collected and made available, and surveys can often not be quickly adapted to new circumstances, especially in times of crisis (**time lag**). Third, not all of the collected data is made available for research (**availability**). Fourth, the possibilities of linking different data are often limited, especially if the data sets are created by different producers (**linkability**). Fifth, technical and practical access to the available data is more difficult than it should be (**data access**).



[↪ BACKGROUND INFO 23](#)

#### **Background: The German Data Forum (RatSWD) and its guiding principles**

The research data infrastructure in Germany that is relevant for economic and social policy is largely coordinated by the RatSWD. The Forum was set up in 2004 as an independent advisory board to the Federal Government on the recommendation of the Commission for the Improvement of the Informational Infrastructure between Science and Statistics. **It is conceived as a joint representation of the interests of data producers**, such as the research data centres of the statistical offices of the FSO/LSOs, the Deutsche Bundesbank and the Socio-Economic Panel (SOEP), as well as data users. It thus plays a key role in the further development of the research data infrastructure in the social sciences. [↪ CHART 143](#) In particular, the Forum assumes the task of accrediting and evaluating the research data centres in Germany according to uniform criteria and minimum standards. The RatSWD's **guiding principle** for modern and sustainable research data centres is the concept of **open data for research**, i.e. unrestricted, free of charge access for research purposes to administrative, economic and scientific data, readable by humans and machines, with metadata and a transparent licence. Collected **data should be findable, easily accessible, automatically linkable and reusable**.



↳ CHART 143

Locations of the accredited research data centres<sup>1</sup> in 2023



Accreditation in the period: ■ before 2009 ■ 2009 - 2013 ■ 2014 - 2018 ■ 2019 - 2023

1 – For exact designations of the research data centres, see Table 25.

Sources: Federal Agency for Cartography and Geodesy, RatSWD, own presentation  
© Sachverständigenrat | 23-315-01

## 1. Incomplete collection of data

537. In many areas, research is restricted because **data** is either **not collected at all** or not of sufficient quality. **For example**, there is no nationwide **register for schoolchildren or an educational progress register** that would make data on grades from intermediate and final examinations, school entrance examinations and the like accessible for scientific purposes and evaluable on a longitudinal basis (RatSWD, 2022a; GCEE Annual Report 2021 items 376 ff.). This is partly because many Länder do not collect these data at all, and partly because the data that are available in the Länder are not pooled (Hertweck et al., 2023), which makes a meaningful evaluation of education policy impossible.
538. Information on wealth distribution is essential for redistributive policy. However, no reliable **data on household wealth** has been collected since the suspension of the wealth tax in 1997. Although the Federal Statistical Office (Destatis) provides survey-based data on wealth in Germany via the sample survey on income and consumption, the German Institute for Economic Research (DIW) via the SOEP, and the Deutsche Bundesbank via the survey of Private Households and their Finances (PHF), these surveys **have some deficits that have long been known** (Löffler et al., 2015). The voluntary nature of participation in the surveys is likely to lead to selection effects that result in a bias between the measured and actual distribution of wealth – especially due to an insufficient inclusion of high-wealth households in the samples. Self-assessments can deviate significantly from the actual value of wealth, especially in the case of low-liquid assets. In addition, the small sample size makes a differentiated view of the wealth situation according to socio-economic characteristics difficult.
539. For a targeted distribution policy, but also for climate and energy policy, **data on the building stock** is currently very relevant. Data on real estate are kept up to date by the public administration, e.g., in the land survey register or land registers. However, the data on buildings and housing units they contain are not yet recorded completely and uniformly nationwide. In addition, there is no uniform identifier in the form of a building and apartment number, like e.g., in Norway, to link the different data (Krause et al., 2022). The poor data availability in this area became particularly salient during the energy crisis, when the deficient data on the energy efficiency of buildings confined the possibilities for developing and implementing targeted support measures.
540. In addition to data that should be collected continuously, there also exist **relevant event-related one-off data** that are not collected systematically. Yet they can be necessary for **evaluating discretionary policy measures**. An example of this is the evaluation of the COVID-19 assistance packages. In France, an evaluation of the support measures, including the necessary data collection, was decided on immediately after their implementation in 2020. The first evaluation already took place with the final report in summer 2021 (GCEE Annual Report 2021 box 25). In Germany, only now, more than three years after the COVID-19 pandemic began, are data being linked and a corresponding data infrastructure being built up for the evaluation of the corporate support measures.



## 2. Time delay in the provision of data

541. If no event-related data are collected, a timely evaluation of policy measures could still be conducted relying on the standard, regularly collected data. However, **officially collected microdata** are only **available for research purposes with a considerable delay**, if at all. This is due to both long statutory reporting periods and a lack of resources for data preparation. The delay makes **timely, up-to-date evaluations almost impossible**. The impact of the COVID-19 crisis or the energy crisis on companies, as well as the effect of the support measures could have been examined, for example, with the help of the Official Company Data for Germany (AFiD). However, these are only made available with a delay of two to three years; depending on the data module, the most recent year available is currently 2021, 2020 or 2019. Similarly, in the case of the microcensus – the most comprehensive household survey in Germany – and the wage- and income-tax statistics, the most up-to-date, available data stem from the reporting years 2020 and 2017 respectively.
542. **Macroeconomic indicators** – such as the national accounts or the indicators of short-term economic statistics – are **also only available after a significant delay**. Against this background, the efforts being made by Destatis to accelerate the flash estimate of the gross domestic product (GDP) using scanner, satellite, mobility, financial-transaction, and transport data and to improve the estimate quality are to be welcomed (Dickopf et al., 2019). Nevertheless, there is a need for additional high-frequency and timely business-cycle indicators. Although these indicators would inevitably be of a slightly lower quality than the final data products of official statistics, they would paint a sufficiently accurate picture of the current situation much more promptly, and thus represent a very helpful source of information, especially in volatile times of crisis.

## 3. Availability of existing data restricted

543. Another problem is that **data that has already been collected is sometimes not made available for research at all or only to a limited extent**. For this reason, comparisons of educational results between Länder, for example, are almost impossible. The extended sample carried out between 2000 and 2006 in the course of the PISA survey, which made Länder comparisons possible, was discontinued in the 2009 survey year. Since then, the PISA data set has not contained any Länder variables, as Länder comparisons were actually supposed to be carried out with the newly introduced data from the Institute for Quality Development in Education (IQB). The IQB's high-quality data do, in principle, contain identifiers for the individual Länder. However, their use is restricted by law; it is subject to a review process, and comparisons of individual Länder are explicitly prohibited.
544. Even more problematic is the **impeded access to official process or register data**. On the one hand, registers are still lacking in many areas. For example, the basic business register is still under construction and an administrative register with data on buildings and apartments simply does not exist. A register-based census is planned for the first time for 2031. On the other hand, **many of the**

existing **register data are not yet intended for research use, or only to a limited extent** (RatSWD, 2022a, 2023a). Such registers include, for example, the Trade Register, the Residents Register and the Central Register of Foreigners.

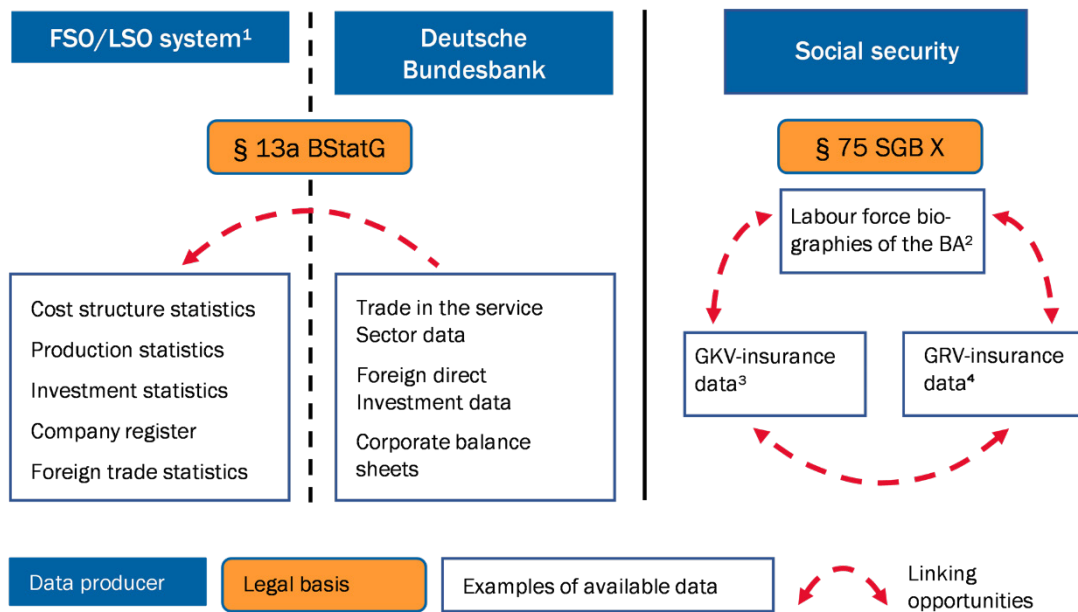
545. Often, **only samples** are available for register data but not the complete data sets. This is the case, for example, with the Central Register of Foreigners of the FDZ Federal Office for Migration and Refugees (BAMF) or with the sample of integrated labour-market biographies (SIAB) at the FDZ of the Institute for Employment Research (IAB) of the Federal Employment Agency. The reason for this restriction is often data-protection concerns but sometimes also limitations of the IT infrastructure. Restriction to samples can prevent certain analyses, for example studies for small subgroups or estimates based on networks between individual observations, for which the largest possible connected network is needed (Bonhomme et al., 2023). In some cases, the complete data sets are only made available to a small number of researchers, which distorts the level playing field in the scientific process.
546. Deletion regulations, e.g., **according to section 13a of BStatG**, are also a key problem and should be abolished. [↪ ITEM 567](#) Section 13a of BStatG currently stipulates that identifiers of companies must be deleted after 30 years. Depending on how the law is interpreted, this can be seen as an obligation to delete all data sets created on the basis of the links after 30 years. Studies based on longer time series, or the application of new methods to older data that promises new insights for the present, are thus made impossible. Such studies are also made more difficult if analogue data from the past have not been digitised and made accessible for research. This is the case, for example, with the data of the Central Register of Foreigners (Altmann et al., 2023).

## 4. Data linkage at best on a case-by-case basis

547. For many economic policy issues, it is necessary to link data from several sources because data collection and storage in Germany are very decentralised. [↪ BACKGROUND INFO 22](#) For example, to examine how financing restrictions affect the investment behaviour of German companies and thus influence wages and productivity, one would have to link three data sources: Most labour market data are collected by the Federal Employment Agency and the IAB on a register basis. Destatis collects production and investment data and the Deutsche Bundesbank collects balance sheet data from larger companies. However, the possibilities for **linking are very limited, as links between different data producers are only possible if they are explicitly permitted by law**. [↪ CHART 144](#) Up to now, however, such permission is lacking in most cases. Accordingly, no data from official statistics may be linked with data from scientific business surveys in Germany (Gottschalk et al., 2023). Moreover, in the case of personal data, the previously obtained informed consent of the data subjects is usually required. Linkages between data from official statistics and social-security funds are therefore also practically impossible (Altmann et al., 2023). The lack of permission to link data also represents a significant obstacle to research in education and health (Altmann et al., 2023; Hertweck et al., 2023).

↘ CHART 144

**Current options for linking important data sources in the public statistics system**



1 – Federal Statistical Office and the statistical offices of the Länder. 2 – Federal Employment Agency. 3 – National health insurance. 4 – Statutory pension scheme.

Source: own depiction

© Sachverständigenrat | 23-468-01

In addition to legal hurdles, which were only partially removed by the reform of the BStatG in 2005 (Brandt et al., 2007; RatSWD, 2023a), there are **institutional and organisational challenges** in data linkage. For example, there is a **lack of uniform identifiers for linking** data sets between the different data producers. Furthermore, the rules for data use are not uniform. As a result, data sets are currently mainly constructed and used as separate silos.

548. **In the past, there have been efforts to explore** a standardised **linking** of different data sets. These included, among others, the project 'Combined Company Data for Germany', which was carried out by the statistical offices of the Federation and the Länder, the IAB, the Federal Employment Agency (BA) and the Deutsche Bundesbank; the project was concluded in 2012 (Bender et al., 2007; Gürke et al., 2011). **However, in addition to the necessary legal foundations, there is a lack of broadly applicable, standardised processes**, so that data linkage is still only possible to a very limited extent and is laborious. Thus, despite the efforts made up to now, statistics from different institutions **are hitherto linked exclusively on a case-by-case basis** for concrete research projects. This is due to the regulatory framework.

For example, **section 75 of the German Social Code (SGB) X** implicitly allows the **linking of various data from social security system** for specific research projects. ↘ CHART 144 However, these must be approved on a case-by-case basis by the competent federal authority. On the basis of **section 13a of BStatG**, **data from the Deutsche Bundesbank** can also be transmitted to the FSO/LSOs system and **linked by them with data from the federal statistics**. ↘ CHART 144 However, this option is only used in the context of specific

research projects (Kruse et al., 2023). At least, the data linked in this way are made available to the scientific community via the FDZ of the Federal Statistical Office.

## 5. Data access not user-friendly

549. Last but not least, **access to data that are available in principle is laborious and complex**: Applying for access to data is time-consuming, while data access itself is costly and not user-friendly. Especially problematic are situations in which the access to data is possible only through workstations offered at the data providers' offices or controlled remote data processing; this applies to many data sets of official statistics but also to other FDZs. Controlled remote data processing, for example, requires the preparation of programme routines by researchers. These routines are sent to the research data centres, checked and executed there, after which the results are examined and finally released to the researchers. This is a lengthy process with additional limitations on the number of variables due to technical restrictions. For example, in the Taxpayer Panel, a maximum of 25 variables out of more than 1,000 available variables can be used at any one time (Advisory Board to the Federal Ministry of Finance, 2021). These **cumbersome access routes lead to higher costs, bottlenecks in physical infrastructure and a lower level of user-friendliness** than fully remote data access via secure online access (RatSWD, 2019, 2022b).
550. **Data protection restrictions** are a major cause of **user-unfriendliness in data access**. The legal principles of data protection are interpreted differently in the individual Länder, making access to register data for research purposes more difficult (RatSWD, 2023a). In this context, the research privilege pursuant to Article 89 of the GDPR (General Data Protection Regulation) has hitherto been interpreted very cautiously, and scientific research has been treated as subordinate to the interests of data protection in the weighing up of interests. For example, a different interpretation of the formal anonymity of the data means that Bavarian company data cannot even be viewed directly by researchers during a physical visit to the research data centre (on-site access). Instead, only controlled remote data processing with all its restrictions is offered for this purpose. [▶ ITEM 549](#)
551. In addition, there is often only **limited information about the structure of the individual data sets or inadequate meta-data and data-set documentation**, so that users have to spend a lot of time exploring and editing data. This is particularly costly in the absence of remote access. For example, data-structure files are not made available for all data sets. Such files would enable researchers to better prepare for data work because they contain randomly generated data that reflect the structure of the real data set. This means that analysis files can be prepared in advance and time on site at the research data centre can be used much more efficiently, taking pressure off the physical infrastructure on site.
552. **Even for less sensitive data that is available as aggregated or micro-aggregated data sets, access in Germany is not very user-friendly**. For example, the standard access GENESIS-ONLINE offered by Destatis provides a much

less user-friendly interface than other data providers, especially for time-series data (Bachmann et al., 2023a). Time-series data, even if they originate from Destatis, can be accessed and visualised much more easily via FRED, the data portal of the St. Louis Federal Reserve, or the time-series database of the Deutsche Bundesbank.

### III. OPTIONS FOR ACTION: A DATA INFRASTRUCTURE FOR THE 21ST CENTURY

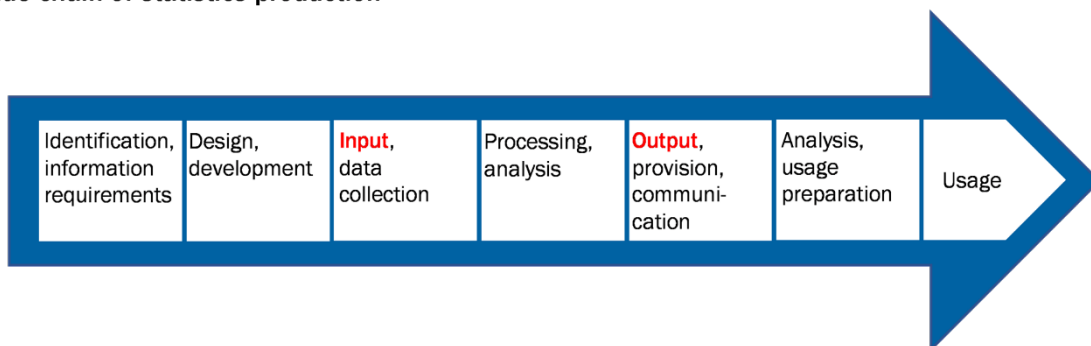
#### 1. Improve the regulatory framework of the FSO/LSOs system

553. The paradigm shift in official statistics from input orientation to output orientation (Hahlen, 2009), which has been an objective since the turn of the millennium, should finally be worked out in detail and implemented. This means focusing on information requirements and on services to be provided rather than on data sources that may be used. However, the **legal framework of official statistics is still input-oriented**, i.e., it focuses on the exact definition of the data that official statistics are allowed to collect. [↪ CHART 145](#) Up to now, according to section 9 of BStatG, legal provisions – often formulated as individual laws for specialised statistics – must be adopted for official statistics, provisions that define in detail which characteristics may be collected from which subjects. However, this means that legal adjustments are often required in order to adapt new statistics and surveys and for new ways of use. This considerably reduces the adaptability of official statistics. [↪ ITEM 562](#)

For this reason, **legislation on statistics**, especially section 9 of BStatG, should **be modified** in accordance with the **output orientation** approach. Such an adjustment would enable the FSO/LSOs system to fulfil the tasks generally defined in the law through suitable surveys without having to legitimise them in

[↪ CHART 145](#)

**Value chain of statistics production**



Source: KomZS (2023)  
© Sachverständigenrat | 23-470-01

detail through individual laws on specialised statistics. Thus, the FSO/LSOs system should be authorised to itself define the surveys that are required – including characteristics and samples, as well as the necessary access to the register and administrative data needed to meet an information requirement defined by law. This would significantly increase the flexibility and responsiveness of the FSO/LSOs system.

- 554. Improving the data infrastructure requires additional financial and human resources.** The resources needed for additional services, and especially for additional data products, can be reduced if existing products are designed more efficiently. Many of the following proposals are therefore based on the adaptation and improvement of existing products or the use of existing, but not yet used, data sources, as well as the conversion of existing data collection methods to register-based methods. However, it will not be possible to provide all products and services in this way as some will involve a lot of additional effort. In particular, the required improvement of the digital data infrastructure and the development and use of administrative data and registers will generate additional expenditure.

Therefore, the **budgets** of the corresponding institutions, especially those responsible for official statistics, should be expanded. Such budget expansion can nevertheless have a cost-saving effect for the state, since the additional funds should be seen as a financial **investment in a better decision-making basis for policy measures, which can reduce the costs of** such measures through a more targeted design and continuous evaluation. If, for example, the additional costs are generously estimated at 50 % of Destatis' average annual budget of €260 million in the years 2021 to 2023, this would result in additional annual expenditure of 130 million euros. If we assume, for example, that in the case of assistance for companies, e.g., COVID-19 assistance, a reduction in deadweight effects amounting to 5 % of the total costs were possible, then the costs for the COVID-19 assistance measures for companies alone, whose disbursement volume has so far amounted to 46 billion euros, would have been 2.3 billion euros lower. This would have made it possible to stem the above-mentioned additional costs for almost 18 years.

- 555.** In order to improve the range of services available to the scientific community, it would be commendable to **enshrine a research mandate for Destatis** in the BStatG. The granting of a research mandate to Destatis could follow the example of research mandates to the U.S. Census Bureau (pursuant to Department of Commerce DOO 35-2B), the German Federal Employment Agency (pursuant to section 280 of SGB III) or the Deutsche Bundesbank (implicitly derived from the monetary-policy mandate and the financial-stability mandate pursuant to section 3 of BBankG [Deutsche Bundesbank Act] as well as the need for a solid research basis to fulfil the mandate). These each operate their own research centres: the Center for Economic Studies, the IAB and the Research Centre of the Deutsche Bundesbank.

**A statutory research mandate for Destatis enshrined in section 3 of BStatG and the establishment of a research centre at Destatis** would have at least three advantages. First, Destatis researchers would work on specific



research issues and could thus immediately develop solutions to the research data problems that inevitably arise in practice. This would contribute to a rapid and continuous improvement of the available research data. Second, Destatis' own interest in the preparation, provision and use of research data would increase if it was also evaluated with regard to the research mission. This would have positive effects on the available research data. Third, it can be assumed that the researchers at the research centre would contribute to high-quality scientific research with German data, and thus to a better understanding of the German economy. ↘ [BOX 30](#)

556. **Furthermore, the FDZs of the Federal Government and the Länder** should be **legally enshrined in the BStatG as a core task** of the FSO/LSOs system. At present, the FDZs are not explicitly mentioned in the BStatG. The provision of data to external researchers is only indirectly regulated via the BStatG. Legal codification would emphasise the importance of the research data centres at Destatis and the statistical offices of the Länder and explicitly point out that a corresponding institution (FDZ) and corresponding personnel and financial resources are needed for the provision of data. First, this would ensure the long-term operation of the FDZs. Second, it would solve the current problem of self-financing of the data centres by the data users, which leads to comparatively high costs of data access. Public funding would be justified insofar as the research results would benefit not only the data users themselves but society as a whole. Positive external effects are thus created. Direct funding would help to internalise these external effects.

## 2. Close data gaps

557. The lack of comparative analyses of the education systems in different Länder could be addressed by means of **nationwide individual pupil statistics** (Hertweck et al., 2023) that would provide panel data **on the entire school career of pupils**. Administrative data should be supplemented by competence measurements and linked across school transitions. Such statistics could form the **core of educational progression register** from kindergarten through school to higher education. In order to be able to compare the effectiveness of different education systems and the quality of different schools, it is essential to establish regular and comprehensive competence measurements. The results could be made available for research as part of individual pupil statistics.
558. The problem of the lack of **official data on wealth distribution** could be addressed comparatively cost-effectively by **linking the microcensus with administrative data** and developing suitable procedures for estimating wealth using administrative data (Bachmann et al., 2023b). Linkage with wage- and income-tax data could relieve the burden on respondents, improve the quality of information on income and make it possible to estimate financial and real estate assets. This would address not only the lack of wealth data, but also the problem of the distorted representation of respondents' income situation that arises with self-reported data (Emmenegger and Münnich, 2023). Business assets, in turn, can be determined on the basis of data from the Deutsche Bundesbank on company accounts. Data from a future register of buildings and apartments could be

used to determine real estate assets. To enable this, efforts should be made as soon as possible to establish consistent identifiers across data sets. The tax ID is likely to be suitable for this purpose. An obligation to provide information on the household equity holdings and to state the respective commercial register number should be suitable for linking the personal data with company balance sheet data. Furthermore, linking the survey data of the microcensus with other administrative data must be permitted by law.

559. The lack of information on the building stock in Germany could improve significantly with the planned amendment of the EU Building Directive (European Parliament, 2023). This amendment provides for the establishment of national databases for energy-certificate data and the introduction of building logbooks, which could form the basis for a building register. **A building register would be useful, and the information it contains should be carefully selected as it can form the data basis for research in the field of energy and social policy.** For example, information on rental apartments, which account for more than half of the housing stock in Germany (Krause et al., 2022), has hardly been recorded so far, yet it is highly relevant for social policy, for example. Data on existing heating systems, which show which energy sources are used to generate heat, are also key.
560. From a macroeconomic and economic policy perspective, there is also a need for a **high-frequency household survey** that maps the economic situation (labour market, consumption, expectations, income, assets, debt, etc.) of German households almost in real time (Bachmann et al., 2021). Especially during the COVID-19 pandemic, the latest economic developments were only known to policy-makers in broad outlines, and economic policy measures could therefore often not be precisely tailored. A systematic switch of the SOEP to a rolling (quarterly) survey frequency could be a step in the direction of such a survey. A similar system exists as part of the rolling survey of the CPS in the United States. The microcensus, being the largest household survey in Germany, is made up of reporting weeks evenly distributed over the year. The information on the reporting weeks in the micro-census has hardly been used to date. The legal framework for the micro-census should furthermore be designed in such a way that new survey modules can be incorporated at short notice and on an ad-hoc basis. In addition, a household survey with an even higher frequency (weekly, fortnightly), which could be activated in times of crisis, would be helpful. In a similar way, a high-frequency business survey could be designed along the lines of the Business Pulse Survey of the U.S. Census Bureau or the Bundesbank Online Panel – Firms (BOP-F) of the Deutsche Bundesbank.
561. In future, new legislative projects should involve an obligation to integrate an accompanying system of data collection, data provision and timely scientific evaluation if the volume of funding yet to be defined is exceeded. For example, **scientific monitoring and evaluation, including data collection**, could be prescribed for **measures with a funding volume of more than 100 million euros per year**. Up to now, there is only an obligation to conduct an internal evaluation for regulatory projects with a compliance cost in excess of 5 million euros (Federal Government, 2013, 2019). In addition, research clauses, i.e.

research-friendly, barrier-free access rules, should be created for independent scientific research. Ideally, in the case of measures that affect companies or households, the identifier of those affected should be collected. In this way, the data could be linked to information from official process or register data.

### 3. Less time delay and greater adaptability

562. Some of the new data products proposed, in combination with output-oriented statistics legislation [▶ ITEM 553](#), could help to **increase the speed and responsiveness of official statistics**, in addition to improving the amount of data available. The survey programme of a high-frequency household survey [▶ ITEM 560](#) could be adapted at short notice to provide useful information in crisis situations. For example, in the wake of the COVID-19 pandemic, the questions in the monthly CPS in the United States were adjusted at short notice to measure the impact of the pandemic. [▶ ITEM 531](#) Similarly, an obligation to collect data and conduct a researcher-designed evaluation in the case of policy measures above a certain volume could ensure that data for evaluating these measures are available in a timely manner.
563. In order to **reduce the time lag** in the provision of **research data and the production of public data products based on official statistics**, data collection should take place in a timelier way. In many cases, this can be achieved by making greater use of administrative process data, which should be made interlinkable with uniform identifiers. In particular, there is potential here to make more administrative process data available by increasing the digitisation of administrative processes. In the field of health data, for example, data from the electronic patient file could be used. Another aspect to be considered is shortening the reporting deadlines for selected statistics and accelerating data transmission. In doing so, however, additional burdens for the reporting entities, e.g., companies, would have to be borne in mind. A timelier processing of the data is necessary, especially in the case of the microcensus, in order to make survey data available as official statistics more quickly. Similarly, research data should also be made available more promptly, which would mean accelerating the preparation of data for the scientific community. Both require additional financial and human resources.

### 4. Improve the availability of data already collected

564. All data sources that have been publicly funded should be available to independent scientific research on principle. Personal and less aggregated data are particularly important for economic research because they can be used to analyse impact channels of economic-policy measures and their heterogeneous effects (RatSWD, 2023a). An **expansion of administrative microdata available for research should therefore be prioritised**. Providing such data is likely more cost-efficient than collecting new data and is not associated with the additional reporting burdens, especially for companies and individuals. The amount of work can be further reduced if **data collection by different data**

**producers is harmonised**, for example by using shared data portals. Harmonisation could be linked to the **once-only principle** to avoid entering the same data several times.

565. Due to their comprehensive coverage and the high quality of data collection, **administrative data** are particularly valuable for **empirical research** (Currie et al., 2020), as they are especially reliable. Areas with good availability of such data offer researchers ideal conditions for particularly deep and thorough investigations (Nagaraj and Tranchero, 2023). International researchers take advantage of this, as shown by the frequently studied data on the German labour market. [↪ BOX 30](#) This high-quality empirical research, in turn, provides **well-evidenced findings on the basis of which policy-makers can make quick and informed decisions**.

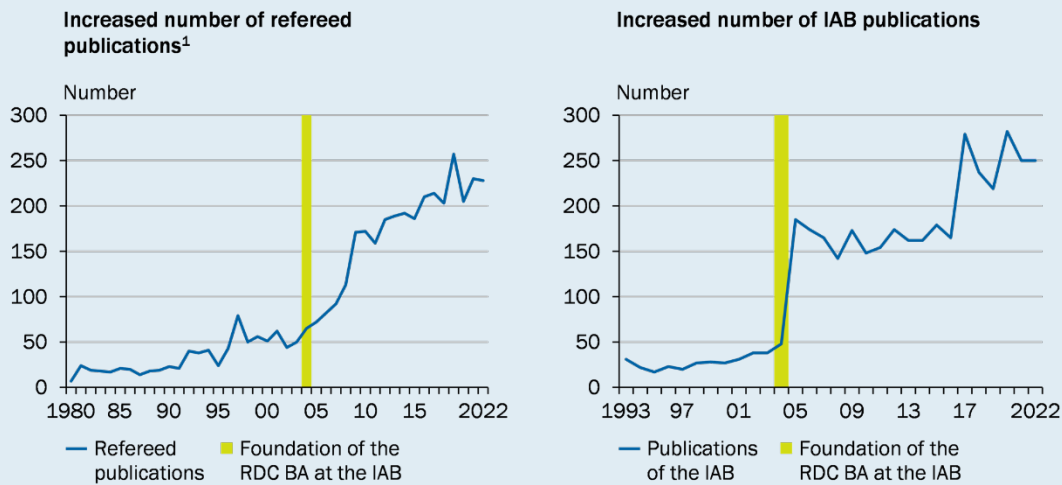
[↪ BOX 30](#)

#### Background: Access to research data and the IAB's research – a success story

The **Institute for Employment Research (IAB)** is a department of the Federal Employment Agency (BA) that has been in operation since 1967. The IAB **was tasked with labour-market and occupational research** and has at its disposal the administrative data from the BA's area of responsibility and from other statistical sources for this purpose. The IAB has made extensive contributions to labour-market and occupational research, and these are reflected in a large number of its own publications. [↪ CHART 146 RIGHT](#)

[↪ CHART 146](#)

#### Publications on the German labour market



1 – Annual number of publications in scientific journals shown through the Web of Science portal when following search criteria are entered: „Germany“ and („employment“ or „unemployment“ or „wage“) in the „Subject“ field and „Economics and Business“ in the „Area“ field.

Sources: IAB, Web of Science

© Sachverständigenrat | 23-465-01

In addition to engaging in research with process data from the BA, the IAB also collects its own data. **Data evaluations were originally carried out within the IAB**, and microdata relevant to the labour market were rarely shared with external actors. This led to an underutilisation of data resources and to very limited cooperation projects with national and international academic researchers. The resulting analyses were mostly descriptive; only in a few cases were

they published in peer-reviewed journals (Müller and Möller, 2019). Over time, however, the demand for empirical evidence on labour markets has increased significantly, especially as a result of the Hartz reforms. In addition, the Commission on Improving the Informational Infrastructure between Science and Statistics (2001) made proposals for using statistical surveys more efficiently, including one to establish FDZs at every government data producer, e.g., also at the IAB.

The **FDZ BA at the IAB** was founded in 2004 and since then has **provided external researchers with the data** available at the IAB through **standardised data products**. Since then, the number of researchers using IAB microdata to study the German labour market has risen rapidly. For example, about 1,400 researchers are currently working on more than 650 research projects. This has resulted in a high number of peer-reviewed publications on the German labour market. [↪ CHART 146 LEFT](#) The creation of access points in other European countries and in the United States from 2012 onwards has significantly increased the number of foreign researchers using IAB data (Müller and Möller, 2019).

566. In addition, **the preparation and provision of historical macroeconomic data** would be **valuable**. For example, access to some national-accounts data has only been available in a user-friendly way since 1991 onwards, even though they were collected by Destatis much earlier in some cases. Making data available from the 1950s onwards, as in the United States, would be important for an improved understanding of macroeconomic relationships. Furthermore, long-running data on variables such as investment and capital-stock series, some of which go back to the 19th or even 18th century, would be helpful for macroeconomic investment and capital-stock research. These data could be made available, for example, within the framework of Destatis' 'EXSTAT – Experimental Statistics' programme.
567. The availability of data collected in the past is made more difficult by deletion regulations such as **section 13a of the Federal Statistics Act (BStatG)**. [↪ ITEM 546](#) This is problematic for empirical economic research and policy advice, as long-running time series are necessary for identification of impact mechanisms. Deletion regulations for data should be **abolished**, especially the deletion regulation in section 13a of BStatG. The GDPR's research privilege should be used, as proposed by the German Data Forum (2023b), to promptly establish long-term storage periods for research data.

## 5. Better linkability across different data providers

568. **Data trustees should be established to link microdata across the existing data silos of** individual data providers. Data trustees are data intermediaries who combine data from different sources and make it available to authorised users in a secure and legally compliant manner. This applies in particular to data from the public administration, official statistics and publicly co-financed research institutions. Macroeconomic data and data from the companies should also be integrated. Data trustees would make such data linking from different sources possible. **Ideally, data trustees would be created by** expanding Destatis' legal mandate, for example by **establishing a microdata centre**

**within Destatis, modelled on the AMDC in Austria.** [↪ BACKGROUND INFO 24](#)  
 Thereby, it should be ensured that sufficient resources are available to provide researchers with cost-effective data access and that the administration and the register-keeping agencies cooperate. Contrary to what was originally envisaged, the data institute established by the Federal Government (BMWK and BMI, 2023) is not designed to take on the task of a data trustee for research.



[↪ BACKGROUND INFO 24](#)

**Background: Austrian Micro Data Center as a model for Germany**

In Austria, all research institutions that meet the access requirements under the Research Organisation Act can apply to use and link administrative data via the **Austrian Micro Data Center (AMDC; Fuchs et al., 2023)**. The range of data offered is very extensive: in addition to all the micro- and register data used by Statistik Austria, numerous other federal administrative data are available (Altmann et al., 2023). In addition, researchers can request a linkage of administrative data with their own data. There are various possibilities for data access for the approved projects, including secure remote access. The AMDC acts as a **central information point** in accordance with Article 8 of the Data Governance Act and as a **data trustee**, which in particular makes it possible to link the various data sets. The establishment of the AMDC in July 2022 was prepared by numerous amendments to the law, which **enshrined the tasks and rights of the AMDC in law** (Altmann et al., 2023).

569. Until the data-trustee strategy is fully implemented, **reducing the legal obstacles to linking registers** with other administrative or scientific data remains an important task. The **expansion of section 13a of BStatG is particularly relevant** here, as it defines the limits for merging official data. Thus, according to the current design, only the economic statistics of the Deutsche Bundesbank can be transmitted to Destatis, but not, for example, the data of German Pension Insurance. The Research Data Act should explicitly include the possibility of linking data for research and extend the linkability to other data producers alongside the Deutsche Bundesbank and Destatis. For this, a **number of individual laws** would have to be amended, as noted by the RatSWD (2023b). **Section 7a** of the Tax Statistics Act (**StStatG**) would have to allow individual data from tax statistics to be merged with data from other public data producers. In the Administrative Data Usage Act (**VwDVG**) and the Statistics Register Act (**StatRegG**), the principle would also have to be enshrined that all data producers of public statistics may exchange and link their data for scientific research purposes, at least via a data trustee.
570. In order to reduce the effort involved in data linkage both within official statistics and with external datasets, **unique identifiers across datasets** need to be introduced as quickly as possible. For example, the commercial register number or the VAT number could be used for companies (Gottschalk et al., 2023); the establishment registration number recorded by the Establishment Registration Number Service (BNS) could serve as the establishment ID. This standardisation will be particularly relevant in the future when the linking options are expanded and new registers, such as educational progression registers, are set up (RatSWD,



2022a). However, the lack of consistency among identifiers also increases processing costs for data links that are already approved (Gottschalk et al., 2023).

571. Since the linked research data are available to scientists but not to political decision-makers or the general public, **social acceptance for data linkage** is likely to be higher than in the case of linkage for administrative purposes. If the public is informed about the **societal benefits of the enabled research**, social acceptance of data linkage is likely to increase.

## 6. Simple and user-friendly data access

572. In order to simplify data access for the scientific community, a **scientific or research privilege** should be enshrined as a key component of the planned Research Data Act. This **should replace existing individual regulations on data access, use and protection** and regulate them in a uniform manner throughout Germany. In particular, the research privilege should ensure that scientific research is given sufficient weight when weighing against data protection. In addition, generous linkage rules for data should be prescribed in the Act. [▶ ITEM 569](#) Since scientific research contributes significantly to the common good, everything that is not explicitly and justifiably prohibited should be permitted in the area of data access, linkage and use for science. A Research Data Act would make a significant contribution to legal certainty, since, at present, data linkages are often prevented by the different legal opinions of the authorities involved on the interpretation of existing data protection law.
573. The highly decentralised research data infrastructure in Germany leads to unnecessary hurdles in data access. For example, in some cases research projects at the FDZs of the Federal Government and the Länder must obtain the consent of all the Länder individually. Instead, **access to research data should be more strongly coordinated and centralised in the FSO/LSOs system**. The standard procedure should be shifted from requiring the consent of all the authorities involved to an objection procedure, or else a centralised procedure should be introduced. In this way, the approval of a central authority, for example at Destatis, could suffice for the approval of research projects. In addition, there should be a nationally uniform interpretation of data-protection rules; up to now, interpretation inconsistencies have been an obstacle to access (RatSWD, 2023b). In order to achieve these goals, **section 16 (6) of BStatG should be formulated in such a way that all actors in the FSO/LSOs system regulate** research data access, confidentiality and data protection **in a uniform manner**. To this end, Destatis, for example, could be given the authority to draw up uniform access and confidentiality guidelines in this area that apply to the entire FSO/LSOs system.
574. The user-friendliness of **data access to microdata** for science could be increased by a **central registration of researchers** and research institutions. A new institution at the German Data Forum (RatSWD) could act as a central registration office for all research data centres. Alternatively, accreditation at one FDZ could automatically grant access to all the others. A prerequisite for such solutions, however, is a **standardisation of the legal basis for data access**.

With the introduction of a new set of indicators – for example on the processing time for applications – selected aspects of the quality of data access could be tracked to identify problem areas. To ensure that the security of the data is not affected by the increase in user-friendliness, diligence is crucial when checking the applicant institutions' authorisation to use the data. In this context, one conceivable solution might be to impose **access restrictions on institutions whose researchers demonstrably do not comply with data-confidentiality rules**. Such a procedure is the rule in Denmark, for example.

575. In order to create a transparent data infrastructure, a **publicly funded, non-restricted access platform to provide a data overview** is necessary. Such a platform was prescribed as a central information point **in the European Data Governance Act**. This regulation should have been transposed into national law by September 2023. Germany has not yet implemented the Data Governance Act and is therefore in default. This could lead to infringement proceedings under Article 258 of the TFEU. Rapid, research-oriented implementation by the Federal Government is therefore urgently required. Parallel to this, an improvement in the available metadata would be helpful. Furthermore, data-structure files should be made available that can be used off-site to prepare research stays. [▶ ITEM 551](#)
576. **Adjustments to statistics legislation** are necessary in order to enable nationwide **access to formally anonymised data via a remote access** system. For example, remote access would have to be regulated in section 16 (6) of BStatG for access to official data, or in section 75 SGB X for access to BA data (RatSWD, 2023a). The legal requirements for linking registers must be created, and there must be a unique identifier for implementation, as is the case in many European countries.
577. The **user-friendliness of access to more aggregated, publicly available data should also be improved**. In the case of macroeconomic time series, especially the publicly accessible databases of the Federal Reserve Bank of St. Louis (FRED), the time-series database of the Deutsche Bundesbank, and the recently overhauled data service of the OECD can serve as models for improving GENESIS-ONLINE, or for creating an additional access path for time-series data (Bachmann et al., 2023a). In addition, the data should be made available in a form that allows the most versatile use possible. **Databases that do not contain ready-made aggregates but rather microdata that can be aggregated on an inquiry-specific basis are particularly helpful** here. Such data access makes it possible for users to construct their own statistics. For example, it would be possible to output the moments of the statistical distribution – such as the median or the standard deviation and not just the mean – and to aggregate the data according to the research question. A corresponding infrastructure is currently being tested by the Rhineland-Palatinate Statistical Office, among others, and could go into regular operation in 2024. In order to realise economies of scale and make data access more user-friendly overall, it would be expedient to extend the infrastructure to include the entire FSO/LSOs system.

# APPENDIX

▾ TABLE 1

## Accredited research data centres in Germany in 2023

Abbreviation	Designation
BAMF-RDC	Research Data Centre of the Federal Office for Migration and Refugees
BIBB-RDC	Research Data Centre of the Federal Institute for Vocational Education and Training
DeZIM.fdz	Research Data Centre of the German Centre for Integration and Migration Research
EBDC	LMU-ifo Economics & Business Data Center
IOER-Monitor	Monitor of Settlement and Open Space Development
RDC ALLBUS	Research Data Centre ALLBUS at GESIS
RDC at KBA	Research Data Centre at Kraftfahrt-Bundesamt
RDC BA at IAB	Research Data Centre of the German Federal Employment Agency (BA) at the Institute for Employment Research (IAB)
RDC Bildung	Research Data Centre for Education at the DIPF   Leibniz Institute for Research and Information in Education
RDC BZgA	Research Data Centre of the Federal Centre for Health Education
RDC eLabour	Centre for Qualitative Data in the Field of Sociology of Labour and Industry
RDC Elections	Research Data Center Elections at GESIS
RDC GePaRD	German Pharmacoepidemiological Research Database
RDC GML	Research Data Centre German Microdata Lab (GML) at GESIS
RDC International Survey Programmes	Research Data Center International Survey Programmes at GESIS
RDC IZA, IDSC	International Data Service Centre (IDSC) at the Institute for the Study of Labour (IZA)
RDC pairfam	Research Data Center of the German Family Panel
RDC PIAAC	Research Data Center PIAAC at GESIS
RDC PsychData at ZPID	Research Data Center of the Leibniz Institute for Psychology
RDC Qualiservice	Research Data Centre Qualiservice
RDC Ruhr at RWI	Research Data Centre Ruhr at the RWI – Leibniz-Institut für Wirtschaftsforschung
RDC SHARE	Research Data Center of the Survey of Health, Ageing and Retirement in Europe
RDC Wissenschaftsstatistik	Research Data Centre Wissenschaftsstatistik of Stifterverband
RDC-AGD	Research Data Centre Archive for Spoken German at the Leibniz Institute for the German Language
RDC-aviDA	Research Data Centre for audio-visual data of qualitative social research
RDC-BAuA	Research Data Centre of the Federal Institute for Occupational Safety and Health
RDC-BO	Research Data Center for Business and Organizational Data
RDC-Bund	Research Data Centre of the Federal Statistical Office
RDC-DJJ	Research Data Center of the German Youth Institute
RDC-DZA	Research Data Centre of the German Centre of Gerontology
RDC-DZHW	Research Data Centre for Higher Education Research and Science Studies
RDC-IQB	Research Data Center of the Institute for Educational Quality Improvement (IQB)
RDC-IWH	Research Data Centre of the Halle Institute for Economic Research
RDC-Land	Research Data Centre of the Statistical Offices of the Länder
RDC-LifBi	Research Data Center of the Leibniz Institute for Educational Trajectories
RDC-RISC	Research Data Centre of the Research Institute Social Cohesion
RDC-RKI	Research Data Centre of the Robert Koch Institute
RDC-RV	Research Data Centre of the German Pension Insurance
RDC-SOEP	Research Data Center of the Socio-Economic Panel Study at DIW Berlin
RDSC Bundesbank	Deutsche Bundesbank Research Data and Service Centre
SAFE RDC	Research Data Center of the Leibniz Institute for Financial Research SAFE
ZEW-RDC	ZEW Research Data Centre

Source: RatSWD

© Sachverständigenrat | 23-319-01

# REFERENCES

- [Advisory Board to the Federal Ministry of Finance](#) (2021), Gutachten des Wissenschaftlichen Beirats beim BMF zur Verbesserung der Dateninfrastruktur für die Steuerpolitik, Monatsbericht des BMF Januar 2021, 30–36.
- [Alstadsæter, A. et al.](#) (2020), The first weeks of the coronavirus crisis: Who got hit, when and why? Evidence from Norway, NBER Working Paper 27131, National Bureau of Economic Research, Cambridge, MA.
- [Altmann, S. et al.](#) (2023), Statement der UAG Arbeitsmarkt und Sozialversicherung, RatSWD Working Paper, German Data Forum (RatSWD), Berlin, in press.
- [Bachmann, R., A. Balleer, Z. Enders, N. Garnadt and H. Stüber](#) (2023a), Antworten zur öffentlichen Konsultation zum Forschungsdatengesetz, Fragebogen zum Forschungsdatengesetz, Verein für Socialpolitik, AG Daten, Unter-AG Makroökonomik, Berlin.
- [Bachmann, R., A. Balleer, Z. Enders, N. Garnadt and H. Stüber](#) (2023b), Für einen besseren Datenzugang: Stellungnahme der Arbeitsgruppe Makroökonomische Daten, RatSWD Working Paper, German Data Forum (RatSWD), Berlin, in press.
- [Bachmann, R., A. Peichl and R.T. Riphahn](#) (2021), Bessere Daten – bessere Politik, Frankfurter Allgemeine Zeitung, Frankfurt am Main, 6 August.
- [Bender, S., J. Wagner and M. Zwick](#) (2007), KombiFiD – Kombinierte Firmendaten für Deutschland: Konzeption der Machbarkeitsstudie für eine Zusammenführung von Unternehmensdaten der Statistischen Ämter, des Instituts für Arbeitsmarkt- und Berufsforschung der Bundesagentur für Arbeit und weiterer Datenproduzenten, FDZ Methodenreport 5/2007, Research Data Centre of the German Federal Employment Agency at the Institute for Employment Research, Nuremberg.
- [BMWK and BMI](#) (2023), Konzept zum Aufbau des Dateninstituts, Federal Ministry for Economic Affairs and Climate Action and Federal Ministry of the Interior and Community, Berlin.
- [Bonhomme, S., K. Holzheu, T. Lamadon, E. Manresa, M. Mogstad and B. Setzler](#) (2023), How much should we trust estimates of firm effects and worker sorting?, *Journal of Labor Economics* 41 (2), 291–322.
- [Brandt, M., D. Oberschachtsiek and R. Pohl](#) (2007), Neue Datenangebote in den Forschungsdatenzentren – Betriebs- und Unternehmensdaten im Längsschnitt –, FDZ-Arbeitspapier 23, Statistische Ämter des Bundes und der Länder, Wiesbaden.
- [Bundesregierung](#) (2019), Fortentwicklung der Evaluierungskonzeption der Bundesregierung, Beschluss des St-Ausschusses Bessere Rechtsetzung und Bürokratieabbau vom 26. November 2019, Press release, Berlin, 27 November.
- [Bundesregierung](#) (2013), Konzeption zur Evaluierung neuer Regelungsvorhaben, Vom St-Ausschuss Bürokratieabbau am 23 January 2013 beschlossene Fassung, Berlin.
- [Currie, J., H. Kleven and E. Zwiars](#) (2020), Technology and big data are changing economics: Mining text to track methods, *AEA Papers and Proceedings* 110, 42–48.
- [Decker, J. and D. Vögrimmer](#) (2021), Die Rolle der amtlichen Statistik innerhalb des Datenökosystems, *WISTA – Wirtschaft und Statistik* 3/2021, 17–26.
- [Dickopf, X., C. Janz and T. Mucha](#) (2019), Vom BIP-Flash zum BIP-Nowcast: Erste Ergebnisse einer Machbarkeitsstudie zur weiteren Beschleunigung der BIP-Schnellschätzung, *WISTA – Wirtschaft und Statistik* 6/2019, 47–58.
- [Emmenegger, J. and R. Münnich](#) (2023), Localising the upper tail: How top income corrections affect measures of regional inequality, *Journal of Economics and Statistics* 243 (3–4), 285–317.
- [European Parliament](#) (2023), Energy performance of buildings (recast), Amendments adopted by the European Parliament on 14 March 2023 on the proposal for a directive of the European Parliament and of the Council on the energy performance of buildings (recast) (COM(2021)0802 – C9-0469/2021 – 2021/0426(COD)).
- [Fuchs, R., T. Göllner, S. Hartmann and T. Thomas](#) (2023), Fostering excellent research by the Austrian Micro Data center (AMDC), *Journal of Economics and Statistics*, im Erscheinen, <https://doi.org/10.1515/jbnst-2023-0043>.

[Gottschalk, S., C. Krolage, G. Licht, A. Peichl, S. Schaffner and S. Wichert \(2023\)](#), Besserer Zugang zu Unternehmensdaten für die empirische Wirtschaftsforschung in Deutschland”, RatSWD Working Paper, German Data Forum (RatSWD), Berlin, in press.

[Gürke, C., A. Gruhl and T. Hethy-Maier \(2011\)](#), Verknüpfung von Unternehmensdaten verschiedener Datenproduzenten Das Projekt „Kombinierte Firmendaten für Deutschland“, WISTA – Wirtschaft und Statistik 2/2011, 91–97.

[Hahlen, J. \(2009\)](#), Zur Rolle der amtlichen Statistik für eine evidenzbasierte Wirtschaftsforschung und -politik, WISTA – Wirtschaft und Statistik 10/2009, 1021–1030.

[Hertweck, F., I.E. Isphording, S.H. Matthewes, K. Schneider and C.K. Spieß \(2023\)](#), Bildungspolitik braucht Forschung – doch Forschung braucht bessere Daten, RatSWD Working Paper, German Data Forum (RatSWD), Berlin, in press.

[Kommission zur Verbesserung der informationellen Infrastruktur zwischen Wissenschaft and Statistik \(Eds.\) \(2001\)](#), Wege zu einer besseren informationellen Infrastruktur: Gutachten der vom Bundesministerium für Bildung und Forschung eingesetzten Kommission zur Verbesserung der Informationellen Infrastruktur zwischen Wissenschaft und Statistik, 1st edition, Nomos Verlagsgesellschaft, Baden-Baden.

[KomZS \(2023\)](#), Abschlussbericht, Kommission Zukunft Statistik, Wiesbaden, in press.

[Krause, A., M. Zimmermann and I. Herda \(2022\)](#), Überlegungen zu einem Gebäude- und Wohnungsregister: Aufbau, Pflege und Nutzung, WISTA – Wirtschaft und Statistik 4/2022, 25–38.

[Kruse, H.W., F. Hieber, H. Limberg, B. Zapf and D. Boddin \(2023\)](#), Außenhandelsaktive Unternehmen: neue Analysemöglichkeiten durch Mikrodatenverknüpfung, WISTA – Wirtschaft und Statistik 5/2023, 13–26.

[Löffler, M., A. Peichl, C. Wittneben and C. Woodage \(2015\)](#), Möglichkeiten zur Verbesserung der statistischen Datengrundlage zur Beschreibung höchster Einkommen und Vermögen: Abschlussbericht, Research Report, ZEW Kurzexpertise 58, Centre for European Economic Research, Mannheim.

[Ludvigsson, J.F., P. Svedberg, O. Olén, G. Bruze and M. Neovius \(2019\)](#), The longitudinal integrated database for health insurance and labour market studies (LISA) and its use in medical research, European Journal of Epidemiology 34, 423–437.

[Müller, D. and J. Möller \(2019\)](#), Giving the international scientific community access to German labor market data: A success story, in: Crato, N. and P. Paruolo (Eds.), Data-Driven Policy Impact Evaluation: How Access to Microdata is Transforming Policy Design, Springer International Publishing, Cham, 101–117.

[Nagaraj, A. and M. Tranchero \(2023\)](#), How does data access shape science? Evidence from the impact of U.S. census's research data centers on economics research, NBER Working Paper 31372, National Bureau of Economic Research, Cambridge, MA.

[Palmer, K., B. Prest, S. Iler and S. Villanueva \(2022\)](#), Options for EIA to publish CO2 emissions rates for electricity, RFF Report 22–8, Resources for the Future, Washington, DC.

[Profit, S. \(2023\)](#), Eine moderne Dateninfrastruktur für die evidenzbasierte Politikberatung in Deutschland, Speech, Wissenschaftliche Fachtagung „Daten.Forschung.Zukunft“, Session "Datenkompetenz und Nationale Dateninfrastruktur: Herausforderungen und Schlüssel für Demokratie und Entscheidungen“, Wiesbaden, 6 July.

[RatSWD \(2023a\)](#), Positionspapier des RatSWD: Nutzung von Registerdaten für Zwecke der Forschung sichern, German Data Forum (RatSWD), Berlin.

[RatSWD \(2023b\)](#), Forschungsdatengesetz: Was zentral ist, German Data Forum (RatSWD), Berlin.

[RatSWD \(2022a\)](#), Positionspapier des RatSWD: Aufbau eines Bildungsverlaufsregisters: Datenschutzkonform und forschungsfreundlich, German Data Forum (RatSWD), Berlin.

[RatSWD \(2022b\)](#), Positionspapier des RatSWD: Eckpunkte für ein Forschungsdatengesetz, Empfehlungen des Rates für Sozial- und Wirtschaftsdaten für die Erarbeitung eines Forschungsdatengesetzes, German Data Forum (RatSWD), Berlin.

[RatSWD \(2019\)](#), Remote Access zu Daten der amtlichen Statistik und der Sozialversicherungsträger, RatSWD Output 5(6), German Data Forum (RatSWD), Berlin.

[U.S. Bureau of Labor Statistics \(2020\)](#), Measuring the effects of the coronavirus (COVID-19) pandemic using the current population survey, <https://www.bls.gov/bls/measuring-the-effects-of-the-coronavirus-covid-19-pandemic-using-the-current-population-survey.htm>, retrieved 28 October 2023.