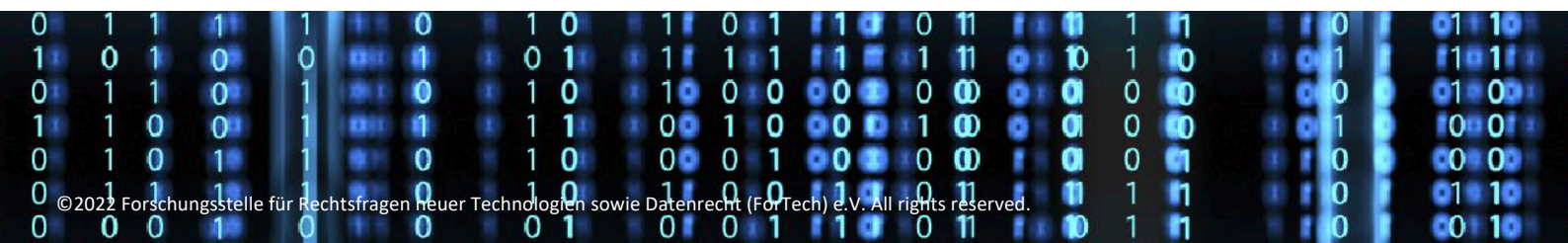


Research data access

Framework conditions, principles, and guidelines for privileged access to data for research and science

*Prof. Dr Louisa Specht-Riemenschneider/Alexander Wehde**

I. Introduction.....	1
II. Course of the investigation	2
III. (Union) Fundamental Rights Framework of a Research Data Access Regime	2
IV. Data protection legal framework	4
1. Consent.....	4
2. Secondary use without consent	4
a) Data processing for the fulfilment of tasks in the public interest.....	5
b) Data processing on the basis of a balancing of interests.....	5
c) Permits under German federal state law.....	6
d) Change of purpose according to Article 6 (4) GDPR	7
V. Principles of research data access.....	7
1. Openness by Design.....	7
2. Five safes model	8
3. FAIR Data Principles.....	8
VI. Regulatory structure of research data access de lege lata.....	9
VII. Analysis of national clauses to research data access	10
VIII. Guidelines for research data access de lege ferenda.....	12
1. Health sector	12
2. Online business sector.....	14
3. Mobility and energy sector.....	14



Abstract:

Access to data for public good research is fundamental in many sectors to solve urgent societal problems. Structurally, this data access can be achieved in various ways, but especially subjective rights to data access for research in the public interest are urgently needed. The standards for research data access should be based on the Five Safes Model and the FAIR principles. In terms of fundamental rights, the legal positions of the data subject, the fundamental rights of third parties, and the fundamental interests of those seeking access must be taken into account and appropriately balanced when structuring data access rights. In this respect, not only the type and scope of the access claim but also its limitation provisions must be considered.

I. Introduction

"Society needs science - science needs data".¹ Access to data for research and science is a basic condition for gaining scientific knowledge, also and especially in the interest of society as a whole. In the health sector, for example, it helps to identify the side effects of medicines or vaccinations and to improve treatment options. Only access to databases of platforms in the online business sector allows research and science to investigate algorithms used, for example, concerning discrimination issues, potential overblocking, the spread of disinformation, or even the practices used to persuade users to make a certain decision. Spotify, for example, has developed voice recognition software that can be used to make recommendations for songs or ads depending on the user's mood.² This at least has the potential to induce bad purchases as a result of projection bias.³ In the mobility sector, improved data access for research can be used to develop and evaluate intelligent transport systems. In the energy sector, energy efficiency could increase through the analysis of electricity consumption figures and factors.

The lack of access to research data is increasingly being complained about, especially in recent times.⁴ The coalition agreement of the new German federal government explicitly states that it wants to take appropriate measures to improve access to research data.⁵ However, Hevers already called for the creation of a comprehensive right to information access for the benefit of research and science in his dissertation "Informationszugangsansprüche des forschenden Wissenschaftlers" (Information access claims of the researching scientist) from 2015,⁶ and *Wielsch* developed media-specific access rules for intellectual property in his habilitation thesis "Zugangsregeln - Die Rechtsverfassung der Wissensteilung" (Access rules - the legal constitution of knowledge sharing).⁷ The German Competition Law 4.0 Commission also spoke out - albeit narrowly - in favour of data access for the benefit of research.⁸

In Germany, access to research data in the social, behavioural and economic sciences is de lege lata based on a system of 39 research data centres accredited by the German Council for Social and

Economic Data (RatSWD). The granting of access via research data centres is an important basis for empirical research, but it is not the focus of this paper. Rather, the investigation carried out here focuses primarily on guaranteeing access to the data generated by public and private bodies that are not already de lege lata stored in research data centres, because these data are of particular relevance to the sectors under investigation.

In terms of substantive law, data access can be guaranteed in various ways: Contractual claims to information lead to data access, as does the exercise of powers under data protection law.⁹ There are also access claims under cartel law.¹⁰ However, all of these access claims do not apply primarily to the scientific community; rather, persons who are not scientifically active may also be entitled to access. The focus of this article, however, is explicitly on research clauses that standardise privileged data access for science and research. Non-specific access claims are thus neglected as a subject of investigation. Beyond contractual, data protection and antitrust obligations, however, data access has so far been granted in an extremely rudimentary manner. Clauses allowing or obliging data access for research (research clauses) are found only sporadically, e.g. in Section 19 (3) Act on the Copyright Liability of Online Content Sharing Service Providers (UrhDaG), in Section 5a Network Enforcement Act (NetzDG), in Section 8 Federal Cancer Registry Data Act (BKRG) or Section 303e German Social Code, Book V (SGB V). In addition, the data covered by the right of access are limited, and the conditions for access vary greatly. Principles and rules for access to research data have not yet been systematically developed. This is the task of this article.

From the currently discussed need for data access rules in the area of B-B, G-B and C-B, which are addressed by the draft Data Act announced in February 2022, only a section of the required data access regime¹¹ is addressed with research data access.

II. Course of the investigation

For research data access, the (Union) fundamental rights framework, as well as the data protection law framework, will first be outlined. Then, the principles that research data access should follow will be presented. The data access claims existing in national law will then be examined with regard to their regulatory structure, their prerequisites and legal consequences, to ultimately be able to develop guidelines for the future design of research clauses from the fundamental rights framework, the principles of research data access and the analyses of the existing data access claims.

III. (Union) Fundamental Rights Framework of a Research Data Access Regime

Depending on the level at which data access claims are introduced, the framework for granting such data access claims is based on national fundamental rights or the CFR and ECHR. In each case, the legal positions of the data subject, the fundamental rights of third parties and the fundamental rights of those seeking access must be appropriately balanced. On the side of the data subject, the protection

of trade secrets, which is enshrined in Article 12 or Article 14 of the Basic Law, depending on the opinion¹² expressed, and in Article 17 of the Charter of Fundamental Rights or Article 6, 15, 16 of the Charter of Fundamental Rights, depending on the opinion¹³ expressed, are affected. Equally, however, the protection of intellectual property standardised in Article 14 of the Basic Law and Article 17 CFR, e.g. the database protection under copyright law according to Section 2 Act on Copyright and Related Rights (UrhG) but also the sui generis right according to Section 87a UrhG as well¹⁴ as the protection of freedom of occupation (Article 12 (1) of the Basic Law or Article 15 (1) CFR) must be taken into account. On the part of the third parties affected by data access, the right to informational self-determination under Article 2 (1), Article 1 (1) of the Basic Law as well as the provisions of Article 7 and 8 CFR must be taken into account, and on the part of those seeking access, the freedom of research and science pursuant to Article 5 (3) of the Basic Law or Article 13 CFR.

The interference in the legal positions of the data access addressees pursues a legitimate purpose with the granting of the activity of research and science; within the framework of appropriateness, the legislature traditionally has a very broad scope of assessment,¹⁵ which should not be exceeded in the case of granting data access in favour of research and science. Voluntary measures as a milder means within the framework of necessity are in any case not likely to be suitable to the same extent. Within the framework of reasonableness, data access in the public interest is much more likely to be justified than a data access claim in the private interest, because the public interest should generally weigh much more heavily. Corresponding compensation possibilities of the data access addressee should also contribute to the appropriateness of the regulation.¹⁶ Within these conflicts of fundamental rights, freedom of research is of considerable importance in the context relevant here. Data access is to be guaranteed in its favour, which is why its substantive scope of protection in Article 13 CFR and Article 5 (3) of the Basic Law requires explanation in the following.

Member States are bound by Community fundamental rights when they transpose Community law into national law, enforce Community law or restrict fundamental freedoms domestically.¹⁷ However, where the Member States do not enforce Community law, the constitutional understanding of the concept of freedom of research applies. In the case of freedom of research, however, both conceptual understandings largely coincide: According to Article 13 CFR, art and research are free. Academic freedom is respected. The ECJ has not yet defined the concept of research, but the literature falls back on the case law definition of the German Bundesverfassungsgericht (Federal Constitutional Court), since Article 13 CFR is considered to be inspired by the German Basic Law.¹⁸ The Bundesverfassungsgericht defines science as "any activity which, according to its content and form, is to be regarded as a serious and planned attempt to determine the truth".¹⁹ According to the Bundesverfassungsgericht, scientific research is the "intellectual activity with the aim of gaining new knowledge in a methodical, systematic and verifiable manner".²⁰ A particular methodical procedure

and a certain level of knowledge are required.²¹ It is irrelevant, however, whether the research takes place within or outside of higher education institutions.²² The term research is to be understood broadly and also includes private research.²³ Even preparatory and supporting activities are protected.²⁴ In addition to natural and legal persons (insofar as Article 5 (3) of the Basic Law is applicable to these legal persons in its essence, Article 19 (3) of the Basic Law), holders of fundamental rights are also institutions of higher education, even if they are institutions or corporations under public law.²⁵

IV. Data protection legal framework

If personal data are made accessible for science and research, the requirements of data protection law must be observed in addition to the requirements of trade secret protection and database protection. However, access to research data is already possible *de lege lata*, even with regard to sensitive personal data as defined in Article 9 GDPR. Essentially, access to research data can be based on the following legal foundations:

1. Consent

The disclosure of non-sensitive personal data can be based on consent, Article 6 (1) (a) GDPR. First of all, the general requirements apply, so that consent must be voluntary, informed and unambiguous, among other things, Article 4 (11), Article 7 GDPR. If sensitive data is involved, explicit consent is required, Article 9 (1) GDPR. The data subject must be adequately informed about the intended data processing and thus also about the disclosure of the data before the data is collected. Consent must always be given for the specific data processing.²⁶ However, when data is collected, it is often not possible to foresee the later research project in concrete terms.²⁷ Therefore, in view of Recital 33, less stringent requirements for the specificity of consent should apply in the context of scientific research.²⁸ However, it is not yet comprehensively differentiated how such a less specific declaration of consent is to be concretely formulated. The German Medical Informatics Initiative has developed a model text for medical research for broad consent to the secondary use of pseudonymised data, to which the Conference of Independent Data Protection Commissioners of the Federation and the Länder declared its agreement on 15 April 2020.²⁹ At the European level, the draft Data Governance Act provides that the applicability of such standardised consent forms for the declaration of "broad consent" can be declared permissible by delegated act.³⁰ Concepts of multi-level or dynamic consent are also conceivable.³¹

2. Secondary use without consent

However, secondary use of data does not necessarily have to be based on consent. Other grounds for permission can also be considered.

a) Data processing for the fulfilment of tasks in the public interest

The processing of non-sensitive data can be based in particular on Article 6 (1) (e) and (f) GDPR. With regard to the secondary use of sensitive data, the requirements of Article 9 GDPR must be taken into account. Article 6 (1) (e) GDPR can only be used to justify data processing by public authorities. These in Germany can be public bodies of the Federation and the Länder, which may be organised under public law as well as under private law, as well as authorised persons, provided that data processing is necessary for a public task assigned to the body. Such an assignment of tasks is made by legal requirements of the Union or the member states, so that Article 6 (1) (e) GDPR is not to be understood as a stand-alone authorisation.³² The norm follows a strictly functional approach.³³ In particular, it justifies data processing for research purposes by universities as research institutions under public law, as they are expressly assigned research tasks by the higher education laws of the Länder.³⁴ In the case of processing sensitive data, the requirements of Article 9 (2) (g) and (i) GDPR must be observed, according to which a simple public interest is no longer sufficient for data processing, but a substantial public interest is required.

b) Data processing on the basis of a balancing of interests

Finally, data processing for scientific purposes can be based on Article 6 (1) (f) GDPR and for sensitive data on Article 9 (2) (j) GDPR in conjunction with Section 27 (1) 2 BDSG.³⁵ Article 6 (1) (f) GDPR requires a careful balancing of the legitimate interests of the controller or a third party and the interests or fundamental rights and freedoms of the data subject, taking into account all the circumstances of the individual case.³⁶ In the context of Article 6 (1) (f) GDPR, it is necessary that the interests are at least of equal importance.³⁷ These legitimate interests may also include the research interests of the controller. Public bodies may only rely on this element of permission to the extent that they are not in a specific governmental relationship with the data subject, but are confronted as participants in private legal transactions (e.g. in the context of auxiliary fiscal transactions).³⁸ For research institutions organised under private law, however, Article 6 (1) (f) GDPR applies without restriction.

Article 9 (2) (j) GDPR in conjunction with Section 27 BDSG does not differentiate between public and non-public bodies. The prerequisite is that the secondary use of sensitive, personal data is necessary for scientific research purposes and that the research interests significantly outweigh the interests of the data subjects. Processing of non-anonymised data can then no longer be based on Section 27 (1) BDSG if the purposes could also be achieved by working with anonymised data.³⁹ This is in line with the anonymisation obligation in Section 27 (3) BDSG, according to which data must be anonymised as soon as this is possible or compatible with the purpose pursued. In any case, a separate storage of the allocation characteristics has to take place according to Section 27 (3) BDSG. For the secondary use of

sensitive health data, appropriate and specific measures must be taken to protect the interests of the data subject within the meaning of Section 22 (2) BDSG.⁴⁰

It has not yet been conclusively clarified when an overriding research interest exists. In any case, the general weighing of interests from Article 6 (1) (f) GDPR is intensified to the effect that, due to the addition of the word "substantial", in case of doubt, the interests of the data subject must be assumed to prevail.⁴¹ In any case, a balancing of the conflicting fundamental rights positions is necessary.⁴² The publication of sensitive, personal data is only possible under additional conditions, cf. e.g. Section 27 (4) BDSG. However, if research data is only disclosed to a limited group of recipients and the disclosure is secured by suitable measures such as non-disclosure agreements or similar, this does not constitute publication.⁴³

c) Permits under German federal state law

At the German federal state level, further authorisation criteria include Section 11 (1) HmbDSG, Section 17 (1) BlnDSG, Section 13 (1) LDSG-BW (both for the processing of sensitive data and non-sensitive data) as well as Article 8 BayDSG, Section 13 (1) BremGDPRAG, Section 24 (1) HDSIG (exclusively for the processing of sensitive data).⁴⁴ Pursuant to Section 11 (1) HmbDSG, public bodies may process personal data for certain projects of scientific and historical research as well as statistics without consent, provided that interests of the persons concerned worthy of protection are not impaired due to the nature of the data, their obviousness or the way they are used. Section 17 (1) BlnDSG also exempts scientific and historical research and statistics from the requirement of consent to process personal data. However, the prerequisite for this is that the public interest in carrying out the project significantly outweighs the interests of the data subject that are worthy of protection and that the purpose cannot be achieved in any other way. According to Article 13 (1) LDSG-BW, public bodies may process personal data for scientific or historical research purposes if the purposes cannot be achieved in any other way or can only be achieved with disproportionate effort and the research interests outweigh the interests of the data subject. The wording of the regulation in Article 8 BayDSG does not provide for any explicit special regulation for research when handling personal data. However, personal data may be processed for the purpose of preventive health care or occupational medicine if this is necessary (Article 8 (1) (3) BayDSG). Section 13 (1) BremDSGVOAG allows the processing of personal data for scientific or historical purposes, insofar as the processing is necessary for these purposes and the interests of the controller in the processing significantly outweigh the interests of the data subject in not having the processing carried out. The same applies to the provision of Section 24 (1) HDSIG.

d) Change of purpose according to Article 6 (4) GDPR

Processing for archiving purposes in the public interest for scientific or historical research purposes or for statistical purposes is also possible under Article 6 (4) GDPR. It is assumed that the secondary use is compatible with the original purpose of the processing of personal data. However, the requirements listed in Article 89 (1) GDPR must be taken into account, in particular, technical and organisational measures must be in place to ensure, among other things, respect for the principle of data minimisation.⁴⁵

V. Principles of research data access

Research data access can and must never be guaranteed across the board, but must follow certain principles in addition to the (Union) fundamental rights framework. These principles have already been developed internationally and will be summarised here. They are "Openness by Design", the "Five Safes Model" and the "FAIR Data Principles".

1. Openness by Design

The disclosure of data is not only a driver for innovation; research also benefits considerably from open data access. In the interest of society as a whole, public sector data should therefore be disclosed as a default setting ("openness by design"). Data openness can be achieved in different degrees, which can be described as follows:⁴⁶

1. Open data: data is made available for anyone to use, modify and distribute without restrictions.
2. Public data: Data is made publicly available, but with some restrictions on use.
3. Shared data: data is made available to a limited group of participants, possibly with some restrictions on use.
4. Closed data: The data is only available within a single organisation.

The degree of possible openness essentially results from the risk to the rights and interests of those affected by the data processing. The disclosure of administrative data can also affect the data protection rights or other interests of third parties that are worthy of protection. However, this risk can often be adequately countered by anonymisation or pseudonymisation. Only if this is not possible is it necessary to restrict the disclosure of data or limit access.⁴⁷ The protection of trade secrets, on the other hand, will be affected much more frequently in the case of access to privately-held data than in the case of administrative data.

2. Five safes model

The so-called "Five Safes Model" was developed for the benchmarks of data access, which sees data access as a spectrum of five risk dimensions. Each of these dimensions raises a specific question for data access, which at best can be answered in a regulatory way:⁴⁸

Safe projects	Is this use of the data appropriate?
Safe people	Can the researchers be trusted to use it in an appropriate manner?
Safe data	Is there a disclosure risk in the data itself?
Safe settings	Does the access facility limit unauthorised use?
Safe outputs	Are the statistical results non-disclosive?

Fig. 1: Designing data access for research (Five Safes)

The risk dimensions influence each other: the more it is ensured that the data is used within "safe projects", e.g. for projects in the public interest, and the more it is also ensured that only certain persons have access, the more extensively risky data, e.g. with personal references, can be made accessible.⁴⁹

3. FAIR Data Principles

In addition, data access for science and research should also be based on the FAIR principles developed for research data management at a workshop in Leiden in 2014, which have since been updated by a working group.⁵⁰ According to these principles, data must first be findable. For this purpose, they and the associated metadata could, for example, be given a unique, permanent label. The data would also have to be given extensive descriptive metadata, which would also contain the label described. Finally, the data, like the metadata, would have to be contained in a searchable database or similar. Furthermore, the data should be accessible. For this purpose, they and the associated metadata could be retrievable on the basis of their labelling, for example, by means of a standardised communication protocol. In addition, the data would have to be interoperable. To this end, they and their associated metadata would need to use formal, accessible, common and widely applicable knowledge representation semantics. They would also need to contain qualified references to other data or metadata. Systems are also required that allow and ideally promote data exchange, especially through appropriate interfaces and syntactic interoperability. Finally, the data must be reusable, which can be ensured through appropriate descriptions and transparent access conditions.⁵¹

VI. Regulatory structure of research data access de lege lata

Data access for science independent of already existing contractual relationships and also independent of data protection rights of the individual researcher can be achieved via five structurally different regulatory instruments, namely through

1. a claim to data access for the benefit of research that is not subject to fundamental rights
2. genuine research clauses
3. open data legislation
4. transparency regulations and reporting obligations
5. permissions to grant access to data

These five paths to data access have varying degrees of intensity in favour of research and science. Open data legislation primarily addresses public administration, for example through the E-Government Gesetz (EGovG), and obliges the disclosure of data to a group of persons not limited to science and research. The same applies to claims under the Freedom of Information Acts, the Environmental Information Act and the Consumer Information Act. Transparency requirements and reporting obligations are spread across many laws, especially vis-à-vis private institutions, for example in Section 2 NetzDG vis-à-vis providers of social networks and video-sharing platform services. They provide quite general information that can be used in science but is not likely to be of significant interest due to the lack of specificity of the data. Transparency regulations and reporting obligations rather serve the general public for rather general information.⁵² For the purposes of science and research, however, more detailed information and raw data are required, e.g. the data explicitly mentioned in Section 5a (6) NetzDG, their respective context and the training data of the automated procedures⁵³, in order to independently gain scientifically sound knowledge about the procedures in question for the automated recognition of content. Granting access to data requires a basis for processing under data protection law, at least if it concerns personal data. This permission to grant data access does not establish a substantive legal claim to data access and therefore also only has a limited effect in favour of science and research. Substantive legal claims to data access are only established by so-called genuine research clauses. More far-reaching could only be a data access claim not directly based on fundamental rights, whose existence could be guaranteed on the basis of freedom of information and research in accordance with the press law information claim recognised by the German Federal Administrative Court, but it is predominantly rejected de lege lata. Neither freedom of information grants a claim to the opening of information sources nor freedom of research. According to their legal intensity, a ladder of research data access emerges:⁵⁴

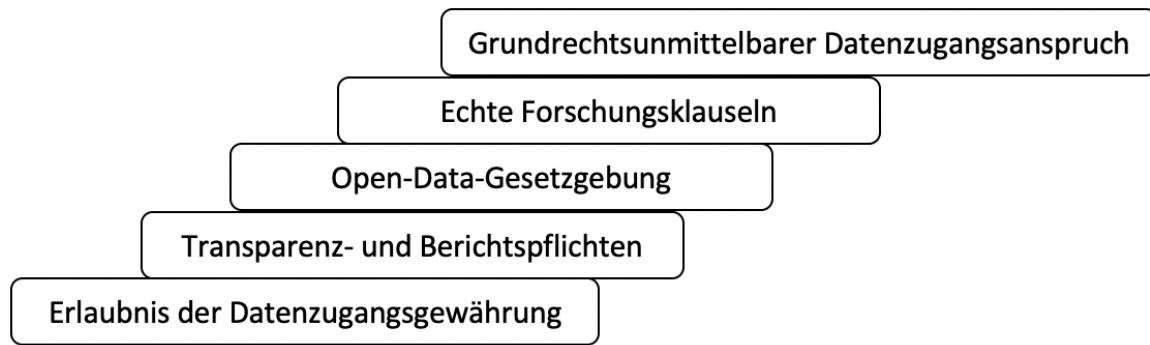


Fig. 2: Ladder of research data access

VII. Analysis of national clauses to research data access

Due to their intensity of effect in favour of science and research, the focus of this study is on genuine research clauses. Genuine research clauses exist *de lege lata* at national level in the online economy, health and mobility sectors, but not in the energy sector. A cross-sectoral data access claim for research and science does not exist.

In this context, a very narrow scope of application of the existing data access claims can be observed, which is specified in two ways: firstly, with regard to the data collected and secondly, with regard to the addressees of the access. If these are directed against public-law bodies, there is regularly a discretionary scope for the access addressee. The higher the protective measures, e.g. through anonymisation solutions, the more likely it is that the legal consequence will be a bound decision.⁵⁵ In the mobility sector, on the other hand, Section 1g (5) StVG (German Road Traffic Act) and Section 63a (5) StVG provide for discretion even in the case of access to non-personal data. This can be explained by potentially conflicting other legal positions, e.g. trade secret protection. If data access claims are directed against private bodies, the protection of the rights and interests of third parties is primarily guaranteed by the limitation provisions of the claim norm. Many - but not all - research clauses require research to be in the public interest, in some cases also in the public interest.⁵⁶ This can be explained in terms of fundamental rights dogma: The public interest is a particularly serious public interest and is therefore particularly suitable in terms of fundamental rights dogma to justify the encroachment on the fundamental rights of the access addressees associated with research data access. The more likely it is that the public interest can also be achieved through commercial research, the more likely it is that this can also be covered by the research clauses as being eligible for access.

Additional requirements are only stipulated by a few research clauses. However, where they are provided for, e.g. within the framework of Section 5a NetzDG in the form of a protection concept to be submitted upon application, they serve to protect the fundamental rights and freedoms of third parties or the access addressee.

Data access for the benefit of science and research is not guaranteed *de lege lata* in the majority of research clauses without restriction but is limited by rules on limits. This, too, is an expression of the balancing of the different fundamental rights positions of those entitled, those obliged to make claims and third parties. The more limited the scope of application of the research clause and the more far-reaching the requirements for a protection concept to be submitted at the time of application, the less restrictive the provisions of the research clauses. The reason for this is that the fundamental rights of the addressees of access and third parties are already taken into account to a large extent at the level of the scope of application, the form of the application and the protection concept so that the provisions on limitations can themselves be more limited.

A number of research clauses tie data access to the purpose of (scientific) research.⁵⁷ Purpose limitation of data access means that these data may initially be used solely for the named purposes, e.g. to carry out scientific research projects. Alternatively or cumulatively, there are regulations on connection use that can affect both the source data (e.g. anonymisation after completion of the research project; connection use that changes the purpose) and the research results. Further prerequisites for data access can be found in some research clauses in the criterion of necessity, in some cases, a "particularly justified case" is also provided for as a restriction, while other research clauses provide for the participation of a scientific committee in the decision on data access. All these additional conditions have in common that they restrict data access for science and research in order to ensure a balance with conflicting legal positions by providing further "safeguards" to protect these conflicting rights. These safeguards can be substantive (through the criterion of necessity) or formal (through the participation of bodies in the decision-making process).

In the health sector, a comprehensive regulation of data access remuneration can be found in the Data Transparency Fee Ordinance (Datentransparenz-Gebührenverordnung), which applies to the granting of data access according to Section 303a SGB V. In the online economic sector, remuneration regulations exist, at least in national law, in the form of a claim for reimbursement of the costs incurred through data access in an appropriate amount. This is an expression of an appropriate compensation for the costs arising from data access. Compensation claims are likely to make data access unattractive for research and science and to make data access more difficult. Section 5a NetzDG, therefore, provides that costs may not constitute a significant obstacle to the use of the data access claim. They are determined following Section 287 (1) ZPO. In addition, a maximum limit of 5,000 euros applies.

The research clauses examined hardly specify the period within which data access must be granted. Only the Digital Services Act (DSA) draft contains a requirement in Article 31 DSA that data access must be guaranteed within a "reasonable period". At least for the public law access provisions that provide for a discretionary decision in the legal consequence, it is also possible to fall back on the general

principles according to which a proper exercise of discretion requires access to be granted within a reasonable time.⁵⁸ It should be noted that data access requests can vary greatly in scope and urgency, which is why the standardisation of general deadlines is generally not possible. On the other hand, the researcher must have an indication of the point in time at which he or she can file an action without the risk of an immediate acknowledgement with negative cost consequences for him. If the researcher is not to be required to set deadlines in each individual case, it is essential to set deadlines that are both variable and precise. At the same time, however, care must be taken to allow the data subject to extend the deadline depending on the type and scope of the data access request, so as not to place an unreasonable burden on the data subject.

Concerning the distribution of the burden of proof, it should be noted that this essentially follows the general principles of the burden of presentation and proof.

The enforcement of rights depends on the opening of legal recourse. Whether civil or administrative legal action is open is assessed according to general principles. Particularities arise only if the party obliged to provide access is a private person, but data access can only be requested by a third party, e.g. the coordinator for digital services within the framework of Article 31 DSA. If the coordinator fails to act, he or she must be obliged to take action against the private access addressee through administrative law.

VIII. Guidelines for research data access de lege ferenda

1. Health sector

From the basic legal framework, the principles of research data access and the analysis of national research clauses, sector-specific guidelines for research data access can be developed, which in the health sector can be transferred into an overall research data access ecosystem as well as a health research data access law. A mixed system of original research clauses with central data repositories (e.g. already existing central registries such as the Federal Cancer Registry), decentralised-centralised data repositories (e.g. distributed registries such as the state cancer registries, whereby distribution does not necessarily have to take place at the level of the federal states) and entirely decentralised data repositories, e.g. at health service providers, is recommended for the health sector. The data of private health service providers could also be fed into this system of centralised, decentralised-centralised and completely decentralised data repositories. In addition, the possibility of data donation should also and especially be opened up via Personal Information Management Systems (PIMS).

In the regulatory design of data access claims, a close purpose limitation to public interest research should be standardised as a rule. If this is the case, the circle of those entitled to data access need not be limited to non-commercial research. Insurance companies should be excluded from eligibility. A

necessity criterion for data access should not be provided. However, in the interest of protecting conflicting rights and interests, a protection concept should be required, irrespective of whether the data is transferred anonymously or on a personal basis. In the interest of effectively guaranteeing data access, a standardised - ideally internationally uniform - procedure should be provided for the data access request. It is also advisable to involve a body that decides on the ethical compatibility of the research, as provided for in some foreign legal systems (France, Canada, Finland, Australia). An additional application requirement should be a positive review by this Research Ethics Board (REB). In Canada, this body also decides whether the data will be used "for approved data purposes". In Australia, this function is performed by the Data Governance Board. Here, a separation between the two bodies seems sensible, because the decision on the ethical compatibility of the research project should primarily be made by ethicists, while the examination of whether the research serves the presumed public interest and whether the substantive legal requirements of the data access claim are also fulfilled is a legal question. The connection use of the data should be clearly defined. They should be pseudonymised, anonymised or deleted as soon as possible. The use of data for commercial advertising purposes should be prohibited, as should the sale of data. A complete prohibition of data transfer is also conceivable. This would reduce the risk of misuse, which is why the data access claim in favour of research and science should, in return, be more far-reaching according to the Five Safes Model than in the case of the permissibility of research data disclosure. Anonymised data should not be allowed to be de-anonymised, whereby the prohibition of de-anonymisation should also be addressed under criminal law.⁵⁹ At the same time, standards for anonymisation should be provided in the interest of legal clarity and to preserve the requirement of certainty. Data access must not unduly restrict the rights and interests of third parties. This should be provided in the form of a restriction clause modelled on Article 15 GDPR. Remuneration regulations should be limited to covering the costs of administrative activities. This should be oriented towards the Data Transparency Fees Ordinance. Deadline regulations in the administrative area should always be accompanied by sufficient coverage of personnel and material resources. At the same time, a careful examination of data access must be guaranteed. Flexible time limit regulations with an upper limit are therefore preferable to rigid time limit regulations, which cannot or can only insufficiently react to the concrete processing effort. It is recommended to follow Finland's Secondary Use of Social and Health Data Act and to make the decision on a data-permit immediately, but no later than 3 months after receipt of the complete application by the authority. Decisions on data access are made in the form of an administrative act. In the event of a negative decision, the legal remedy is therefore an action for an obligation to issue an administrative act granting access. In principle, the general rules on the burden of proof apply. However, the public interest should be presumed if the research is carried out at public research institutions and the research results (anonymised) are made available to the public. In addition to a

system of centralised, decentralised-centralised and completely decentralised storage of research data, flexible data trust structures should be provided for and placed on a legally secure basis.

2. Online business sector

For the online business sector, a model research data access clause can be derived from the analysis of research clauses, which should be standardised in a number of identified laws as derivative and original research clauses. In the online economic sector, too, access to data from private and public bodies should be guaranteed through indirect data access structures in state organisations. In concrete terms, this means that, according to Article 31 DSA, the Digital Services Coordinator or a similar body should decide on data access to relieve the private bodies of the data access decision. A limitation of the data access claim to specific research projects does not have to take place. If such a restriction is not made, however, asymmetrical regulation is recommended, i.e. addressing private-sector companies only above a certain size that precludes economic overburdening by the obligation to provide data access. For the rest, the data access requirements in the health sector can essentially be used as a model. Unlike in the health sector, however, in the online business sector, private entities are obliged to provide the data in the event of a justified data access claim. This can involve considerable effort, which must be compensated. At the same time, excessive compensation can significantly impair the effectiveness of the data access claim. Section 5a NetzDG, therefore, goes the way of a maximum limit of 5,000 EUR and otherwise leaves the appropriateness of the amount to the discretion of the court pursuant to Section 287 Code of Civil Procedure (ZPO). This seems to be the only feasible way, whereby the concrete quantification of the maximum limit per request should be determined and at best evidence-based. Both researchers and access providers must be able to appeal against the decision of the digital services coordinator. In the event of a refusal to grant data access by the access provider, the rightful claimant should be able to claim directly against the access provider without having to wait for further action by the digital services coordinator. If the digital service coordinator refuses to grant the claimant the right, the claimant can take action against the coordinator by way of an obligation action.

3. Mobility and energy sector

In the mobility sector and the energy sector, due to the lack of national and international best practice regulations, a commission of experts should first be set up, following the example of Great Britain, which will develop recommendations in particular for the data covered by a data access claim and the appropriate data access infrastructure.⁶⁰ In the mobility sector, the approaches of mobility data spaces and platforms already pursued by various bodies should be thought together.

-
- ¹ Guiding theme of the 8th Social and Economic Data Conference of the Social and Economic Data Council (RatSWD) 2020.
- ² *Savage*, Spotify wants to suggest songs based on your emotions, available at: <https://www.bbc.com/news/entertainment-arts-55839655>, last accessed on 01.12.2021.
- ³ *Heidhues/Köster/Közegi*, Steering Fallible Consumers, 2021, pp. 10-11, available at: http://www.personal.ceu.hu/staff/Botond_Koszegi/steering.pdf, last accessed 01.12.2021.
- ⁴ *Peichl/Bachmann/Riphahn*, FAZ (06.08.2021), available at: <https://www.faz.net/aktuell/wirtschaft/forschern-stehen-in-deutschland-zu-wenige-daten-zur-verfuegung-17471899.html>, last accessed on 01.12.2021; *Hevers*, Informationszugangsansprüche des forschenden Wissenschaftlers, p. 470 et seqq.
- ⁵ Coalition agreement between SPD, Bündnis 90/Die Grünen and FDP, Mehr Fortschritt wagen - Bündnis für Freiheit, Gerechtigkeit und Nachhaltigkeit, para. 615-620.
- ⁶ *Hevers*, Informationszugangsansprüche des forschenden Wissenschaftlers, Berlin 2015, p. 453.
- ⁷ *Wielsch*, Zugangsregeln - Die Rechtsverfassung der Wissensteilung, 2008.
- ⁸ Report of the German Commission Competition Law 4.0, A new competition framework for the digital economy, p. 44, available at: https://www.bmwi.de/Redaktion/DE/Publikationen/Wirtschaft/bericht-der-kommission-wettbewerbsrecht-4-0.pdf?__blob=publicationFile&v=14, last accessed 01.12.2021.
- ⁹ *Specht-Riemenschneider*, in: Federal Ministry of Justice and Consumer Protection/Max Planck Institute for Innovation and Competition, Data Access, Consumer Interests and Public Welfare, 2021, p. 402 et seqq.
- ¹⁰ On this, comprehensively *Podszun*, Handwerk in der digitalen Ökonomie, 2021, p. 76 et seqq.
- ¹¹ Inception Impact Assessment v. 28.05.2021 on the EU Commission's planned Data Act, available at: https://ec.europa.eu/info/law/better-regulation/have-your-say/initiatives/13045-Data-Act-&-amended-rules-on-the-legal-protection-of-databases_en, last accessed on 01.12.2021.
- ¹² Cf. BVerfGE 115, 205 (229) = NVwZ 2006, 1041; BVerfGE 128, 1 (56) = NVwZ 2011, 94 = NJW 2011, 441 Ls; BVerfGE 137, 185 (255 et seqq.) = NVwZ 2014, 1652.
- ¹³ *Wollenschläger*, in: von der Groeben/Schwarze/Hatje, Europäisches Unionsrecht, 7th ed. 2015, Article 17 GRC marginal no. 16; *Wollenschläger* (footnote 13), Article 16 GRC marginal no. 8; *Aplin*, in: Geiger, Research Handbook on Human Rights and Intellectual Property, 2015, pp. 421-437; *Breuer*, in: Isensee/Kirchhof, Handbuch des Staatsrechts VIII, Fundamental rights: Economy, Procedure, Equality, 3rd edition 2010, Section 171 marginal no. 38.
- ¹⁴ *Wismeyer/Herzog* NJW 2020, 288 (290): "In this respect, there are good reasons for recognising the position of the database producer as property within the meaning of Article 14"; Fechner, Geistiges Eigentum und Verfassung, 1999, p. 371.
- ¹⁵ Explicitly for cases of granting data access *Wismeyer/Herzog* NJW 2020, 288 (293).
- ¹⁶ *Wismeyer/Herzog* NJW 2020, 288 (293).
- ¹⁷ *Ehlers*, in: Ehlers, Europäische Grundrechte und Grundfreiheiten, 3rd edition 2009, Section 14 marginal no. 50; *Wissenschaftliche Dienste des Deutschen Bundestages*, Die Wissenschaftsfreiheit im Grundgesetz und in der Charta der Grundrechte der Europäischen Union, elaboration, p. 5, available at: <https://www.bundestag.de/resource/blob/420386/07891b8c2e2b3a104b0ffd0128619ba1/WD-3-149-10-pdf-data.pdf>, last accessed on 01.12.2021.
- ¹⁸ *Schlösser-Rost/Koch*, in: Brink/Wolff, BeckOK Datenschutzrecht, 36th Edition 2021, BDSG Section 27 marginal no. 15a-20.
- ¹⁹ BVerfGE 47, 327 (367); BVerfGE 35, 79 (113); see also *Jarass*, in: Jarass, EU-Grundrechte-Charta, 4th edition 2021, Article 13 marginal no. 8.
- ²⁰ BVerfGE 35, 79 (113).

²¹ *Pernice*, in: Dreier, Grundgesetz-Kommentar, 2nd edition 2004, Article 5 (3) marginal no. 27; *Wissenschaftliche Dienste des Deutschen Bundestages*, Die Wissenschaftsfreiheit im Grundgesetz und in der Charta der Grundrechte der Europäischen Union, elaboration, p. 5, available at:

<https://www.bundestag.de/resource/blob/420386/07891b8c2e2b3a104b0ffd0128619ba1/WD-3-149-10-pdf-data.pdf>, last accessed on 01.12.2021.

²² *Jarass*, in: Jarass (footnote 19), Article 13 marginal no. 8; *Bernsdorff*, in: Meyer/Hölscheidt, Charta der Grundrechte der Europäischen Union, 5th edition 2019, Article 13 marginal no. 14.

²³ *Bernsdorff*, in: Meyer/Hölscheidt (footnote 22), Article 13 marginal no. 14.

²⁴ *Jarass*, in: Jarass (footnote 19), Article 13 marginal no. 8; *Bernsdorff*, in: Meyer/Hölscheidt (footnote 22), Article 13 marginal no. 14.

²⁵ *Starck*, in: v. Mangoldt/Klein/Starck, Kommentar zum Grundgesetz, 5th edition 2005, Article 5 III marginal no. 408; *Kempen*, in: Epping/Hillgruber, BeckOK Grundgesetz, 47th edition 2021, Article 5 marginal no. 179-182.

²⁶ *Stemmer*, in: Brink/Wolff (footnote 18), GDPR Article 7 Rn. 74 et seqq.

²⁷ With regard to the health sector and epidemiological research *Pigeot/Buchner* DuD 2014, 816 (818).

²⁸ *Lauber-Rönsberg*, in: Putnings/Neuroth/Neumann, Praxishandbuch Forschungsdatenmanagement, 2021, p. 104; *Stemmer*, in: Brink/Wolff (footnote 18), GDPR Article 7 marginal no. 78.

²⁹ The sample text is available at: <https://www.medizininformatik-initiative.de/de/mustertext-zur-patienteneinwilligung>, last accessed on 01.12.2021.

³⁰ Proposal of the EU Commission (25.11.2020) for a Regulation of the European Parliament and of the Council on European Data Governance (Data Governance Act), available at: <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52020PC0767>, last accessed on 01.12.2021.

³¹ *Hummel/Braun/Augsberg/von Ulmenstein/Dabrock*, Data Sovereignty - Governance Approaches for the Health Sector, 2021, pp. 10 and 15; in contrast, *Stemmer*, in: Brink/Wolff (footnote 18), GDPR Article 7 marginal no. 78.

³² *Taeger*, in: Taeger/Gabel, GDPR BDSG, 3rd edition 2019, GDPR Article 6 Rn. 79 f.

³³ *Albers/Veit*, in: Brink/Wolff (footnote 18), GDPR Article 6 marginal no. 41; *Taeger*, in: Taeger/Gabel (footnote 32), GDPR Article 6 marginal no. 87.

³⁴ *Golla*, in: Specht/Mantz, Handbuch Europäisches und deutsches Datenschutzrecht, 2019, Section 23 marginal no. 43.

³⁵ See also *Schlösser-Rost/Koch*, in: Brink/Wolff (footnote 18), BDSG Section 27 marginal no. 35 et seqq.

³⁶ *Taeger*, in: Taeger/Gabel (footnote 32) GDPR Article 6 marginal no. 98.

³⁷ *Schulz*, in: Gola, Datenschutz-Grundverordnung, 2nd edition 2018, Article 6 Rn. 58; *Specht*, in: Specht/Mantz (footnote 34), Section 9 marginal no. 50.

³⁸ *Albers/Veit*, in: Brink/Wolff (footnote 18), GDPR Article 6 marginal no. 46; cf. the BVerwG ruling of 27.09.2018 - 7 C 5/17.

³⁹ *Taeger*, in: Taeger/Gabel (footnote 32), BDSG Section 27 marginal no. 7.

⁴⁰ See also *Schlösser-Rost/Koch*, in: Brink/Wolff (footnote 18), BDSG Section 27 marginal no. 35 et seqq.

⁴¹ *Taeger*, in: Taeger/Gabel (footnote 32), BDSG Section 27 marginal no. 8.

⁴² *Taeger*, in: Taeger/Gabel (footnote 32), BDSG Section 27 marginal no. 9; cf. also *Schlösser-Rost/Koch*, in: Brink/Wolff (footnote 18), BDSG Section 27 marginal no. 31 et seqq.

⁴³ Thus *Lauber-Rönsberg*, in: Putnings/Neuroth/Neumann (footnote 28), p. 106.

⁴⁴ In addition, reference can also be made, for example, to Section 303e (1) (8) SGB V, which, as a special provision, can also be made the basis for data processing, namely the transfer of data by the research data centre.

⁴⁵ *Taeger*, in: Taeger/Gabel (footnote 32), GDPR Article 6 marginal no. 154.

⁴⁶ A strategy for a Modern Digitised Energy System, Energy Data Taskforce report, 2019, p. 25, available at: <https://esc-non-prod.s3.eu-west-2.amazonaws.com/2019/06/Catapult-Energy-Data-Taskforce-Report-A4-v4AW-Digital.pdf>, last accessed on 01.12.2021.

⁴⁷ A strategy for a Modern Digitised Energy System, Energy Data Taskforce report, 2019, p. 25, available at: <https://esc-non-prod.s3.eu-west-2.amazonaws.com/2019/06/Catapult-Energy-Data-Taskforce-Report-A4-v4AW-Digital.pdf>, last accessed on 01.12.2021.

⁴⁸ *Desai/Ritchie/Welpton*, Five Safes: Designing data access for research, Economics Working Paper Series 1601, University of the West of England, 2016, p. 5, available at: <https://www2.uwe.ac.uk/faculties/bbs/Documents/1601.pdf>, last accessed on 01.12.2021.

⁴⁹ <https://www.ukdataservice.ac.uk/manage-data/legal-ethical/access-control/five-safes>, last accessed on 01.12.2021.

⁵⁰ *Wilkinson et al.* Nature: Scientific Data 3, 2016, 1 (3); on the FAIR Data Principles in the context of data sovereignty considerations, cf. *Hummel/Braun/Augsberg/von Ulmenstein/Dabrock*, Datensouveränität - Governance-Ansätze für den Gesundheitsbereich, 2021, p. 29.

⁵¹ *Wilkinson et al.* Nature - Scientific Data, 2016, p. 3, available at: <https://www.nature.com/articles/sdata201618.pdf>, last accessed on 01.12.2021; *Hummel/Braun/Augsberg/von Ulmenstein/Dabrock*, Data Sovereignty - Governance Approaches for the Health Sector, 2021, p. 29.

⁵² Cf. BT-Drs. 19/29392, p. 18.

⁵³ BT-Drs. 19/29392, p. 18.

⁵⁴ Terminology based on the ladder of permissions in *Ohly*, Volenti non fit iniuria - Die Einwilligung im Privatrecht, 2002, p. 141 et seqq.

⁵⁵ From the research point of view *Hevers* also argues in favour of bound decisions (footnote 6), p. 453.

⁵⁶ Cf. for example the explanatory memorandum to Section 8 BKRG, which speaks of a public interest in knowledge, BT-Drs. 19/28185, p. 43.

⁵⁷ For example, Section 303e (1) (8) SGB V.

⁵⁸ *Hevers* (footnote 6), p. 445.

⁵⁹ *Specht* GRUR Int. 2017, 1040 (1046-1047); Expert Opinion of the Data Ethics Commission, 2019, p. 132, available at: https://www.bmi.bund.de/SharedDocs/downloads/DE/publikationen/themen/it-digitalpolitik/gutachten-datenethikkommission.pdf?__blob=publicationFile&v=6, last accessed 01.12.2021.

⁶⁰ <https://www.gov.uk/government/groups/modernising-energy-data>, last accessed 01.12.2021.