8 Output
6th Appointment Period

RatSWD.
German Data Forum

Data Protection Guide

2nd fully revised edition

SPONSORED BY THE

Federal Ministry
of Education
and Research

German Data Forum (RatSWD)

Data Protection Guide

2nd fully revised edition

Revised by
Prof. Dr. Matthias Bäcker and Dr. Sebastian Golla

In collaboration with the
German Data Forum (RatSWD)

List of abbreviations

para.	Paragraph
Art.	Article
BDSG	German Federal Data Protection Act
BStatG	German Federal Statistics Act
BVerfG	German Federal Constitutional Court
DFG	Deutsche Forschungsgemeinschaft [German Research Foundation]
GDPR	EU General Data Protection Regulation
DSK	Conference of the Independent Federal and State Data Protection Supervisory Authorities
CJEU	European Court of Justice
GG	Basic Law of Germany
LDA	German State Data Protection and Data Access Officer
LDSG	German State Data Protection Act
LfDI	German State Data Protection and Freedom of Information Officer
RatSWD	German Data Forum
SGB	German Social Security Code
subpara.	subparagraph
ULD	Independent Data Protection Centre of the State of Schleswig-Holstein

Table of Contents

Preface

The German Data Forum (RatSWD) is an independent board comprised of empirical researchers and representatives of data producers. Established by the Federal Ministry of Education and Research in 2004, it seeks to make lasting improvements to the data infrastructure for empirical research in order to make it internationally more competitive. The German Data Forum (RatSWD) works at the intersection between research, data production, and data protection, bringing the expertise of numerous accredited research data centres together. One of its tasks is to provide advice to the academic and policy-making communities.

Within this mandate, the German Data Forum (RatSWD) created a first version of the guide on data protection during its fifth appointment period (2014–2017). In early 2020, the guide was fundamentally revised to take into account the new requirements for data protection resulting from the EU General Data Protection Regulations (GDPR). The guide's main focus is on the handling of personal data within the social, behavioural, and economic sciences.

1 Introduction

The purpose of this guide is to familiarise the interested reader – researchers in the social, behavioural, and economic sciences in particular – with the data protection regulations that are relevant to empirical research. It explains the basic principles of data protection according to the GDPR, identifies other relevant regulations, introduces the key concepts used in data protection in research, and elaborates on the obligations of data processors. The guide deals in detail with the aspects of data protection related to the collection, processing, publication, and storage of data.

The social, behavioural, and economic sciences are interested in and require data that are as comprehensive and accurate as possible, that can be used for researching a wide array of problems, and that allow results to be reproduced. These requirements often conflict with the principles of data protection, such as storage limitation, data minimisation, and purpose limitation. Handling data responsibly is crucial for the integrity of research and the willingness to entrust researchers with data. Therefore, compliance with data protection regulations is essential for empirical researchers.

The goal of this brochure is to present the major features of data protection law and to highlight its challenges. The multifaceted nature of research in the social, behavioural, and economic sciences makes it impossible to present all data protection issues in a sweeping or brief way. In most cases, individual issues will need to be addressed in more depth. Research projects that process large amounts of personal data regularly need to pay special attention to data protection measures to ensure compliance with the regulations.

The text includes references to additional literature and other resources. Particularly valuable in this regard are the websites of data protection supervisory agencies.[1] Since these agencies are the primary enforcers of data protection legislation, their information is highly relevant. Although they are nonbinding, the guidelines provided by the European Data Protection Board have become essential for interpreting the GDPR.[2] In addition, the European Data Protection Officer issued a provisional statement on data protection and scientific research in January 2020.[3]

Although the information from the regulators has been complemented by a large body of literature on the interpretation of data protection legislation, many important practical problems are still unresolved or at the very least require more in-depth examination. Court decisions, of which there are presently few, are unlikely to change this situation in the near future. This is because the legality of processing data, especially for research purposes, often depends on weighing up interests on a case-by-case basis and other considerations that are impossible to anticipate without knowledge of the concrete practical context.

In the current phase, the scientific community can help shape good data protection practices in research. Two concrete examples are the development of ethical standards for statements of consent (see section 4.5) and of criteria for the weighing of interests in the context of Art. 6 para. 1 letter f GDPR (see section 2.2) and Sec. 27 para. 1 sentence 1 BDSG (see section 4.1).

1 An overview of data regulation agencies, including their contact details and website addresses, can be found at https://www.bfdi.bund.de/DE/Infothek/Anschriften_Links/anschriften_links-node.html (in German).
2 Available at https://edpb.europa.eu/edpb_de.
3 Available at https://europa.eu/!WR73NV.

2 Basic principles of the GDPR

The EU General Data Protection Regulation (GDPR), which came into effect on 25 May 2018, marks a new era in data protection legislation. Although the GDPR maintains many of the established principles in such legislation (Albrecht and Jotzo 2017: Part 2 marginal no. 1), it also introduces some new obligations for data processors and, in particular, high penalties for violations. As a result, these regulations have garnered an unprecedented amount of attention for data protection, making the implementation of the regulations seem urgent.

A new aspect of the GDPR is that the rules on processing personal data primarily and directly follow European Union law. Regulations under national laws are merely supplementary (see section 3). This section will introduce the most important principles of data protection in accordance with the GDPR.

The GDPR privileges data processing for the purpose of academic research in several ways. They include exemptions from the principles of purpose limitation (Art. 5 para. 1 letter b half-sentence 2 GDPR, see section 2.4) and storage minimisation (Art. 5 para. 1 letter e half-sentence 2 GDPR, see section 2.6), options for restricting certain rights of data subjects (Art. 89 para. 2 GDPR, Art. 14 para. 5 letter b GDPR, Art. 17 para. 3 letter d GDPR, Art. 21 para. 6 GDPR, see section 5.3), and the option to establish the grounds for processing special categories of personal data for research purposes (Art. 9 para. 2 letter j GDPR, see section 4.1). On the reverse side of these privileges, the GDPR also makes "appropriate safeguards for the rights and freedoms of data subjects" obligatory when processing data for academic research (Art. 89 para. 1 sentence 1 GDPR, see section 5.1).

2.1 Scope

> *„This Regulation applies to the processing of personal data wholly or partly by automated means and to the processing other than by automated means of personal data which form part of a filing system or are intended to form part of a filing system.“*
>
> *Art. 2 para. 1 GDPR*

The GDPR applies when

- personal data are concerned,
- these data are processed,
- and such processing is to be carried out by automated means or within a file system.

Personal data, as defined under Art. 4 No. 1 GDPR, are “any information relating to an identified or identifiable natural person”. A natural person is deemed to be identifiable if “the person can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person.” Consequently, a name need not be recorded for data to be personal. For instance, it is sufficient if a person can be identified using an identification number in conjunction with additional information. It is not necessary that the content of the information is of any relevance.

The term “personal” is understood to be comprehensive and context-specific. Information is considered personal even if it requires some effort to connect it to a person. The European Court of Justice recently ruled that, among others, dynamic IP addresses from the perspective of the website operator (CJEU, judgement of 19-Oct-2016 – C-582/14) as well as examination answers and corrections (CJEU, judgement 20-Dec-2017 – C-434/16), are personal data.

Data are also considered personal when they are **pseudonymised**, so long as the person can be re-identified (see section 4.4). Only **anonymisation** makes data non-personal (see section 4.3).

According to Art. 4 No. 2 GDPR, **processing** means “any operation or set of operations which is performed on personal data or on sets of personal data, whether or not by automated means, such as collection, recording, organisation, structuring, storage, adaptation or alteration, retrieval, consultation, use, disclosure by transmission, dissemination or otherwise making available, alignment or combination, restriction, erasure or destruction.” In short, the term “processing” describes any form or manner of handling personal data. This also includes the deletion and the publication of data.

The fact that processing must be **automated or take place within a file system** excludes some data processing operations from the scope of the GDPR. Non-systematic note-taking of personal information or the verbal exchange of personal data are such cases. However, for public research institutions, this limitation of the scope of the GDPR is practically irrelevant. The German Federal Data Protection Act (BDSG) and the state data protection laws (LDSG) have declared that for public institutions – which include public universities and non-university research institutions – the GDPR also applies to non-automated processing and processing outside of file systems.

2.2 Principle of non-admissibility, lawfulness of processing

> *„Personal data must be processed lawfully, [...].“*
>
> *Art. 5 para. 1 letter a GDPR*

According to the GDPR, the processing of personal data always requires justification, i.e., processing is **prima facie prohibited**. In Art. 6 para 1, the GDPR stipulates in which exceptional cases processing is allowed. This list is exhaustive. Other statutory provisions, such as the federal or state-level data protection laws, however, may be more specific.

It is relevant for practice that there must be a reason for **every single data processing operation**. When, for example, data are collected, stored, analysed, and finally published, every single processing step must be reviewed for admissibility.

Both the data subject's **consent** (Art. 6 para. 1 letter a GDPR, see section 4.6) and **statutory authorisation** may constitute a legal basis for the processing of data. Relevant statutory authorisation for researchers are work which is in the public interest (Art. 6 para. 1 letter e GDPR) and the authorisation to process data after a general weighing of interests (Art. 6 para. 1 letter f GDPR). The GDPR does not grant special authorisation to process data specifically for research purposes. Art. 89 GDPR does not contain any such authorisation for processing for research purposes but allows member states to issue special regulations (with regard to the rights of data subjects, see section 5.3) and requires them to provide special protections or guarantees with respect to processing data for research purposes. The national law contains special research provisions under Sec. 27 para. 1 BDSG and in the respective provisions of the state data protection laws (LDSG). However, these authorisations do not concern the processing of all personal data (according to Art. 6 GDPR), but only of specific categories of such data (according to Art. 9 GDPR, see section 4.1).

Which authorisations are applicable depends on how a research institution is organised. **Public research institutions**, as a matter of principle, cannot rely on authorisation after a general weighing of interests. They must rely on the performance of their duties, i.e. the performance of tasks in the public interest. The legal basis for processing personal data for research purposes is Art. 6 para. 1 letter e GDPR in addition to national regulations that specify the role of research. For example, public universities (universities and universities of applied sciences) and non-university research institutions are expressly assigned a research function by the state university acts of the different Länder. The authorisations granted to universities and other public research institutions for processing data for research purposes are thus primarily limited by their research tasks. Since the processing of data must be necessary to fulfil their research tasks, secondly, a dedicated weighing of interests must be carried out. This weighing of interests can be based on the criteria listed below.

Private research institutions,[4] however, generally cannot invoke the public interest. For these institutions, the legal basis for processing personal data for research purposes is generally Art. 6 para. 1 letter f GDPR. This provision allows for personal data to be processed insofar as it is necessary for pursuing the legitimate interests of those responsible or of a third party. This provision applies unless such interests are overridden by the data subject's interests or fundamental rights and freedoms which require personal data be protected.

Academic research is a legitimate interest in itself and must be **weighed** against the interests of the data subjects. Considerations on the part of the data subject that must be taken into account include the sensitivity of the data (see section 4.1), whether the data originate from publicly accessible information or from other sources, and whether the data subject anticipates that the data would

4 The differentiation between private and public research institutions does not depend on the organisational form but what kind of tasks to the respective institution has been assigned. For example, a (g)GmbH [limited company (with public utility)] that is controlled by a public agency and that has been assigned a research task must also be considered a public research institution. And when a public research institution solicits funds from private entities (e.g., a company), any research project funded in this way will still be part of a public task. A wholly different situation arises when individual researchers carry out research projects for private entities within the framework of secondary employment.

be processed (further). The relationship between the researcher processing the data and the data subject must also be considered. For the research interests, it is relevant to consider how important the processing of specific personal data is for the realisation of a research project.

2.3 Transparency

> *„Personal data must be processed [...] in a transparent manner in relation to the data subject.“*
>
> *Art. 5 para. 1 letter a GDPR*

A fundamental concern of the data protection legislation is to ensure that data subjects do **not lose track** of how their personal data are handled, despite the complexity of modern data processing. The principle of transparency therefore is included in several provisions of the GDPR. For instance, the requirement that a data subject's consent must make the purposes of data processing clear (see section 4.5 for more detail) both establishes clarity and fulfils the obligation to inform data subjects proactively (see section 5.2 for more details). The data subject's right to access (Art. 15 GDPR) also promotes transparency.

2.4 Purpose limitation

> *„Personal data must be collected for specified, explicit and legitimate purposes and must not be further processed in a manner that is incompatible with those purposes.“*
>
> *Art. 5 para. 1 letter b GDPR*

The principle of purpose limitation aims to make the handling of personal data **transparent and controllable**. Firstly, it requires specifying the purpose of data processing as precisely as possible when the data are initially collected, for example in a statement of consent or in the research design. Moreover, any further processing, as a matter of principle, is limited to the stated purpose. Changing the purpose is possible but requires authorisation.

However, the GDPR's principle of purpose limitation has been **relaxed specifically for research**. In this way, the law takes into account the fact that research goals and challenges often cannot be determined conclusively in advance. Art. 5 para. 1 letter b half-sentence 2 GDPR states that any further processing, for the purpose of scientific or historic research, of data originally collected for other purposes is not incompatible with the original purpose. This provision creates the fiction of processing compatibility: even if the research purposes, using conventional assessment criteria, are not compatible with the original processing purpose, the GDPR legally provides such compatibility. When processing the data further, the processor can therefore rely on the same legal basis they relied on when first processing the data.

Example 1: Files from trade supervisory authorities are to be analysed for a research project examining the susceptibility to corruption in certain industries. Even though the personal data were originally collected by the authorities to fulfil official duties, their further processing is generally admissible despite the fact that the research purposes prima facie differ substantially from the authorities' administrative tasks. However, as always, processing the data additionally requires a respective authorisation as set out under Art. 6 and, if applicable, Art. 9 GDPR.

Example 2: Within the framework of a long-term research project on changes in worldviews, data were collected that later proved useful for a criminological analysis of the likelihood of subsets of respondents to become offenders. The criminological analysis represents a new processing purpose because it is a new research project with a new scientific objective. However, the change in purpose again is privileged, and the processing of the data again requires authorisation, which must be reviewed separately for the new research purpose.

2.5 Data minimisation

> *„Personal data must be adequate, relevant and limited to what is necessary in relation to the purposes for which they are processed."*
>
> *Art. 5 para 1 letter c GDPR*

The principle of data minimisation requires that dealing with personal data be limited to the bare **minimum** needed to fulfil the intended purpose. Among others, this principle is reflected in the authorisation to process data as necessary to fulfil official duties or to protect legitimate interests (Art. 6 para. 1 letters e and f GDPR).

Example: In a study examining consumer preference variation of same-sex partners, information about religious views may not be collected if that information is not relevant to the stated research purpose.

However, the principle of data minimisation should not be interpreted so narrowly as to limit the research purpose strictly to a specific research question. The amount of data needed may also be determined by an overarching research project with a broader purpose. Nevertheless, the project must still adhere to the data protection obligation to define the purpose for processing data as precisely as possible. It is not possible to circumvent data protection regulations by stating an overly broad purpose (an extreme example: "data processing for social scientific research").

The principle of data minimisation can conflict with research projects involving large amounts of data (**big data**). It may be argued, however, that a mass analysis of personal data is necessary to carry out such research. In these cases, it is especially important to describe the purpose of the research project as precisely as possible.

Example: The mass analysis of personal data originating from archives or social networks may also be considered to be within the necessary scope if it is scientifically plausible that such an analysis helps to identify correlations and arrive at insights that would not be possible with smaller amounts of data. In contrast, analysing correlations with no specific goal would not be permissible if such an analysis were not based on a research objective that was formulated clearly in advance.

2.6 Storage limitation

> *„Personal data must be kept in a form which permits identification of data subjects for no longer than is necessary for the purposes for which the personal data are processed.“*
>
> *Art. 5 para 1 letter e GDPR*

The principle of storage limitation is related to the principle of data minimisation. It requires that personal data be deleted if they are no longer necessary for the intended processing purposes. Therefore, it imposes a **time limit** on the processing of personal data.

For research purposes, however, there is an exemption from the principle of storage minimisation, as set out under Art. 5 para. 1 letter e half-sentence 2 GDPR. This rule allows a longer storage of personal data if, in addition to their initial purpose, there are research purposes that require prolonged storage times, e.g., for secondary analyses. However, this is also subject to the general obligation on academic research that personal data must be anonymised or pseudonymised whenever possible (see sections 4.2 and 4.3).

2.7 Integrity and confidentiality

> *„Personal data must be processed in a manner that ensures appropriate security of the personal data, including protection against unauthorised or unlawful processing and against accidental loss, destruction or damage, using appropriate technical or organisational measures.“*
>
> *Art. 5 para 1 letter f GDPR*

The principle of integrity and confidentiality primarily addresses the **technical security** of the data. Art. 32 GDPR contains more general requirements and Art. 89 GDPR contains other research-specific requirements. They are elaborated on in detail below (see section 5.2)..

2.8 Technical and organisational measures for implementing these principles

According to Art. 25 para. 1 GDPR, the controller “shall [...] implement appropriate technical and organisational measures [...], which are designed to implement data-protection principles, such as data minimisation, in an effective manner and to integrate the necessary safeguards into the processing in order to meet the requirements of this Regulation and protect the rights of data subjects.”

This is the basis for the obligation to observe the aforementioned principles of data protection law already in the technical and organisational design of data processing procedures (**Privacy by Design**). This way, the GDPR responds to the fact that the law only has limited effectiveness on individual data processing operations if the existing technical and organisational design poses practical constraints.

The technical principles of data avoidance and data minimisation are based on the same rationale: Already in the development of data processing technologies and their default settings, the goal should be to minimise the amount of personal data being processed.

Example: A research app for mobile devices may not have the technical capability to access the device’s camera and location tracking technology if such information is not necessary for the research purpose.

3 Applicable regulations besides the GDPR

While the GDPR may be the most prominent regulation for (research) data protection, it is not the only relevant one. The constitutional framework of the EU and Germany must be observed as well (see section 3.1).

In some cases, national regulations supplement the GDPR regarding the processing of personal data for research purposes (see sections 3.2 - 3.4). However, this is only possible to the extent that the GDPR, which takes priority, expressly allows leeway for such supplementary regulations.

3.1 Framework of EU and constitutional law

> *„Everyone has the right to the protection of personal data concerning him or her.“*
>
> *Article 8 para. 1 of the Charter of Fundamental Rights of the European Union*

At the European level, the **constitutional basis for data protection legislation** is the right to protection of personal data as set out in Article 8 of the Charter of Fundamental Rights of the European Union. According to Art. 8 sec. 2 sentence 1 of the Charter, personal data can only be processed in good faith, for specified purposes, and on the basis of the consent of the person concerned or on some other legitimate basis laid down by law. Under German constitutional law, it is the right to informational self-determination that is relevant for data protection. This is not made explicit in the German Basic Law but results from Art. 2 para. 1 (right to free development of one's personality) together with Art. 1 para. 1 (protection of human dignity) of the German Basic Law. While the GDPR of the European Union has been governing data protection as a legal act since May 2018, it is the provisions of the Charter of Fundamental Rights of the European Union that prevail.

> *„The arts and scientific research shall be free of constraint. Academic freedom shall be respected.“*
>
> *Article 13 of the Charter of Fundamental Rights of the European Union*

Scientific research is also protected by the constitution. Art. 3 para. 3 subpara. 1 sentence 3 of the Treaty on European Union stipulates that the promotion of scientific advancement is a goal of the EU on equal footing with sustainable economic growth and social progress. In this vein, Art. 179 para. 1 of the Treaty on the Functioning of the European Union stipulates the creation of a European research area. The processing of personal data for research purposes is a fundamental requirement for attaining this goal. In addition, Art. 11 (Freedom of expression and information) and Art. 13 (Freedom of the arts and sciences) of the Charter of Fundamental Rights of the European Union protect individual academic freedom. The freedom of the arts and sciences protects any activity aimed at gaining knowledge through a methodical, systematic, and verifiable manner, as well as the publication of scientific findings. Both natural persons who engage in research on their own behalf and legal entities, particularly universities and non-university research institutions, can rely on freedom of research. In Germany, freedom of research is guaranteed by Art. 5 para. 3 sentence 1 in the German Basic Law.

3.2 German Federal Data Protection Act (BDSG) and State Data Protection Acts (LDSGs)

The GDPR includes clauses that allow national legislators to make supplementary regulations or restrictions on a number of specified issues. National legislators may, for example, issue regulations on data processing to fulfil statutory obligations and public tasks (Art. 6 para. 2 and 3 GDPR), on the freedom of expression and information (Art. 85 GDPR), on restrictions of the rights of data subjects (Art. 23 GDPR), or on scientific research (e.g., see Art. 9 para. 2 letter j GDPR and Art. 89 para. 2 GDPR). The data protection law of both the federal and state-level in Germany features such regulations. As a rule, the German Federal Data Protection Act (BDSG) applies to public institutions of the federal government[5] and for non-public institutions (Sec. 1 para. 1 BDSG). The state data protection acts (LDSG) apply to public institutions of the federated states (*Länder*).

Both the BDSG and the LDSGs contain **special regulations for data processing for research purposes**. As these regulations are very similar in substance, Sec. 27 BDSG can serve as an example of corresponding regulations that exist in the state data protection acts. Sec. 27 para. 1 BDSG allows processing of special personal data (see section 4.1) if the research interests substantially outweigh those of the data subjects. Sec. 27 para. 4 BDSG provides for additional stricter regulations for the publishing of data (see section 8). Sec. 27 para. 2 BDSG makes exceptions to certain rights of data subjects granted by GDPR if data are processed for research purposes (see section 5.3). Sec. 27 para. 3 BDSG contains a special anonymisation requirement if data are processed for research purposes (see section 4.2).

3.3 Special regulations under federal and state law

In addition to the BDSG and the LDSG, there are a number of special regulations under German federal law (BDSG) and laws of the federated states (LDSG) that only apply to the processing of data for research purposes in certain research areas or when using specific research methods.

In **German federal law**, such special regulations can be found, for example, in the Medicinal Products Act, the Genetic Diagnostics Act, the Transplantation Act, or the Stasi Records Act. The provisions of the Federal Act of Registration must be observed when the research includes registration data (see Von Lewinski 2017: 1ff.).

The protection of **social data**[6] is comprehensively regulated in Sec. 67 et seq. SGB X. This protection primarily follows the principles that apply to the protection of other personal data but is subject to stricter requirements in specific cases (see Kipker and Pollmann 2019: 718ff.). Sec. 67a et seq. SGB X regulate special authorisations for processing social security data. If the data are not collected directly from the data subjects, the legitimacy of their collection largely depends on the weighing of interests (cf. Sec. 67a para. 2 SGB X).

5 Federal law additionally applies to the public institutions of the states if data protection is not regulated by state law, and insofar as the institutions implement federal law or act as bodies of the judicature, except when this is an administrative matter.

6 Social security data are personal data that are processed by an agency mentioned under Sec. 35 SGB I fulfilling its tasks under SGB [Social Security Code] I.

Sec. 75 para. 1 sentence 1 No. 1 SGB X authorises the **transfer** of social data for research purposes if these data are necessary for research projects on welfare benefits or academic research on the labour market and occupation. Here, too, the interests of the data subjects must be considered on a case-by-case basis. However, sentence 2 of the provision generally obligates the transmitting agency[7] to obtain consent if this is reasonably feasible and if more than basic information, such as names and addresses, are to be passed on. Sec. 75 para. 4 SGB X additionally makes the transfer of these data contingent on official authorisation.

Sec. 67b para. 2 and para. 3 SGB X also contains **special regulations for consent** to the processing of social security data. According to Sec. 67b para. 3 sentence 1, consent to the processing of personal data for research purposes can be given for certain projects or for certain areas of academic research. This regulation therefore allows the processing purposes to be determined broadly (see "broad consent" in section 4.5 for more detail).

Some **state laws** for specific areas contain detailed provisions for the processing of personal data for research purposes. Examples include school acts, state hospital acts, and state archive acts. The data of the federal and state statistical offices are especially protected under statistics acts. A special secrecy obligation applies to information that is provided for statistical purposes (acc. to Sec. 16 BStatG and the respective provisions in the state laws). According to Sec. 16 para. 6 BStatG, data can only be transferred or made accessible for research purposes if they are de facto or formally anonymised. The purpose for processing data must be strictly limited and the data must be deleted after the research project is completed (Sec. 16 para. 8 BStatG).

7 For example, the research data centre of the Federal Employment Agency (BA) at the Institute for Employment Research (FDZ BA at IAB) or of the German pension insurance (FDZ-RV).

4 Key concepts of data protection in research

This section introduces the key data protection concepts that are important for processing personal data for research purposes.

4.1 Special categories of personal data

> *„Processing of personal data revealing racial or ethnic origin, political opinions, religious or philosophical beliefs, or trade union membership, and the processing of genetic data, biometric data for the purpose of uniquely identifying a natural person, data concerning health or data concerning a natural person's sex life or sexual orientation shall be prohibited.“*
>
> *Art. 9 para. 1 GDPR*

The GDPR provides **additional protections** for several types of personal data. The data mentioned in Art. 9 para. 1 GDPR are considered to be especially sensitive. This regulation is also relevant for research in the social sciences; information about political beliefs, religious affiliation, or health data from health insurance companies could be relevant for social milieu studies, for example.

For the processing of such data to be admissible, the requirements set out under Art. 9 para. 2 GDPR must be fulfilled **in addition** to the general requirements of Art. 6. This provision provides several options for justifying the processing of these data. One option is to acquire consent that expressly refers to the processing of special personal data (Art. 9 para. 2 letter a GDPR, see section 4.5).

Another possible justification for processing certain **publicly available data** arises from Art. 9 para 2 letter e GDPR. This provision allows for the processing of special personal data that evidently were made public by the data subject. In these cases, the data lack the need for special protection. This norm allows, for example, the analysis of posts published in openly accessible parts of social networking sites for research purposes, insofar as this only concerns information about the authors of the posts themselves (Golla/Hofmann/Bäcker 2018: 92). Note that such an authorisation cannot be used if a person other than the data subject published the data.

Even more relevant for researchers, however, is the provision in Art. 9 para. 2 letter j GDPR. It allows member states to issue their own regulations of processing personal data for **research purposes**. The German legislators used this option in Sec. 27 BDSG (see section 3.2), and similar regulations can be found in the data protection acts of the German states (LDSG). The following elaboration on Sec. 27 BDSG can therefore be largely applied to the regulations in the state laws.

Sec. 27 para. 1 sentence 1 BDSG stipulates that three requirements must be met for processing of special categories of personal data for research purposes to be admissible:

1. The **purpose must be for academic research**. To fulfil this requirement, there must be a concrete research project that meets scientific standards in terms of design and content.
2. Processing the data must be **necessary to carry out this project**. This means that the project cannot be carried out without the processing of these concrete personal data. In this regard, the possibility of pseudonymising and anonymising the data must also be considered. Processing personal data would not necessary if the project could also be carried out using pseudonymised or anonymised data.

3. Interests must **be weighed on a case-by-case basis**. The research interest must significantly outweigh the interests of the data subjects. By using the single word "significantly", Sec. 27 para. 1 sentence 1 BDSG imposes higher demands on data processing than the general weighing of interests under Art. 6 para. 1 letter e or f GDPR. In order to reliably assess whether the research interest significantly outweighs other interests, practical guidelines for each research area will need to be developed. Researchers themselves, together with the data protection officers of their institutes and, where appropriate, representatives from the supervisory authorities can (and should) create such guidelines. This gives them the opportunity to establish their own standards of data protection-friendly research.

Weighing the interests will usually be the **focal point in a review** of whether data processing is legitimate in light of Sec. 27 para. 1 sentence 1 BDSG. Such a review will need to assess both the research interests and the interests of the data subjects. For the research interest, a key factor is how important processing special categories of personal data is for realising a research project. To make a general assessment of the significance of a project, on the other hand, is difficult and would conflict with the freedom of research. An important question for assessing the interests of data subjects is the degree of intensity with which Art 9 para. 1 GDPR is concerned, i.e., to what extent the characteristics mentioned in the regulation can be determined from the processed data (for example, datasets may directly state or only contain indications of a person's religious beliefs.)

Example: An essential part of a research project involves information about political opinions. However, this information is only processed in a very limited scope for each data subject. This indicates that the research interest significantly outweighs those of the data subjects.

Furthermore, the controller, who is responsible for data processing, must take appropriate and specific measures to protect the interests of the data subjects according to Sec. 27 para. 1 sentence 2 BDSG. These measures are largely identical to the suitable safeguards stipulated by Art. 89 GDPR (see section 5.1) and the technical and organisational measures required by Art. 32 GDPR (see section 5.2).

The data protection supervisory agencies provide further information on the requirements for handling special categories of personal data (DSK 2018h).

4.2 Anonymisation

Anonymisation means the **erasure of any personal references** in datasets. If data are anonymised, the GDPR no longer applies to their processing (see section 2.1). In contrast to the former version of the BDSG, the GDPR does not define the term anonymisation.[8] although it does refer to the concept of anonymisation. No. 26 of the recitals of the GDPR suggests that de facto anonymisation is necessary for the erasure of personal references, and that this is achieved if personal references can only be restored with disproportionate effort.

The data protection legislation contains special **requirements for anonymisation in research**. Art. 89 para. 1 sentence 4 GDPR stipulates that, if possible, any further processing of data for research purposes must be carried out in such a way as to ensure that data subjects cannot or can no longer be identified. This must be interpreted as a requirement to pseudonymise or anonymise data processed for research purposes, insofar as this is feasible (Golla 2019: 658f.). Sec. 27 para. 3 sentence 1 BDSG also contains an anonymisation requirement for special categories of personal data (see section 4.1).

8 According to Sec. 3 para. 6 of the former version of BDSG [Federal Data Protection Act], anonymisation means the "alteration of personal data so that information concerning personal or material circumstances cannot be attributed to an identified or identifiable natural person or that such attribution would require a disproportionate amount of time, expense, and effort."

These anonymisation requirements can be considered expressions of the principles of **data minimisation** (see section 2.5) and **storage limitation** (see section 2.7) (Johannes and Richter 2017: 302). However, in many cases, it may prove difficult to achieve the intended research results using anonymised data, especially in the social sciences (Arning/Forgó/Krügel 2006: 701). Reproducibility of research results often requires that data not be anonymised because such an operation is irreversible. As a result, often only pseudonymisation, the milder alternative to anonymisation, will be considered.

From a **technical point of view**, methods that are suitable for anonymisation modify or diminish the content of the data, for example by deleting identifiers, aggregating attributes, or masking data. Effective anonymisation is not a trivial operation because it must be ensured that data subjects cannot be re-identified or the data de-anonymised using outside information or new technological methods. A single anonymisation therefore cannot be considered to be permanently effective. Taking into account the current state of technological advancement, the effectiveness of anonymisation must be reviewed at regular intervals.

Example: The data of a survey are stored using codes that are generated by personal information about the participants. These codes have four digits, consisting of the first letter each of the first names of the participant, their father and their mother, and of the last digit of their year of birth (e.g. MAS8 for Monika, Anton, Sandra, born in 1968). When matched with a database, this information might be used to identify survey participants. Consequently, this is not a safe method of anonymisation.

A viable option would be the following: A index number or a random number is assigned to the collected data. So long as a direct attribution to individual test persons might be necessary for the research purpose (e.g., because additional data is to be collected later), an attribution table is kept; during this period, the data remain pseudonymised. Once no direct attribution of the data is necessary anymore, the table is deleted. This way, the data are anonymised, provided the content of the data cannot be used to for inferences about the identity of the data subjects. However, the data are not anonymised if such re-identification remains possible.

4.3 Pseudonymisation

> *„'Pseudonymisation' means the processing of personal data in such a manner that the personal data can no longer be attributed to a specific data subject without the use of additional information, provided that such additional information is kept separately and is subject to technical and organisational measures to ensure that the personal data are not attributed to an identified or identifiable natural person.“*
>
> *Art. 4 No. 5 GDPR*

In contrast to anonymisation, pseudonymisation does not completely erase personal references in the data. Identifying information is replaced by tokens, which are stored separately. Data protection regulations remain applicable to pseudonymised data.

Example: The names of the data subjects in a dataset are replaced by code numbers. The persons cannot be identified by other attributes in the dataset. The list that identifies which code number was allocated to which person is deposited with a trustee.

If pseudonymised data suffice to attain the research purpose, **pseudonymisation is required** according to Art. 89 para. 1 sentence 3 GDPR. A similar requirement is contained in Sec. 27 para. 3 sentence 2 BDSG. These requirements also highlight the principles of data minimisation (see section 2.5) and storage limitation (see section 2.7). Pseudonymisation is often a more realistic alternative to processing personal data than anonymisation because it allows researchers to maintain the traceability of their databases.

4.4 Responsibility

> *„For the purposes of this Regulation, 'controller' means the natural or legal person, public authority, agency or other body which, alone or jointly with others, determines the purposes and means of the processing of personal data."*
>
> *Art. 4 No. 7 GDPR*

Selecting a controller determines who must comply with the data protection regulation and against who data subjects can exercise their rights. In research, it is not easy to determine the controller responsible for data processing. According to Art. 4 No. 7 GDPR, the relevant criterion for being the controller is **who defines the purpose and the means of data processing**.

Consequently, the decisive factor is how much an individual influences the processing of data. Criteria that help determine who is ultimately responsible include the authority to issue instructions, being the point of contact for the data subjects, the type of role, professional expertise, and the latitude to act in individual cases. The controller may utilize their own assistants (particularly employees and civil servants) or external processors (entrusted with data processing tasks under service contracts) for processing the data. They might enjoy **considerable latitude**.

As long as these criteria are met, it can be assumed that even larger research institutions (such as universities) have **central responsibility**. Although the individual researchers have a great deal of leeway, they act as representatives of their institution and fulfil the tasks assigned to the institution.

According to Art. 26 GDPR, multiple entities may be **jointly responsible** for processing data. Joint responsibility may be considered,

- if institutions cooperate with one another within the framework of larger research projects,
- if survey programmes entrust data collection to survey institutes that have some leeway in planning and conducting the survey, or
- if researchers store their research data in repositories and make them available for potential secondary analyses.

4.5 Consent

> *„Within the meaning of this regulation, 'consent' of the data subject means any freely given, specific, informed and unambiguous indication of the data subject's wishes by which he or she, by a statement or by a clear affirmative action, signifies agreement to the processing of personal data relating to him or her.“*
>
> *Art. 4 No. 11 GDPR*

In research, consent is an important legal basis for processing personal data. It can justify both the processing of special categories of personal data (Art. 9 para. 2 letter a GDPR, see section 4.1) and other personal data (Art. 6 para. 1 letter a GDPR). For reasons of legal security, consent is often preferred to the authorisations for data processing provided by law. While the potential of such statutory authorisation should not be underestimated (see section 6.3), consent may also be advisable for **ethical reasons** (following the principle of "informed consent").

Art. 4 No. 11 GDPR requires that consent be **given freely**. That means there must not be any (perceived) coercion or a significant power imbalance between the processor and the data subject that could influence the act of giving consent.

Consent can only be freely if it is **informed** consent. In research, it is important to provide a generally comprehensible description of research concepts and objectives. In some cases, providing the data subject with detailed information can conflict with the research method – for example, psychological experiments that involve targeted manipulations. But such cases are, at the very least, still required to disclose the goal of the research project to the data subject. Manipulation in an experiment may be admissible as long as it is not used to mislead the data subjects into disclosing personal data. In other cases, the requirement for informed consent can be an insurmountable obstacle, for example if the research project cannot be explained or made comprehensible to the data subjects due to cultural or intellectual reasons.

Example: A professor at a university calls on the students in his course to take part in an experiment for behavioural research and asks them for their consent to process their data. He only intends to inform them afterwards about how their data was handled and the objective of the experiment. In this case, it is doubtful that consent was freely given because the relationship between the professor and his students is hierarchic and the latter may feel obligated to participate. In any case, consent is not informed since information must be given before data are processed.

In addition, consent must generally be given for a **specific case**. This is challenging in research because research precise objectives and questions cannot always be determined in advance. Therefore, it is important for research that consent can also be given for a broadly formulated purpose (**"broad consent"**). This does not weaken the purpose limitation requirement (see section 2.4 above). However, it does affect the requirement to limit consent to only one or more concretely defined purposes.

According to recital No. 33 sentence 2 GDPR, data subjects "should be allowed to give their consent to certain areas of scientific research when in keeping with recognised ethical standards for scientific research." (emphasis added). Compliance with the accepted **ethical standards** is supposed to minimise the risks of broad consent. These standards are not legally binding but do at least indicate how to make the legal requirements concrete. In practise, ethical standards develop from codes, guidelines, and declarations. However, standards that apply to processing personal data are not equally established in all areas of research. Through collaboration between researchers, ethics commissions, and data protection officers these standards could be developed from practical cases.

The German Data Forum (RatSWD) dedicated an output paper (RatSWD 2017: 16 ff.) with more information on ethical principles and standards in research in the social and economic sciences.

If ethical standards have been established and are adhered to, broad consent is an important basis for processing data for research purposes. Note, however, that this consent may be freely revoked in the future.

Example: For specific data or collections of data which are limited temporally, spatially, or regarding their content, consent may be given for behavioural sciences (or a sub-discipline) without initially determining the research purpose.

An implied requirement for consent is that the person giving it must be **capable of understanding** the significance of doing so. Minors can also effectively consent to the processing of their data, but they must be able to understand the consequences of their consent. There is no fixed age limit, and the capacity to understand must be determined on an individual basis. Generally, this capacity should be present from age 14 to 16 onwards. Regardless of this, the consent of a parent or guardian may be necessary.

Written consent is not required under the GDPR, although it is still advisable to use the text form for documentation and as proof. As opposed to the written form [as set out under Sec. 126 BGB], the text form does not require that a document is signed. It only requires a statement to be on a permanent data medium, i.e., on paper or in a digital format.

Note that consent can be **withdrawn** by the data subject at any time (Art. 7 para. 3 GDPR). Such withdrawal is effective for future cases only. Any past processing of data that was admissible at the time remains so. From the time of the withdrawal, however, the data may no longer be processed on the basis of the former consent.

The revocability of consent can have adverse consequences, especially for long-term research projects (e.g., panel surveys). In such cases, it is worth considering whether consent can be grounded on other legal bases from the onset. If this is impossible or not desired, the processing of data could still be based on a different legal basis after the withdrawal of consent. The participants should be informed about this when consent is obtained in order to prevent situations in which recourse to other legal bases would be denied based on the principle of good faith.

The further processing of formerly personal data that have been anonymised remains unaffected by a withdrawal of consent. Since such data are not personal data, their processing falls out of the scope of data protection regulations. A mere pseudonymisation, however, is not sufficient to allow further processing because pseudonymised data are still deemed personal data.

More detailed information on the requirements for effective consent can be obtained from the data protection supervisory agency (DSK 2019a).

5 Obligations of the data processor

In the following section, some of the key obligations of data processors in research will be explored.

5.1 Safeguards for the rights and freedoms of data subjects

As briefly described above, the GDPR privileges processing personal data for research purposes in several ways. In return, however, the GDPR also imposes an obligation on data processors that carry out research to provide "**appropriate safeguards** [...] for the rights and freedoms of data subjects" as stipulated by Art. 89 para. 1 sentence 1 GDPR (emphasis added).

According to Art. 89 para. 1 sentence 2 GDPR, such safeguards must include technical and organisational measures that ensure compliance with the principle of data minimisation in particular (see section 2.5). They include, among others, defining storage time limits, access options, and the requirement to anonymise (see section 4.2) or pseudonymise the data (4.3). Which **technical and organisational measures** are required strongly depends on the concrete research project. Among other things, considerations must include what personal data are processed and who is involved in the processing (Johannes and Richter 2017: 302).

Example: One organisational measure could be the contractual obligation requiring compliance with data protection regulations (non-disclosure agreement), as practised by the majority of the research data centres, for example. The choice of access path (download, guest researcher workstations, remote access) and its concrete design is particularly significant. One technical measure could be, for example, the encryption of data prior to their transfer.

In addition to Art. 89 GDPR, controllers are as required by Sec. 27 para. 1 sentence 2 BDSG to take appropriate and specific measures to protect the interests of data subjects.[9] Sec. 27 para. 3 sentence 1 BDSG additionally provides an **anonymisation requirement** for special categories of personal data that are processed for research purposes. However, this requirement already arises from the anonymisation requirement as set out under Art. 89 para. 1 sentence 4 GDPR (see section 4.2 above).

5.2 Data security/technical and organisational measures

In line with previous legislation, Art. 32 GDPR stipulates that data processing operations must be made secure by means of technical and organisational measures. Besides **physical protections** such as lockers and fire prevention measures, this also includes **IT security measures** such as encryption and the use of secure services for the storage and transfer of data. As a rule, technical measures must be in line with the state of the art.

Regarding data processing services, it is generally advisable for researchers to use solutions that are **provided by their research institution** (for the storage of research data in repositories, see section 9.1 below). External services are not recommended, especially if their data processing operations take place outside the European Union. This is true for services like Dropbox (cloud storage), Onedrive (cloud storage), and WhatsApp (messenger). All these services are offered by US providers whose handling of user data is only documented to a limited extent, making data protection rights claims against them particularly difficult.

9 See Sec. 22 para. 2 sentence 2 BDSG.

5.3 Protecting the rights of data subjects

Data processors also need to protect the rights of data subjects that GDPR has expanded, compared to previous legislation. Here, the data subject's right to information (Art. 13–15 GDPR), which ensures transparency, and their **right to influence** the processing of the data (Art. 15–21 GDPR) need to be distinguished. Among the **right to information**, one must additionally differentiate between a controller's obligation to proactively provide information (Art. 13 and 14 GDPR) and the right of access (Art. 15 GDPR), which requires data subjects to request the information.

There are two different provisions for the **obligation to provide information**. Art. 13 GDPR regulates the obligation to provide information when personal data are collected from data subjects (for example, by interviewing data subjects or observing their behaviour). In contrast, Art. 14 GDPR applies to the collection of data without active participation of the data subject (for example, interviews with third parties or research in publicly accessible sources). In short, determining the difference between the cases is a complex task.

Example: If location data are collected using an app on a mobile device that was installed for research purposes, the data subjects themselves serve as a data source and Art. 13 GDPR applies with respect to the obligation to provide information. If content or metadata from public social networks (e.g., Twitter) are processed, the data are not collected directly from the data subjects and Art. 14 GDPR applies.

Both regulations are largely identical with regard to the information that controllers must provide to the data subjects. One difference, however, is at what point in time the information is provided: The obligation to provide information under Art. 13 GDPR must be fulfilled at the beginning of the data collection process, while Art. 14 GDPR allows that information be provided later (within one month at the latest). Another difference is that Art. 14 GDPR contains some additional **exceptions** from the obligation to provide information, in particular for research purposes, when data are collected without the involvement of the data subjects, as set out under Art. 14 para. 5 letter b 2nd half-sentence GDPR. According to this provision, information need not be provided if doing so would make the objectives of the scientific or historical research impossible to realise. This exception is of high practical relevance because it can eliminate the time-consuming obligation for researchers to proactively inform the data subjects. However, this exception applies only if the data are not collected directly from the data subjects.

Example: For large-scale quantitative analyses of openly accessible personal data, such as from social networks, it is often practically impossible to inform all data subjects. Especially when the data are not particularly sensitive, the effort that would be required to provide information to the data subjects would be disproportionate to the value of that information to the data subjects. Therefore, the obligation to provide information as stipulated under Art. 14 GDPR would normally not apply.

The German data protection supervisory authorities provide more detailed information on how to fulfil the obligation to provide information (e.g. DSK 2018e, ULD 2019a).

In addition, data subjects also have a **right of access** according to Art. 15 GDPR, which includes the right to be sent a copy of the data stored about them. The GDPR does not include an exception relevant for academic research. However, one exception can be found in Sec. 27 para. 2 sentence 2 BDSG. This provision restricts the right of access to data which is necessary for purposes of academic research if the effort required to provide the information would be disproportionately high. Influence rights of the data subjects that are relevant to research include the right to rectification of incorrect data (Art. 16 GDPR), to erasure of illegally processed data or data that are no longer required (Art. 17 GDPR – the so-called right to be forgotten), and to limit the processing of data, particularly in cases of doubt (Art. 18 GDPR). Under special circumstances, data subjects can also object to the processing of their data (Art. 21 GDPR).

The data protection supervisory authorities provide guidance on how to implement the right of access (DSK 2018d) and the right to erasure (DSK 2018f).

The GDPR and the national data protection regulations contain exceptions from these influence rights to protect academic freedom. Generally, these exceptions are wide-ranging and usually allow the restriction of data subjects' rights if realising these rights would make conducting the research impossible.

With regards to the **right to be forgotten**, provided under Art. 17 GDPR, Para. 3 letter d of the provision stipulates an exception. It applies if processing the data is necessary for research purposes insofar the right is likely to render impossible or seriously impede the achievement of the objectives.

Art. 21 para. 6 GDPR privileges academic research purposes with respect to the **right to object**. Scientific data processing may be continued despite objection if such processing is necessary to fulfil a task in the public interest. The right to object according to Sec. 27 para. 2 sentence 1 BDSG as well as the right of access (Art. 15 GDPR), the right to rectification (Art. 16 GDPR), and the right to limit data processing (Art. 18 GDPR) are restricted if they render impossible or seriously impede the research objectives. All such assessments are to be documented.

5.4 Documenting processing activities

According to the GDPR, controllers are obligated to keep records of their processing activities (Art. 30 GDPR). Such records serve as an inventory of activities involving personal data. They must be made available to the supervisory authorities on request (Art. 30 para. 4 GDPR). The obligation to keep records of processing activities can be understood as an expression of the controller's general accountability (Art. 5 para. 2 GDPR). The controller must be able to prove compliance with the principles of data protection regulation.

The data protection supervisory authorities provide more detailed information (DSK 2018, DSK 2018a) and templates (DSK 2018b) for keeping records of processing activities.

5.5 Appointment of a data protection officer

Art. 37 GDPR requires data processors to appoint data protection officers. This obligation generally applies to all **public entities** and to **private entities** that either carry out particularly risky data processing activities or normally employ a minimum of ten people who work full time on the automated processing of personal data (Art. 37 para. 1 GDPR, Sec. 38 BDSG).

The data protection officers' tasks include **instructing** and **advising** data processors about their obligations under the GDPR and **monitoring** compliance with the regulation (Art. 39 para. 1 GDPR).

Data protection officers at research institutions are important contacts for questions about the implementation of the GDPR. They should be **consulted** when planning and realising research projects that require processing personal data. Data protection guidelines should be developed in coordination with them.

Supervisory authorities provide more detailed information on the process of appointing data protection officers and their rights and obligations (DSK 2018g, ULD 2019).

5.6 Data protection impact assessment

For especially risky data processing operations, the GDPR requires a data protection impact assessment (Art. 35 GDPR). This can be necessary, for example, if large amounts of **special categories**[10] **of personal data** (Art. 9 GDPR, see section 4.1) are processed for a research project – for example, health data or data regarding religious affiliation.

Data protection impact assessments must contain, among other things, descriptions of the processing operations and their purpose, and an assessment of the existing risks and remedies planned (Art. 35 para. 7 GDPR). As of yet, there are no best practises in conducting a data protection impact assessment, as this is a new instrument introduced by the GDPR.

Papers of the data protection authorities provide some orientation in which cases (LDA Brandenburg 2018) and in what manner (DSK 2018c) data protection impact assessments should be carried out.

Upon request, the institution's **data protection officer** will be involved in the preparation of the impact assessment (Art. 39 para 1 letter c GDPR). According to Art. 36 GDPR, it may also be necessary to involve the responsible **supervisory authority** if a data protection impact assessment concludes that the processing would pose a high risk, should the controller fail to mitigate that risk.

10 Data revealing racial or ethnic origin, political opinions, religious or philosophical beliefs, or trade union membership, genetic data, biometric data for the purpose of uniquely identifying a natural person, data concerning health or data concerning a natural person's sex life or sexual orientation. Data revealing racial or ethnic origin, political opinions, religious or philosophical beliefs, or trade union membership, genetic data, biometric data for the purpose of uniquely identifying a natural person, data concerning health or data concerning a natural person's sex life or sexual orientation.

6 Data collection (during field work)

The following section focuses on the first-time collection of personal data by researchers (**collection of primary data**). Prior to the first-time collection of data, the research purpose should be defined precisely and researchers should check exactly what data are needed (with the principle of data minimisation in mind). Even before collecting the data, there has to be a concept of how the data will be saved, stored (see section 9), and protected using technical and organisational measures (see sections 5.1 and 5.2). In some cases, the research institution's data protection officer should also be involved in this process.

The legal basis for processing the data can either be consent or statutory authorisation.

6.1 Processing data based on consent

If data are collected based on consent, the first step will be to ensure that consent qualifies as **informed and specific** (see section 4.5). The intended and possible use of the data should be described as comprehensively and with as much foresight as possible since consent is granted solely for the stipulated purposes. Any planned necessary transfer of data to research partners, repositories, or other entities for the purpose of joint processing, long-term storage or secondary use, or the publication in anonymised form, must already be noted in the statement of consent.

Data subjects must be informed that participating in the data collection is **voluntary** and that consent can be withdrawn. Anyone who refuses to participate must not be disadvantaged and there must not be any pressure to participate.

If the researchers plan to collect data about people other than the respondents themselves (e.g., survey questions about the spouses or children of respondents), it does not suffice to obtain the consent of the respondents alone.[11] Either additional consent must be obtained from the **third parties** or another legal basis must be used (see section 6.2).

The data subjects should be informed if a **survey is to be repeated**. In order to contact them again, their consent should also include the storage of their contact information for this purpose.

If new data are to be processed in a **"record linkage"** procedure, in which data from different sources pertaining to the same person are linked to one another, this must be mentioned separately in the statement of consent.

In addition to the data collected directly in a survey, statements of consent can also refer to other data that are incidental to the collection process (**paradata**). Paradata relate to, for example, the type and number of contacts, the answer time per question (when computer-assisted data collection systems are used), the rate of speech, or other characteristics of an interview. These data can, for example, be used for quality assurance in a study. Consent can be extended to the paradata if this is made sufficiently clear in the statement.

As proof that consent was given, it should be **documented** at least in text form, even if this form is not required for consent to be effective (see section 4.5). This documentation should be kept at least as long as the data are being processed. Valid consent provides legal certainty for the processing of the research data. However, note that data subjects have the right to **withdraw** their consent for future use at any time.

11 Although it has not yet clear whether data protection regulations require that consent be given personally, a valid authorisation would be necessary in any case for respondents to give consent on behalf of third parties.

6.2 Statutory legal basis

The collection of data can not only be based on consent but also on a statutory legal basis. The general authorisation criteria provided under the GDPR need to be met: While public entities can refer to the **fulfilment of their tasks** (Art. 6 para 1 letter e GDPR), private data processors will mostly use **overriding interests** (Art. 6 para 1 letter f GDPR) as a legal basis. If special categories of personal data are collected, Art. 9 GDPR must be observed as well (see section 4.1).

Consent **does not generally take priority** over other forms of authorisation. In some cases, the collection of data based on the statutory authorisation can be more practical than doing so based on consent. This is true, for example, if the data are not collected in a survey but from other sources. The statutory legal bases can also be convenient for the processing of paradata (data accruing not directly from a survey, but during data collection).

Compared to the collection of data based on consent, the collection and processing of data based on statutory legal bases has the advantage that data subjects cannot revoke it. However, they **may still object** to processing their data on some grounds (Art. 21 GDPR). Note also that obtaining consent can prevent recourse to statutory legal bases if the data subject trusts that their data are processed only on the given basis and within the scope of the consent. If obtaining consent is desirable or necessary, it may be reasonable to state on the consent form that the data could be processed on another legal basis if consent is withdrawn.

Example: A researcher asks an interviewee to consent to processing the information obtained in the interview for their research project. She tells the interviewee that this information would only be processed with their consent. Two weeks later, the interviewee decides to revoke her consent because in hindsight, she is uncomfortable with her answers. In that case, the researcher cannot process the information based on Art. 6 para. 1 letter e or f GDPR (public task or legitimate interests) because this would violate the interviewee's trust.

7 Data preparation and data analysis (following field work)

Data preparation and data analysis are **new processing steps** that also require justification. To the extent that preparation and analysis serve the same research purpose as the collection of data, the authorisation on which the collection of data was based will usually also allow the subsequent processing steps. Problems can arise if new methods or targets of inquiry are added during the analysis that, while still within the scope of the research purpose, heavily burden the data subjects. In such cases, it can be necessary to obtain additional consent or weigh the interests of everyone involved again.

Example: Over the course of a long-term research project on changes in worldviews, test subjects are regularly interviewed about their opinions on current affairs. After some time, it becomes apparent that not only do the data in aggregate form allow drawing general conclusions, but the data can also be used for a prognosis of individual processes of radicalisation – for example by employing complex new analytical methods or by linking the information with personal data in other available databases. A new target of inquiry may still be considered to be in line with the initial research purpose. However, based on the depth and breadth of the analysis, as well as the potentially stigmatising new objective, this data analysis will not simply be covered by the original consent or the initial weighing of interests, which justified the collection of data and the previous, less sensitive analyses. Instead, it must be independently verified whether statutory authorisation criteria are met or if new consent must be obtained.

While the data are being prepared, suitable **technical and organisational measures** must be taken to ensure the protection of the data subjects (see sections 5.1 and 5.2). The data must be pseudonymised as early as possible. After the preparation of data is completed, processors must check to what extent raw data (i.e., including metadata) and contact data need to be kept to ensure traceability or for other reasons (such as for panel surveys). As soon as the research purpose allows it, these data must be anonymised (Art. 89 para. 1 sentence 4 GDPR, Sec. 27 para. 3 sentence 1 BDSG). If the data subjects were promised their data would be deleted within a certain time limit, this must be done.

8 Publications (of data)

The publication of personal data is also a form of data processing (see section 2.1) that requires its own justification. An example of this are academic discussions of the behaviour, biographies, or illnesses of identifiable individuals. When data are published in journals or made available for re-analysis in repositories, the data are also considered published. The publication of data must be justified by consent, if they can still be attributed to individuals (i.e., are not anonymous) (Sec. 27 para. 4 BDSG).

Publications make information accessible to a potentially unlimited audience and therefore can be **especially severe infringements** on the rights of data subjects. From a data protection perspective, this requires special care. If possible without impeding the research purpose and without distorting the findings, the results obtained during the research project must be prepared for publication in such a way as to ensure that they can no longer be attributed to individuals (anonymisation, see section 4.2).

It is much easier to anonymise quantitative data effectively than qualitative data because in the case of former, the data themselves are not published but rather only the results from the analysis. In qualitative data analyses, individual data subjects might be identifiable from specific characteristics – such as a person's style of writing or way of speaking (see Winter/Battis/Halvani 2019: 491f.) – that cannot easily be erased without reducing the informative value of a publication. In such a case, finding the right balance between the conflicting concerns of effectively protecting the data through anonymisation on the one hand and ensuring academic accuracy on the other hand proves to be exceedingly difficult.

Example: A dialectologist publishes clips from interviews with speakers of rare dialects. Because of the individual styles of speech and the content of the interviews, the speakers are identifiable.

Sec. 27 para. 4 BDSG contains a separate provision for **publications of special categories of personal data** (see section 4.1) for research purposes. It stipulates that a publication requires either consent by the data subject (see section 4.5) or that publishing the data is essential for presenting research findings about events in contemporary history.

To assess whether the publication of personal data is **essential for presenting contemporary history**, the interest in historic research must be weighed against the interest in protecting personal privacy rights. It needs to be assessed what informational value the personal data have regarding the relevant event in contemporary history.

Unfortunately, with its focus on contemporary history, the exemption given in Sec. 27 para. 4 BDSG is quite narrow. In conjunction with Art. 9 GDPR, it makes publications more difficult when sensitive data are required but the protection of data subjects' identities cannot be guaranteed (for example, playing back interviews or even just parts thereof on political attitudes or individual medical histories that, together with additional information, could be used to identify the interviewees). For research projects of this kind, the only way to ensure that publishing results is legally safe is through consent, which should include publication in pseudonymised form from the beginning.

The publication of **other personal data** is subject to the general provisions under Art. 6 GDPR. They can also be published based on consent (Art. 6 para. 1 letter a GDPR), in the context of a public task or based on overriding interests (Art. 6 para. 1 letters e and f GDPR). The weighing of interests must consider that publishing data is a severe infringement on the rights of data subjects. Nevertheless, Art. 6 para. 1 letters e and f GDPR allow much greater leeway for the publication of personal data than Art. 9 GDPR and Sec. 27 para. 4 BDSG.

.

9 Storage and secondary use of research data

9.1 Storage and archiving

After a research project has been completed, the continued storage of personal data that have been collected raises data protection issues. The rules of **good academic practice** require that research data be stored after completing a project – according to the guidelines of the German Research Foundation (Deutsche Forschungsgemeinschaft, DFG), for example, normally for a period of ten years (DFG 2019: 22).

A legal basis is required, however, for storing data that are still personal data. This should be taken into consideration, if possible, when **consent** is obtained for data collection, and the maximum storage time should be stated. Otherwise, **statutory legal bases** can be used. Storage for the purpose of documenting research will generally be covered by the research obligations of public institutions. For private research, the decisive question when weighing interests will be whether data subjects have special interests that conflict with long-term storage.

Furthermore, also for archiving, **technical and organisational measures** to make data secure and prevent unauthorised access are indispensable and must continuously be updated.

In practice, repositories and research data centres both within or outside of a researcher's own institution are good options for storing research data. Data protection and data security should always be taken into consideration when selecting repositories and research data centres.

Detailed practical advice regarding archiving and repositories can be found in the German Data Forum's (RatSWD) guidelines on research data management in the social, behavioural, and economic sciences (RatSWD 2018). A current summary of information on data protection in research data centres accredited by the German Data Forum (RatSWD) can be found in their 2018 activities report (RatSWD 2019: 33ff.).

9.2 Secondary use

Finally, a recurring question is under which conditions stored (pseudonymised) research data may be used for new research projects. Processing data on the basis of consent or statutory legal bases is initially restricted to the purposes that were determined at the beginning (**purpose limitation**).

However, it is possible to continue processing the data on a new legal basis. According to Art. 5 para. 1 letter b half-sentence 2 GDPR, further processing of data – originally collected for other purposes – is possible for research purposes not considered incompatible with the original purpose. Accordingly, data may be processed further if the requirements for the legal basis are met. These requirements can arise from of the **same legal norm** that was used for the original processing of the data.

Example: Based on Art. 6 para 1 letter f GDPR, researchers collected data in the form of posts and interactions from social networks in order to gain insight into the language-specific characteristics of users from each network. Now, other researchers would like to use these data to research the group-specific discussion culture on social media. The work of these researchers is also covered by Art. 6 para 1 letter f GDPR, but the interests need to be weighed again. However, there is no need for a special permission for the change in purpose.

10 Checklists and best practices

The following checklists and bullet points are intended to provide a very brief overview of important aspects to be considered when personal data are processed for research purposes. They **are not exhaustive** and should be understood in the context of the information above.

10.1 General questions for processing data

- Are personal data being processed?
  - Definition under Art. 4 No. 1 GDPR. No personal data in case of effective anonymisation.
  - Every processing operation requires justification according to Art. 6 para. 1 GDPR: consent or statutory authorisation.
- Are the data special personal data (Art. 9 GDPR)?
  - Listed under Art. 9 para. 1 GDPR and defined in more detail under Art. 4 No. 13 et seq. GDPR.
  - If so, additional justification is necessary in accordance with Art. 9 para. 2 GDPR besides the regular justification (Art. 6 para. 1 GDPR).
- Have technical and organisational measures been taken to protect the data during processing?
  - This includes measures to protect against destruction, modification, loss, or disclosure of data using state-of-the-art technology.
  - Physical security measures, technological security measures for virtual storage capacities, training for personnel, written agreements (licence agreements), rights management.
- Is a data protection impact assessment required?

The following must be reviewed for every processing step:

- For public entities: Is the processing of the data necessary to fulfil a research obligation?
- For non-public entities: Is the processing of the data necessary in the interest of research, which is not overridden by the interests of the data subjects?
- OR: Is effective (voluntary) consent given (Art. 4 No. 11 GDPR)?
- Are there any important interests that nevertheless oppose the processing of the data?
- Could the task be completed using less data?
- Could it be carried out with anonymous/pseudonymous data?

10.2 Further considerations

- How are data being stored?
  - Safe storage solutions must be selected that prevent, as much as possible, the loss of or unauthorised access to data.
- Who can access the data? (Research team, assistants, office)
  - Access to the data must be restricted to include only those who need to work with the data.

- How long will the data be stored?
  - The maximum storage duration should be defined in all cases.
- Have the data subjects been informed about the processing of the data?
  - The obligation to provide information arises out of Art. 13 and Art. 14 GDPR and must be reviewed on a case-by-case basis.
- Is it possible to object to the processing of the data (erasure)?
  - This option must be given in every case.
- Will the data be transferred to third parties? If so, to whom? Does this include third parties outside the European Union or the European Economic Area?
  - The transfer of data is an independent processing step and must be reviewed for its legitimacy. The transfer of data to third countries must be reviewed independently in accordance with Art. 44 et seq. GDPR (see DSK 2019).

10.3 Best practices

- The research question and the methods used in a research project must be stated in a research design, which meets the standards of the research community regarding its content and approach. It should explain the types and scope of personal data that are collected, processed, and stored, and the technological means used to do so.
- The data protection officer of the applicable institution should be involved in the project as early as possible and be informed of any changes in the research design over the course of the project.
- The legal bases for processing the data for the project should be reviewed. This includes considering the option of obtaining consent even if consent does not take priority over other justifications. The reasons why a particular legal basis was deemed relevant for the planned processing of data should be documented. The key criteria for the weighing of interests should also be noted.
- The data collected will regularly be reviewed for their quality, security, and necessity.
- Technical and organisational measures should be taken to protect the data. They include – but are not limited to – precautions to minimise data, anonymisation or pseudonymisation, the definition of and adherence to time limits for storage, the erasure of data that are obsolete or not useful, the implementation of role concepts, and secure access solutions. Security mechanisms must be installed to prevent the data from being scraped or manipulated.
- To protect the rights of data subjects, a suitable technical and organisational framework must be established, for example through the order of the datasets.
- The research results and the databases on which they are based must be archived for the long term in compliance with data protection regulations if they are needed for other projects or for research reproducibility.
- Research results must be communicated in compliance with data protection regulations. In this regard, the possibility that modern technology may identify persons by using ostensibly harmless information should be taken into consideration.

11 References

Albrecht, Jan Philipp, and Florian Jotzo (2017): Das neue Datenschutzrecht der EU. Baden-Baden, Nomos.

Arning, Marian; Nikolaus Forgó, and Tina Krügel (2006): Datenschutzrechtliche Aspekte der Forschung mit genetischen Daten. Datenschutz und Datensicherheit 30(11), 700–705.

DFG [Deutsche Forschungsgemeinschaft] (2019): Leitlinien zur Sicherung guter wissenschaftlicher Praxis. Kodex. https://www.dfg.de/download/pdf/foerderung/rechtliche_rahmenbedingungen/gute_wissenschaftliche_praxis/kodex_gwp.pdf (Last accessed: 21.12.2019).

DSK [Datenschutzkonferenz] (2018): Kurzpapier Nr. 1. Verzeichnis von Verarbeitungstätigkeiten – Art. 30 DS-GVO. Stand: 17.12.2018. https://www.datenschutzkonferenz-online.de/kurzpapiere.html (Last accessed: 21.12.2019).

DSK [Datenschutzkonferenz] (2018a): Hinweise zum Verzeichnis von Verarbeitungstätigkeiten, Art. 30 DS-GVO. https://www.datenschutzkonferenz-online.de/media/ah/201802_ah_verzeichnis_verarbeitungstaetigkeiten.pdf (Last accessed: 21.12.2019).

DSK [Datenschutzkonferenz] (2018b): Muster für Verantwortliche gemäß Artikel 30 Abs. 1 DSGVO. https://www.datenschutzkonferenz-online.de/anwendungshinweise.html (Last accessed: 21.12.2019).

DSK [Datenschutzkonferenz] (2018c): Kurzpapier Nr. 5. Datenschutz-Folgenabschätzung nach Art. 35 DS-GVO. Stand: 17.12.2018. https://www.datenschutzkonferenz-online.de/kurzpapiere.html (Last accessed: 21.12.2019).

DSK [Datenschutzkonferenz] (2018d): Kurzpapier Nr. 6. Auskunftsrecht der betroffenen Person, Art. 15 DS-GVO. Stand: 17.12.2018. verfügbar unter https://www.datenschutzkonferenz-online.de/kurzpapiere.html (Last accessed: 21.12.2019).

DSK [Datenschutzkonferenz] (2018e): Kurzpapier Nr. 10. Informationspflichten bei Dritt- und Direkterhebung. Stand: 16.01.2018. https://www.datenschutzkonferenz-online.de/kurzpapiere.html (Last accessed: 21.12.2019).

DSK [Datenschutzkonferenz] (2018f): Kurzpapier Nr. 11. Recht auf Löschung / „Recht auf Vergessenwerden". Stand: 17.12.2018. https://www.datenschutzkonferenz-online.de/kurzpapiere.html (Last accessed: 21.12.2019).

DSK [Datenschutzkonferenz] (2018g): Kurzpapier Nr. 12. Datenschutzbeauftragte bei Verantwortlichen und Auftragsverarbeitern. Stand: 17.12.2018. https://www.datenschutzkonferenz-online.de/kurzpapiere.html (Last accessed: 21.12.2019).

DSK [Datenschutzkonferenz] (2018h): Kurzpapier Nr. 17. Besondere Kategorien personenbezogener Daten. Stand: 27.03.2018. https://www.datenschutzkonferenz-online.de/kurzpapiere.html (Last accessed: 21.12.2019).

DSK [Datenschutzkonferenz] (2019): Kurzpapier Nr. 4. Datenübermittlung in Drittländer. Stand: 22.07.2019. https://www.datenschutzkonferenz-online.de/kurzpapiere.html (Last accessed: 21.12.2019).

DSK [Datenschutzkonferenz] (2019a): Kurzpapier Nr. 20. Einwilligung nach der DS-GVO. Stand: 22.02.2019. https://www.datenschutzkonferenz-online.de/kurzpapiere.html (Last accessed: 21.12.2019).

Golla, Sebastian; Henning Hofmann, and Matthias Bäcker (2018): Connecting the Dots – Sozialwissenschaftliche Forschung in Sozialen Online-Medien im Lichte von DS-GVO und BDSG-neu. Datenschutz und Datensicherheit 42(2), 89–100.

Golla, Sebastian (2019): Datenschutz in der Forschung und Hochschullehre. In: Louisa Specht und Reto Mantz (Hrsg.): Handbuch Europäisches und deutsches Datenschutzrecht. München, C.H. Beck, 646–671.

Johannes, Paul C., and Philipp Richter (2017): Privilegierte Verarbeitung im BDSG-E, Regeln für Archivierung, Forschung und Statistik. Datenschutz und Datensicherheit 41(5), 300–305.

Kipker, Dennis-Kenji, and Maren Pollmann (2019): Sozialdatenschutz. In: Louisa Specht und Reto Mantz (Hrsg.): Handbuch Europäisches und deutsches Datenschutzrecht. München, C.H. Beck, 718–761.

LDA Brandenburg [Die Landesbeauftragte für den Datenschutz und für das Recht auf Akteneinsicht] (2018): Liste von Verarbeitungsvorgängen nach Art. 35 Abs. 4 DS-GVO. https://www.lda.brandenburg.de/media_fast/4055/DSFA_Muss_Liste_Allgemein_17102018.pdf (Last accessed: 21.12.2019).

RatSWD [Rat für Sozial- und Wirtschaftsdaten] (2017): Forschungsethische Grundsätze und Prüfverfahren in den Sozial- und Wirtschaftswissenschaften. RatSWD Output 9(5). Berlin, Rat für Sozial- und Wirtschaftsdaten (RatSWD). https://doi.org/10.17620/02671.1 (Last accessed: 21.12.2019).

RatSWD [Rat für Sozial- und Wirtschaftsdaten] (2018): Forschungsdatenmanagement in den Sozial-, Verhaltens- und Wirtschaftswissenschaften – Orientierungshilfen für die Beantragung und Begutachtung datengenerierender und datennutzender Forschungsprojekte. RatSWD Output 3(5). Berlin, Rat für Sozial- und Wirtschaftsdaten (RatSWD). https://doi.org/10.17620/02671.7 (Last accessed: 21.12.2019).

RatSWD [Rat für Sozial- und Wirtschaftsdaten] (2019): Tätigkeitsbericht 2018 der vom RatSWD akkreditierten Forschungsdatenzentren (FDZ). Berlin, Rat für Sozial- und Wirtschaftsdaten (RatSWD). https://doi.org/10.17620/02671.40 (Last accessed: 21.12.2019).

ULD [Unabhängiges Landeszentrum für Datenschutz Schleswig-Holstein] (2019): Datenschutzbeauftragte. Praxis-Reihe: Datenschutzbestimmungen praktisch umsetzen 2. https://www.datenschutzzentrum.de/praxisreihe (Last accessed: 21.12.2019).

ULD [Unabhängiges Landeszentrum für Datenschutz Schleswig-Holstein] (2019a): Informationspflichten. Praxis-Reihe: Datenschutzbestimmungen praktisch umsetzen 4. https://www.datenschutzzentrum.de/praxisreihe (Last accessed: 21.12.2019).

Von Lewinski, Kai (2017): Melderegisterdaten als Grundlage für empirische Sozialstudien. Verwaltungsrundschau 63(1), 1–7.

Winter, Christian; Verena Battis, and Oren Halvani (2019): Herausforderungen für die Anonymisierung von Daten. Technische Defizite, konzeptuelle Lücken und rechtliche Fragen bei der Anonymisierung von Daten. Zeitschrift für Datenschutz 9(11), 489–493.

Appendix

Overview of the legal basis for data provisions in the accredited research data centres (RDC)

Data Centre	GDPR	BDSG	LDSG	SGB	BStatG	Credit Act	Copyright Act	Other
German General Social Survey (ALLBUS/GGSS)	x	x						
Archive for Spoken German (AGD)	x							
Business and Organizational Data (BO)	x	x	x				x	
German Family Panel pairfam	x							
German Federal Employment Agency (BA) at the Institute for Employment Research (IAB)	x			x				
Deutsche Bundesbank	x				x	x		x[a]
Federal Institute for Vocational Education and Training (BIBB)	x	x		x				x[b]
Federal Centre for Health Education (BZgA)	x	x						
German Pension Insurance (RV)	x			x				
DIPF | Leibniz Institute for Research and Information in Education	x	x	x				x	
German Youth Institute (DJI)	x							x[c]
German Centre of Gerontology (DZA)	x	x						
German Centre for Higher Education Research and Science Studies (DZHW)	x	x						
German Centre for Integration and Migration Research (DeZIM)	x							
Interdisciplinary Centre for Qualitative Data in the Field of Sociology of Labour and Industry (eLabour)	x	x						
Institute for the Study of Labor (IZA)	x							
German Microdata Lab (GML)					x			
German Pharmacoepidemiological Research Database (GePaRD)				x				
Ifo Institute - Leibniz Institute for Economic Research	x	x	x					

Data Centre	GDPR	BDSG	LDSG	SGB	BStatG	Credit Act	Copyright Act	Other
Institute for Educational Quality Improvement (IQB)	x	x						
International Survey Programmes	x	x						
Kraftfahrt-Bundesamt (KBA)					x			
LIfBi – Leibniz Institute for Educational Trajectories	x	x						
Leibniz Institute for Financial Research SAFE	x	x	x					
Leibniz Institute of Ecological Urban and Regional Development (IOER)		x						x[d]
Halle Institute for Economic Research (IWH)	x	x			x			
Leibniz Institute for Psychology (ZPID)	x	x	x					
Programme for the International Assessment of Adult Competencies (PIAAC)	x	x						
Qualiservice	x	x	x				x	
Robert Koch Institute (RKI)	x	x						
RWI – Leibniz Institute for Economic Research	x	x		x	x			
German Socio-Economic Panel (SOEP)	x							
Statistical Offices of the Länder					x			
Federal Statistical Office (destatis)	x				x			x[e]
Stifterverband	x	x			x			
Survey of Health, Ageing and Retirement in Europe (SHARE)	x	x						
Elections	x							
ZEW – Leibniz Centre for European Economic	x	x			x			

Note: As of May 2020.
GDPR = EU General Data Protection Regulation, BDSG = German Federal Data Protection Act, LDSG = German State Data Protection Act, SGB = German Social Security Code, BStatG = German Federal Statistics Law, KWG = Banking Act, UrhG = Copyright Act

a EU Statistics Regulations, Foreign Trade and Payments Act, private law contracts with business partners
b Sec. 90 Vocational Training Act (BBIG)
c Federal Budget Code (BHO): Grant notification
d Geodata Access Act (GeoZG)
e EU VO Nr. 557/2013

Contributors

to the revision of the 2nd edition

Revision

Prof. Dr. Matthias Bäcker
Johannes Gutenberg University Mainz

Dr. Sebastian Golla
Johannes Gutenberg University Mainz

Consultation group of the RatSWD

Prof. Dr. Cordula Artelt
Leibniz Institute for Educational Trajectories (LIfBi)

Prof. Stefan Bender
Deutsche Bundesbank

Dr. Jan Goebel
Socio-Economic Panel (SOEP) at DIW Berlin

Heike Habla
Federal Statistical Office

Eckart Hohmann

Dr. Cornelia Lange
Robert Koch Institute (RKI)

Bertram Raum
Expert for Data Protection Law

Prof. Regina T. Riphahn, Ph.D.
University of Erlangen-Nürnberg

Dr. Heike Wirth
GESIS – Leibniz Institute for the Social Sciences

German Data Forum (RatSWD) business office

Dr. Mathias Bug

Thomas Runge

Dr. Katrin Schaar

Contributors to the publication of the 1st edition

Authors

Tobias Gebel
German Institute for Economic Research (DIW Berlin)

Heike Habla
Federal Statistical Office

Dr. Cornelia Lange
Robert Koch Institute (RKI)

Alexia Meyermann
DIPF | Leibniz Institute for Research and Information in Education

Prof. Regina T. Riphahn, Ph.D.
University of Erlangen-Nürnberg

Daniel Schmidutz
Max Planck Institute for Social Law and Social Policy (MPISOC)

German Data Forum (RatSWD) business office

Claudia Oellers

Thomas Runge

Imprint

Publisher:
Rat für Sozial- und Wirtschaftsdaten (RatSWD)
Am Friedrichshain 22
10407 Berlin
Germany
office@ratswd.de
https://www.ratswd.de

Editors:
Thomas Runge, Dr. Mathias Bug, Dr. Katrin Schaar

Layout:
Claudia Kreutz

Translation | proof-reading:
Regina Seelos, https://www.seelos.de | Jonas Huggins

Icons:
made by Freepik from https://www.flaticon.com
Font Awesome, fontawesome.com (modified)

Berlin, October 2020

RatSWD Output:
The RatSWD Output Series documents the German Data Forum's (RatSWD) activities during its 6th appointment period (2017–2020). It serves to publish its statements and recommendations and to make them available to a broad audience.

This report is the result of a project that is funded by the Federal Ministry for Education and Research (reference number: 01UW1802). Unless otherwise stated, the responsibility for this publication lies with the German Data Forum (RatSWD).

doi: 10.17620/02671.57

Suggested citation:
RatSWD [German Data Forum] (2020): Data Protection Guide. 2nd fully revised edition. RatSWD Output 8 (6). Berlin, German Data Forum (RatSWD). https://doi.org/10.17620/02671.57.

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 38 research data centres (as of May 2020) and fosters their interaction and collaboration.

www.ratswd.de