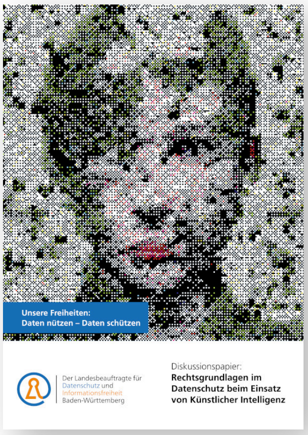
Der Landesbeauftragte für
Datenschutz und
Informationsfreiheit
Baden-Württemberg

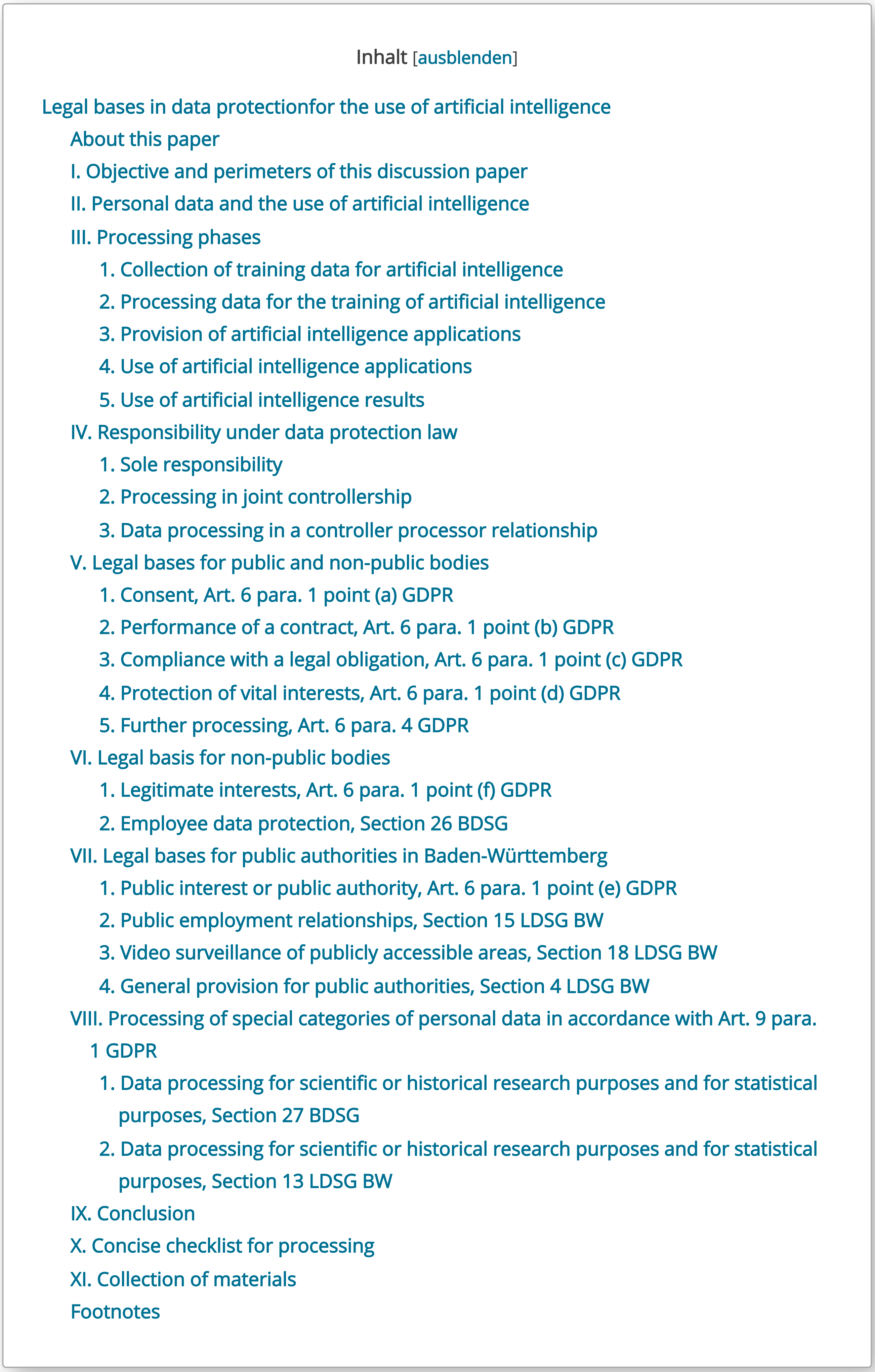
Legal bases in data protection for the use of artificial intelligence

When and how may personal data be processed for the training and application of artificial intelligence?

– Discussion paper. Version 1.0 from 07.11.2023 –



zur deutschen Version



Inhalt [ausblenden]

About this paper

This discussion paper was originally drafted in German and published with the title *Rechtsgrundlagen im Datenschutz beim Einsatz von Künstlicher Intelligenz* on November 7, 2023. This paper is a non-official translation; references and links to external websites, especially in the footnotes, may lead to information in German.

If you have any questions or suggestions regarding this paper, please contact us at ✉ pressestelle@lfdi.bwl.de.

I. Objective and perimeters of this discussion paper

The use of artificial intelligence offers enormous potential for our society. As with any new technology, it also comes with challenges. Following the motto saying “use data, protect data”, data protection and artificial intelligence should be considered together from the outset to strengthen the civil freedoms of citizens and to enable an integration of data protection in innovations. The State Commissioner for Data Protection and Freedom of Information Baden-Württemberg (hereinafter: LfDI BW) would like to contribute to the utilisation of this potential with this discussion paper.

This paper is intended to help controllers in Baden-Württemberg to familiarise themselves with the legal bases that data protection law provides for the use of artificial intelligence systems[1] (hereinafter referred to as AI). The starting point is the current law. The provisions of the European Union’s Regulation laying down harmonised rules on Artificial Intelligence (hereinafter: AI Regulation)[2], which have not yet been finally negotiated, are briefly addressed.

The term “discussion paper” is intended to emphasise that this paper is not a final definition – also with regard to individual aspects – but rather is intended to reflect the current state of discussion. Together with a further collection of materials at the end, the discussion paper is ultimately to be understood as a working guideline to better locate specific application scenarios within the legal framework.

With this “living document”, we hope to create added value for companies and associations (non-public authorities), as well as for authorities (public bodies) by explaining key terms, providing an overview of the legal bases in the General Data Protection Regulation (hereinafter: GDPR), the Federal Data Protection Act (hereinafter: BDSG), and the Baden-Württemberg State Data Protection Act (hereinafter: LDSG BW), and facilitating legal evaluation by formulating key questions.

The discussion paper as well as the checklist at the end of the summary are not exhaustive. The considerations also within the perspective of data protection law are thematically limited. The subject matter only covers the permissions required for the processing of personal data, which may come into play when using AI[3]. The topics of transparency and non-discrimination of the systems (keyword: "bias"), data subject rights (Art. 12-22 GDPR), data security and data protection by design (Art. 25, 32 GDPR), data protection impact assessment (Art. 35 GDPR), and specific challenges posed by data transfers outside the European Union are expressly not addressed and may be relevant for further discussion papers.

This discussion paper contains a version number and the date of the latest version to indicate its editing and updating status.

II. Personal data and the use of artificial intelligence

The GDPR does not contain specific provisions for AI, which is why the definition of personal data is based on the general definition in Art. 4 no. 1 GDPR.[4] According to this, personal data means any information relating to an identified or identifiable natural person. The wording of this provision shows that the concept of personal data is based on a broad understanding.[5]

It may therefore be sufficient to create a personal identifiability if the natural person can be identified on the basis of further (additional) information, see Art. 4 no. 1, clause 2 GDPR. Identifiability can result from the fact that numbers are used to code and represent data.[6] Alongside this, all means that are generally likely to be used for this should be taken into account, see Recital 26 sentence 3. Considering this background, the technology available at the time of processing and technological developments should also be considered. However, the assumption of a personal identifiability requires an existing or potential possibility of access to the additional information necessary for the identification of the natural person.[7] The information required for identification does not necessarily have to be in the hands of a single organisation.[8]

The question of personal identifiability is also of particular importance for AI, since this stipulates the applicability of the GDPR. Thus, the primary question is whether and to what extent a fully trained AI model[9] permits the identifiability of natural persons now and in the future.[10]

Such a personal identifiability could arise, e.g., from the fact that the AI model itself contains the personal data. However, indirect identifiability could also be likely outside the direct storage of personal data. In this context, attacks referred to as model attacks must be considered.[11] For example, so-called membership inference attacks aim at discovering which personal data was in the training data to derive characteristics of natural persons. In contrast, model inversion attacks directly attempt to obtain information about the training data from the learning results of the model. If such attacks on AI are possible, the model itself could in turn be regarded as personal data.[12]

Likewise, whether such model attacks are generally considered to be likely, must also be taken into account.[13] This means that recurring risk assessments are required. In addition to the legal evaluation of the (re-)identifiability of natural persons, this must include the technical methods in terms of data protection-compliant technology design, see Art. 25 para. 1 GDPR. This includes preventive measures in the technical design of an AI in such a way that, e.g., model attacks are avoided, which is why the “differential privacy method” is being discussed.[14] The technical method of “unlearning“[15] can also be used for erasure and the right to be forgotten, see Art. 17 GDPR.

> **Brief:**
> The extent to which AI process personal data depends on the point in time of the evaluation: It may be likely that natural persons can be identified from the outset or only at a later point in time with additional information. In each case, the machine learning methods used must be analysed, as well as the probability that natural persons can be (re-) identified through atypical influence on the systems.

III. Processing phases

Since the term “processing” in Art. 4 no. 2 GDPR covers nearly every process in connection with personal data, processing operations relevant under data protection law in the context of artificial intelligence can be correspondingly diverse. Five processing phases are being presented below as examples.

1. Collection of training data for artificial intelligence

The collection, generation, structuring, or categorisation of training, testing, and application data is a regular process at the beginning of artificial intelligence applications. This can result from the independent collection of personal data, such as the creation of image data with a camera, as well as the downloading of data from publicly accessible sources, particularly the internet.

2. Processing data for the training of artificial intelligence

A further processing step can be the production or development of an AI. Here, the personal data is processed for the initial training of the AI. Improving or specifying (so-called fine-tuning) the AI, for which the same personal data is further or repeatedly processed to increase the quality of the results of the AI, is also to be included under the term ‘processing’.

3. Provision of artificial intelligence applications

Whether the provision of AI trained with personal data constitutes the processing of personal data requires a differentiated evaluation. On the one hand, the free and account-based provision of such applications may constitute the processing of personal data previously used for their development and further enhancement. Whether this is the case depends on the extent to which the training data can be regarded as still “contained” within the AI, since it is further processed when the application is used.[16] On the other hand, personal data collected during the use may also be further processed by the AI, particularly through further training of the application. Such processing would require a separate legal basis.

4. Use of artificial intelligence applications

According to the so-called double-door model of data protection law, a separate legal basis is required for each legal relationship when processing personal data with multiple parties. This means that processes in the third processing phase must be evaluated from the perspective of both the provider and the user.[17]

Insofar as an AI is designed in such a way that its training data is further processed each time it is used, the legal basis with which the users can access this data must be examined. This legal basis would also include the subsequent processing by means of AI use.

5. Use of artificial intelligence results

Finally, the output created by AI applications may also constitute processing that is relevant under data protection law. This applies, e.g., to cases in which personal data is output in the text of a language model or images of real people are generated using an AI image generator.

At this level, special examination and attention must be paid to the processing operation in which the personal identifiability to an AI result is only established by the organisation using the AI (e.g., a draft text is provided with salutation and address data, or a diagnosis is accepted and stored). In this case, the legal basis must also refer to the fact that the AI result is linked to a natural person and a new processing operation with a new risk to the rights and freedoms of natural persons is created

.

> **Brief:**
> The processing operations relevant under data protection law must be differentiated with regard to the aforementioned phases of processing in connection with the AI and evaluated under data protection law from the perspective of the providers, users, and data subjects.

IV. Responsibility under data protection law

1. Sole responsibility

The controller is obligated to comply with the data protection principles of lawfulness, fairness, transparency, purpose limitation, data minimisation, accuracy, storage limitation, integrity, and confidentiality of the processing of personal data, cf. Art. 5 para. 1 GDPR. They are also accountable and must demonstrate fulfilment of the aforementioned requirements and their application to the AI.

In general, persons, companies, authorities, or other bodies involved in artificial intelligence can be considered data controllers under data protection law.

According to Art. 4 para. 1 no. 7 GDPR, the controller is the person who alone or jointly with others determines the purposes and means of processing personal data. The decision on “purposes” and “means” can also be understood as a decision on the “why” and “how” of processing.[18] This may include, e.g., processes in which an organisation develops, provides, or uses an AI.

2. Processing in joint controllership

Joint controllership exists in accordance with Art. 26 para. 1 sentence 1 GDPR if two or more controllers jointly decide on the purposes and means of processing. Thus, at least two parties must work together. They may jointly decide about the processing of personal data. Alternatively, complementary decisions can also lead to joint controllership if the respective decisions have a significant impact on the determination of the purposes and means of the processing. Another important criterion for joint controllership is that the processing would not be possible without the involvement of both parties, in the sense that the processing operations of each of the parties are inextricably linked.[19]

Joint controllership might be conceivable e.g., if data sets from two companies were used to train a joint AI.

In accordance with Art. 26 para. 1 sentence 2 GDPR, joint controllers shall transparently determine in an agreement[20] who is responsible for the fulfilment of the rights of the data subject and who complies with which information obligations in accordance with Articles 12, 13, and 14 GDPR.[21]

KEY QUESTIONS:

- Are there several parties who have jointly decided on the processing of personal data?
- Is there at least one supplementary decision by several parties that has a significant impact on the purposes and means of processing?
- Would processing be possible without the involvement of one of the parties?

3. Data processing in a controller processor relationship

In an effective a controller processor relationship, the processor receives the legal basis for processing the personal data from the client. In this case, processors do not need an own legal basis for their processing.

In accordance with Art. 4 no. 8 GDPR and Art. 28 GDPR, a controller processor relationship is deemed to exist if an organisation processes personal data on behalf of the controller. The processor is therefore an independent body in relation to the controller and is at the same time bound by the instructions of the controller.[22] According to Art. 28 para. 1 GDPR, the controller may only work with processors who can sufficiently guarantee through appropriate technical and organisational measures that processing is carried out in accordance with the requirements of the GDPR and ensure the protection of the data subject right. Processing by a processor is carried out on the basis of a contract with the controller[23] or another legal instrument, as per Art. 28 para. 3 sentence 1 GDPR.

Data processing in the framework of a controller processor relationship is possible to the extent that an AI is trained with personal data under the instructions of the controller and exclusively for the purposes of the controller. In addition, a controller processor relationship comes into consideration if a controller uses an existing AI made available online by a cloud service provider for the processing of personal data, e.g., for diagnostics. In this case, however, the limit of a mere controller processor relationship is exceeded if the personal data that is entered into the AI also benefits the improvement of the application, and the processing therefore also serves the provider's own purposes.

> **Brief:**
> In the context of the use of AI, a differentiated factual analysis may be necessary to determine responsibility under data protection law. In this context, whether personal data is only processed based on instructions and on behalf of another organisation and whether a party has its own interest in the processing of personal data must be assessed.

KEY QUESTIONS:

- What has been agreed between the parties and how is the actual processing of personal data organised?
- Is the agreement as concluded between the parties reflected in the actual processing?

V. Legal bases for public and non-public bodies

At first, the legal bases of the GDPR for the processing of personal data, which are applicable to both public and non-public bodies, are being presented.

1. Consent, Art. 6 para. 1 point (a) GDPR

According to Article 6 para. 1 point (a) GDPR, the processing of personal data is lawful if the data subject has given their consent to the processing of their personal data for clearly defined purposes.

According to the definition in Art. 4 no. 11 GDPR, consent is defined as "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". In addition to a voluntary and unequivocal indication of will, this particularly requires that the consent is given in a sufficiently specific and informed manner.[24]

In order for consent to be sufficiently specific, it has to be specified which data, among other things, is processed for which purposes and by whom, what the type of data processing is, and who the recipients are.[25] This enables the data subject to check whether they wish to give their consent to data processing. The specific requirements for certainty must be determined for each individual case while taking the respective intensity of the data processing into account.[26] Closely linked to this is the requirement for an informed declaration, according to which the data controller must inform the data subject of the key aspects of the data processing.[27] According to Recital 42 sentence 4, this at least includes information about the controller and the purposes of the data processing.
Compliance with data protection requirements for consent-based data processing by AI can pose a challenge in practice.[28] For example beacause of the revocability of consent in accordance with Art. 7 para. 3 sentence 1 GDPR. If the data subject exercises their right of withdrawal, the controller must immediately delete their personal data in accordance with Art. 17 para. 1 point (b) GDPR if there is no other legal basis for the data processing.[29] Under certain circumstances, this can have an impact on the functionality of the AI if it was trained on the basis of this data or if separating the data records concerned to fulfil the erasure obligation would involve disproportionate effort.[30]

Another difficulty can be the lack of transparency and traceability of complex AI,[31] if this raises questions about compliance with data protection requirements in the form of a sufficiently specific and informed declaration of consent.[32] The information must be given in a precise, comprehensible, and easily accessible form in clear and simple language so that the data subject can understand how the data processing works.[33]

It can be particularly challenging for data controllers to fulfil this requirement when even experts can no longer clearly understand the AI and their data processing due to their complexity and architecture (e.g., when using deep neural networks).

However, the lack of transparency and traceability can be prevented to a certain extent by at least providing the data subject with information on the essential aspects of data processing – such as information on the purposes of data processing and the identity of the controller (e.g., in the data protection notices).[34]

KEY QUESTIONS:

- Is there an informed, clearly confirming declaration of consent from the data subject?
- In which phase of the processing should consent be given? When using the (complex) AI, can the data subject be informed in such a way that they can evaluate the effects and scope of the data processing?
- Could the functionality of the AI be jeopardised if data subjects withdraw their consent and exercise their right to erasure? Can an exercise of their rights presently be guaranteed at all?

2. Performance of a contract, Art. 6 para. 1 point (b) GDPR

Art. 6 para. 1 point (b) alt. 1 GDPR permits the processing of personal data to the extent that it is necessary for the performance of a contract to which the data subject is party.[35] On the other hand, processing for the performance of pre-contractual measures in accordance with Art. 6 para. 1 point (b) alt. 2 GDPR may be permissible.

A common requirement is that there must be a specific contractual or pre-contractual relationship between the parties affected by the data processing.[36] It is hence not sufficient that processing is merely mentioned in a contract or is useful for its performance.[37]

The fact alone that data processing is included in a user agreement does not make it lawful.

The processing of third parties' personal data who are not party to a contract or pre-contractual measure is not covered by the legal basis. This means that parties which produce, provide, or use an AI cannot conclude a contractual agreement that legitimises the processing of third parties' personal data.

For example, if a person allows an AI speech generator to be created that is trained with their speech, Art. 6 para. 1 point (b) GDPR would appear to be a conceivable legal basis for the processing of the speech data required for this. However, the use of the speech data provided to further improve a basic AI model would at best be useful for the performance of the contract and will therefore not be covered by Article 6 para. 1 point (b) GDPR.

Similarly, the use of AI within a medical treatment could be covered by the legal basis in accordance with Art. 6 para. 1 point (b) GDPR in conjunction with Section 630a para. 1 of the German Civil Code as part of the primary contractual obligation. This is because the processing of personal data associated with the use of the AI for diagnostic support to fulfil the treatment contract would be covered by this legal basis under certain circumstances. At the same time, the use of the AI would need to be reasonably expected from the data subjects' perspective and the data subjects would need to be informed about the functionality of the AI.

KEY QUESTIONS:

- Is there a contractual relationship to which the data subject is a party, or is a pre-contractual relationship established at the request of the data subject?
- Is the processing an objectively necessary component in the sense that the primary object of the (pre-)contractual relationship would cease to exist without this component?
- Is the processing merely useful for the (pre-)contractual purposes and therefore not necessary for these purposes?

3. Compliance with a legal obligation, Art. 6 para. 1 point (c) GDPR

The processing of personal data may be lawful under Art. 6 para. 1 point (c) GDPR.[38] Such legal obligation needs to be a genuine statutory obligation, i.e., a "must" for data processing.

The controller generally has no freedom of choice in this matter. In addition, there is a stricter requirement with regard to the legal basis, as well as the necessity to limit the processing within this context to what is "absolutely necessary". The legal basis of Art. 6 para. 1 point (c) GDPR offers a limited scope of application in the context of AI.

KEY QUESTIONS:

- Is there an obligation that requires data processing without leeway in the decision-making for the controller?
- Is the processing truly limited to what is absolutely necessary to fulfil the legal obligation?

4. Protection of vital interests, Art. 6 para. 1 point (d) GDPR

The following constellation shows why the distinction between training and application can be significant in terms of the legal basis. The processing of personal data for the training of the AI in the form of collecting, generating, structuring, or categorising data for the protection of vital interests cannot form a legal basis due to the short-term emergency situation that this provision requires. This is because this legal basis only applies to emergency situations with regard to protection against a specific threat to physical integrity and life, meaning that it is subsidiary, see Recital 46 sentence 2 GDPR. This legal basis serves the need to create a legal basis for processing personal data that is (vitally) necessary in the short term. This must concern the vital interests of the data subject and less restrictive means must not be possible. At present, it is hardly conceivable that this could require the training of an AI.

The question of whether the use of an AI by the controller with the input of personal data of the data subject (without the AI using the input data for training) can be based on Art. 6 para. 1 point (d) GDPR could be evaluated differently in individual cases. Use cases – for example in the case of life-saving use with an AI in the context of an unresponsive emergency patient – appear imaginable which would also be subject to the requirements of Art. 9 para. 2 point (c) GDPR

> **Brief:**
> In principle, this legal basis can only be considered for the use of artificial intelligence in emergency situations for short-term measures to protect the vital interests of the data subject.

5. Further processing, Art. 6 para. 4 GDPR

Article 6 para. 4 GDPR regulates the further processing of personal data in the event of a change of purpose. This means processing for a purpose other than that for which the personal data was originally collected. The provision is particularly important regarding training AI if the underlying training data was previously collected for a different purpose (e.g., to fulfil a contract) and is now to be used for training.[39] The crucial question is then whether the further processing of the personal data is permitted under Art. 6 para. 4 GDPR.[40] A further legal basis is also required (according to the controversial legal opinion)[41] for further processing for a new purpose within the meaning of Art. 6 para. 4 GDPR and Art. 5 para. 1 point (b) GDPR.

KEY QUESTIONS:

- For what purpose was the training data originally collected and is it now to be processed for a different purpose?
- Is further processing for another purpose compatible with the original purpose in accordance with the requirements of Art. 6 para. 4 GDPR?
- Is there a legal basis for further processing?

VI. Legal basis for non-public bodies

The following legal standards with regard to Art. 6 para. 1 point (f) GDPR as a legal basis for the processing of personal data in connection with artificial intelligence are generally only applicable to non-public bodies.

According to Art. 6 para. 1 subpara. 2 GDPR, the legal basis of Art. 6 para. 1 point (f) GDPR does not apply to processing by public authorities in the fulfilment of their tasks. The provision is therefore not applicable for public authorities, such as municipalities in this respect.[42]

1. Legitimate interests, Art. 6 para. 1 point (f) GDPR

According to Article 6 para. 1 point (f) GDPR, the processing of personal data is lawful if it is necessary for the purposes of the legitimate interests pursued by the controller or by a third party, except where such interests are overridden by the interests or fundamental rights and freedoms of the data subject which require protection of personal data, particularly when the data subject is a child.

In the area of data processing by AI, the legal basis in accordance with Art. 6 para. 1 point (f) GDPR is likely to be of particular importance. This is mainly because the provision offers a certain degree of flexibility due to its wording which is formulated openly (to innovation): It is based on a so far unspecified legitimate interest of the controller or a third party. Furthermore, unlike other legal bases, it does not require any further legal determination in European or national law. However, the open wording can also lead to a certain legal uncertainty in practice. In the case of more complex processing operations, many circumstances can influence the balancing process inherent in the provision. Because data subjects do not expect their data to be processed in every situation, this can lead to unpredictability for the data subjects, as well as legal uncertainty for the controller.[43]

According to the case law of the ECJ[44], processing of personal data in accordance with Art. 6 (1) (f) GDPR is lawful under three cumulative conditions:

1) Legitimate interest

The term “legitimate interest” is understood broadly.[45] Accordingly, the legitimate interest pursued by the controller may in principle consist of any legal, economic, or non-material interest of the controller or third parties.[46]

In the development and use of AI, a legitimate interest of the controller can at first be assumed. A legitimate interest may exist, e.g., in the development of AI. In a commercial context, data controllers will regularly pursue the goal of offering continually improved and more innovative products, which, e.g., may be the development of autonomous vehicles or the error-free recognition of human interactions.

A legitimate interest for the production, provision, or use of AI could also arise from the interests expressly mentioned in the General Data Protection Regulation, such as fraud prevention or direct marketing.[47]

2) Necessity

In the context of necessity, the controller must examine whether the legitimate interest in processing the data cannot be “reasonably and equally effectively met by other means that are less intrusive on the fundamental rights and freedoms of data subjects”.[48] Processing in the context of AI must therefore be compared with possible alternatives, particularly those that process less data.

If, e.g., the development of an AI at the time of evaluation is also possible without personal data or with anonymised data (and therefore does not allow any conclusions to be drawn about individual persons), the (more intrusive) processing of personal data is not necessary. Particularly with regard to training data, the question therefore always arises as to whether personal data needs to be processed.

The evaluation of necessity also takes the principle of data minimisation as per Art. 5 para. 1 point (c) GDPR into account, which requires, among other things, that personal data is not processed beyond what is necessary.[49] Simply put, when handling personal data in connection with AI, the underlying principle is not “the more data the better”, but rather to stick to the principle that only what is strictly necessary (in relation to the respective processing purpose) shall be processed.

3) Balancing

Finally, the interests or fundamental rights and freedoms of the data subject must not take precedence over the legitimate interests of the controller or a third party.[50] The balancing of the respective opposing rights and interests generally depends on the specific circumstances of the individual case.[51] Among other things, the scope of the processing in question and its impact on the data subjects must be considered.[52] Other criteria may include, e.g., the information value and the number of data subjects.[53] It is also important to consider whether the data subject is a child who requires special protection.[54]

The processing of personal data must generally be proportionate in relation to the legitimate interest with regard to all balancing criteria. A crucial element is also whether the data subject can expect their personal data to be processed in the specific situation, see Recital 47 sentence 1 cl. 2, sentence 3, 4. The processing may be prohibited if this is not the case.[55] The question of the extent to which a controller fulfils their transparency and information obligations is as well likely to have an indirect effect on the balancing process to be carried out.

In the case of data processing by AI, in addition to the level of detail and scope of the training data, circumstances such as the effect on the data subjects or the guarantees to ensure proper training must also be included in the balancing of interests.[56] The level of the interference depends on the specific processing. The training of a so-called large language model could have a greater interference with the data subject rights than the training of a traditional statistical model (e.g., Generalised Linear Mixed Models). Furthermore, it also depends on the category of data to be processed (e.g., processing of special categories of personal data, Art. 9 para. 1 GDPR, for which a legal basis in accordance with Art. 9 para. 2 GDPR is also needed in addition to Art. 6 para. 1 point (f) GDPR).[57]

> **Brief:**
> Overall, Art. 6 para. 1 point (f) GDPR is a particularly suitable legal basis for most processing operations in the AI context due to its openly phrased conditions. However, due to the mandatory balancing of interests, this provision can only provide legal certainty to a limited extent, since it will always be necessary to comprehensively evaluate the specific individual case.

KEY QUESTIONS:

- What is the specific legitimate interest in the processing of personal data? Could the goal pursued also be achieved without processing personal data (e.g., by using synthetic training data as part of the development or production of an AI or by sufficiently disassociating the data used for training)? Is the respective procedure necessary at all or can the purposes pursued not also be achieved in other ways? (For example, AI-based facial recognition would not be necessary for a simple view of the door using a doorbell camera).
- Do the interests of the data subjects outweigh the interests of the controller?
- Based on a specific situation, is it foreseeable to the data subjects that their personal data may be processed?

2. Employee data protection, Section 26 BDSG

The use of artificial intelligence is also becoming increasingly significant in the working environment. The day-to-day work of many employees is characterised by data-based applications. While in the employment context Section 26 BDSG offers the possibility of processing personal data in employment relationships on the basis of the opening clause in Art. 88 GDPR, the applicability of Section 26 para. 1 sentence 1 BDSG is controversial following the ruling of the European Court of Justice of 30 March 2023.[58]

In the employment environment, however, data processing for the performance of the employment contract, or during a job application process with the data subject, can generally be directly legitimised with the legal basis of Art. 6 para. 1 sentence 1 point (b) GDPR without using the legal basis of Section 26 para. 1 sentence 1 BDSG.

However, this presupposes that the data processing with an AI is necessary for the performance of the employment relationship or the application process. There must not be an applicable reasonable and comparably effective data protection-friendly alternative for this processing. In addition, the interest in using artificial intelligence must outweigh the interest of the data subject. The use of AI can also be regulated in more detail through collective agreements, although this must not undermine the level of protection provided by the GDPR. Alongside this, the principles of works constitution law and collective bargaining law must of course be observed, particularly when using AI in the employment context.

It should also be noted that, in terms of employee data protection, strict standards must be applied when examining consent[59] adue to the subordination relationship. According to the wording of Section 26 (2) BDSG, the employee's level of dependency on the company must be considered when examining freely given consent. With regard to the practically significant issues of AI analyses in the application process or "HR management", it should be noted that a remarkable depth of analysis can be achieved through personality profiles. This means that consent might be ruled out as a legal basis in such a case. In addition, the requirements of Art. 22 GDPR may also need to be observed.[60]

> **Brief:**
> In the context of employee data protection, strict standards must be applied to the examination of consent due to the subordination relationship.

KEY QUESTIONS:

- What is the legal basis for the use of the AI in the employment environment?
- Is the use of AI necessary for the performance of the employment relationship or the job application process?
- Is there a reasonable, effective, and more data protection-friendly alternative to using AI?
- Do the employees' interests outweigh the interests of the controller?
- Have the principles of works constitution law and collective bargaining law been considered?
- In the case of consent as a legal basis: Is the consent to be regarded as voluntarily given, even when considering the subordination of the relationship?

VII. Legal bases for public authorities in Baden-Württemberg

In accordance with Art. 6 para. 1 point (e), para. 3 GDPR, the legal bases of the LDSG BW could also be taken into account for public authorities of the state when using artificial intelligence. The special regulations in the areas of law enforcement and justice within the Law Enforcement Directive[61] and the State Data Protection Act for Justice and Fines Authorities (LDSG-JBBW) are not subject of this discussion paper.

1. Public interest or public authority, Art. 6 para. 1 point (e) GDPR

The general wording of Art. 6 para. 1 point (e) GDPR lists two possibilities for processing. Either the processing has to be necessary for the performance of a task carried out in the public interest or in the exercise of official authority. In both cases, however, a task must be delegated to the controller. According to Art. 6 para. 3 GDPR, nevertheless, a legal basis in Union law or Member State law is required. The provision of Art. 6 para. 1 point (e) GDPR does not create a legal basis for the processing of personal data, but is only valid in conjunction with, e.g., the legal bases in national and state law specifically presented here.[62]

2. Public employment relationships, Section 15 LDSG BW

Similar to Section 26 BDSG, Section 15 LDSG BW constitutes a legal basis for data processing in the public employment environment. As with Section 26 para. 1 sentence 1 BDSG, data processing must be necessary to enter into, perform, or terminate the respective employment relationship or to carry out internal planning, organisational, personnel, social, or budgetary and cost accounting measures.

However, a strict examination against the background of a subordinate relationship with the public authority could partially rule out the use of AI in the employment context on the basis of Section 15 LDSG BW. According to Section 84 of the State Civil Service Act, which also applies to the processing of personnel file data of employees and trainees in the public sector in accordance with Section 15 para. 4 LDSG BW, a decision under civil service law may only be based on the exclusively automated processing of personal data if there is neither discretion nor room for assessment.[63]

> **Brief:**
> The subordinate relationship of public service employees could prevent the use of AI on the basis of Section 15 LDSG BW. Nevertheless, a vigorous and comprehensive examination is necessary.

KEY QUESTIONS:

Is there discretion or room for assessment in the decision to be made? If this is the case, the use of an AI is not permitted. If this is not the case, the following questions are to be answered:

- What is the legal basis for the use of AI-supported processing operations in public employment relationships? (legal basis or collective agreement)
- Is the use of AI necessary for the performance of the public employment relationship or the job application process?
- Is there a reasonable and effective data protection-friendly alternative to the use of AI?

- Do the interests of the data subjects outweigh the interests and rights of the controller?

3. Video surveillance of publicly accessible areas, Section 18 LDSG BW

Whether Section 18 LDSG BW can be considered as a legal basis for the use of artificial intelligence in connection with a video surveillance system must be examined more closely taking into account the purpose of the provision.

Monitoring of publicly accessible areas with the aid of a video surveillance system by public authorities in accordance with Section 18 para. 1 LDSG BW is only permitted if this is necessary to fulfil public tasks or in the exercise of domiciliary rights in individual cases: to protect the life, health, freedom, or property of persons (alt. 1), or cultural assets, public facilities, public transport, official buildings and to protect other public locations (alt. 2). There must be no recognisable indications of an overriding interest on the part of the data subjects worthy of protection.

According to the purpose of this provision, Section 18 LDSG BW clearly also serves to protect against criminal behaviour. However, Section 18 LDSG BW is not applicable to local authorities acting as local police authorities and law enforcement for policing purposes. While they are public authorities, Section 44 of the Baden-Württemberg Police Act (PolG BW) regulates video surveillance by the police as a more specific standard, see also Section 2 para. 1 sentence 3 LDSG BW. The content of Section 44 para. 4 PolG BW also raises doubts about the permissibility of the use of AI within the scope of Section 18 LDSG BW. This is because Section 44 para. 4 PolG BW allows law enforcement (but not the local police authority) to analyse video images collected using AI in accordance with Section 44 para. 1 PolG BW. The state legislator has thus expressed that it does not consider the provision for the collection of video images (Section 44 para. 1 PolG BW) to be sufficient to also support an algorithmic evaluation of the image data for the detection of potentially criminal behaviour. Instead, it considered an explicit legal basis in Section 44 Para. 4 PolG BW to be necessary. The reverse conclusion can be drawn for Section 18 LDSG BW: If the legislator had wished to allow algorithmic detection of certain human behaviour for other public authorities, it likely would have also regulated this in Section 18 LDSG BW.

Any video surveillance constitutes an interference of the right to informational self-determination.[64] Therefore, whether the use of artificial intelligence in the context of video surveillance can be proportionate in accordance with Section 18 LDSG BW is questionable as well. Such use would need to serve a legitimate purpose and be suitable, necessary, and appropriate.

Its use generally serves the purpose of supporting the surveillance of publicly accessible areas and is suitable for this purpose. However, whether the use of AI-based video surveillance systems can also be necessary and appropriate is questionable. Particularly considering that the use of AI-based video systems in public spaces should be the mildest means available to achieve the purpose and that the interests of the responsible authority should not be disproportionate to the interests of the data subjects, we assume that Section 18 LDSG BW cannot be a suitable legal basis for the processing of personal data in connection with intelligent video surveillance technology in public spaces.

> **Brief:**
> Under the current legal framework, Section 18 LDSG BW is not a suitable legal basis for the use of artificial intelligence in the context of video surveillance.

4. General provision for public authorities, Section 4 LDSG BW

In principle, the provision in Section 4 LDSG BW can be considered as the legal basis for public authorities to process personal data in an AI. However, it has to be noted that as a generally phrased provision, Section 4 LDSG BW has to be seen as ultima ratio option which only applies if other specific regulations do not conclusively regulate the processing in this respect.

The provision first requires that the public authority is acting in fulfilment of a task within its responsibility or in the exercise of its official authority that has been delegated to it. A further – statutory – provision that assigns the task in question to the public authority is therefore required. Alongside this, the data processing by means of the AI must be necessary for the fulfilment of this task or for the exercise of official authority upon it in a restrictive sense.[65] If the data processing is merely helpful, e.g., because it makes processing easier or cheaper, this does not particularly indicate necessity.[66] Also, the development or training of an AI will not be necessary for the fulfilment of the official task pursued with the AI, on a regular basis. Similarly, only data processing with a low level of interference can be based on this legal basis, since the regulation would otherwise be in conflict with the principles of certainty and proportionality.[67] However, when using an artificial intelligence application, the fact that the processing cannot be fully explained may already indicate a serious interference that excludes relying on this general provision and requires a specific statutory regulation. It will hence be decisive which individual phases of processing are carried out for which purpose and by which means.

Particular care must be taken when examining whether third-party providers of an AI are to be integrated. If the data generated during use is to be used by the AI for further training, the transfer of this data to the provider would generally be a transfer that cannot be based on the legal basis under Section 4 LDSG BW. However, if the AI only processes the data for the respective purpose of its input, processing within the framework of a controller processor relationship could be considered under the conditions described if the public authority itself is authorised to process the data using the AI in question.

> **Brief:**
> If public authorities base processing on Section 4 LDSG BW, a detailed legal justification is required in each case.

KEY QUESTIONS:

- Is there a legal basis other than Section 4 LDSG BW for a public authority?
- Does the processing of personal data serve a task assigned to the public authority by law?
- Is the processing necessary for the fulfilment of the task? Are there less intrusive processing methods? Or is the processing only simpler or more cost-effective than the alternatives?

- Is the processing via the AI associated with a high level of interference that can no longer be based on this general provision?

VIII. Processing of special categories of personal data in accordance with Art. 9 para. 1 GDPR

The General Data Protection Regulation provides for increased protection requirements if special categories of personal data within the meaning of Art. 9 para. 1 GDPR are the subject of processing. Once personal data within the meaning of Art. 4 no. 1 GDPR is processed for the training and use of AI, the possibility must also be considered and examined that sensitive information with a high risk in the form of special categories of personal data in accordance with Art. 9 para. 1 GDPR can be derived from this personal data with a low risk to the rights and freedoms of natural persons in the course of the data's life cycle. And as soon as sensitive findings can be derived, this is accompanied by an increased need for protection and confidentiality, since the processing can pose considerable risks to the protection of the rights and freedoms of natural persons.[68]

Therefore, the potential data types – necessary for the extensive data processing of machine learning – must as well be evaluated to determine whether information relating to special categories of personal data in accordance with Art. 9 para. 1 GDPR can be derived from the personal data in the course of the life cycle.

In view of the increased level of confidentiality and protection for the processing of special categories of personal data, a differentiated and careful evaluation of the accuracy and quality of the data within the meaning of Article 5 para. 1 letter (d) GDPR is required. Due to the extensive processing of special categories of personal data necessary for the training of AI, the presumed high risks to rights and freedoms associated with this training data must be identified and minimised in a timely manner. For the evaluation and assurance of the quality of the training data, "state-of-the-art security and privacy-preserving measures, such as pseudonymisation, or encryption" are provided for in terms the use of special categories of personal data in accordance with Art. 10 para. 5 of the draft AI Regulation.

Exceptions to the prohibition of processing special categories of personal data arise from the legal bases under Art. 9 paras. 2–4 GDPR, which must be narrowly interpreted. Depending on the specific purpose of the processing of the special categories of personal data for the preparation of the training of the AI
(e.g., collection and categorisation of data) and the use of the AI, the legal basis in accordance with Art. 6 para. 1 in conjunction with Art. 9 para. 2 GDPR may vary.

Depending on the context of the processing, the legal basis of explicit consent in accordance with Art. 6 para. 1 point (a) in conjunction with Art. 9 para. 2 point (a) GDPR may represent a suitable legal basis for the collection and structuring of training data and the use of the AI. Freely given consent may be questioned in individual cases due to the influence of lock-in effects, nudging, and cognitive distortions (deceptive design patterns).[69] [70] At the same time, it represents an increase in influence compared to the statutory legal basis.

Consent under data protection law as the legal basis for the use of personal data for scientific research in the healthcare sector also directly expresses the right to informational self-determination.[71] Once the processing of personal data for the training and use of the AI for research purposes is legitimised with a legal basis, the data subjects' options for exerting influence, particularly the right to object (e.g., via dashboard systems or other management systems), must be portrayed.[72] In particular, it is important to implement the safeguards to protect rights and freedoms and to ensure a high level of protection and confidentiality in processing in connection with AI. In this respect, the principle of the 'Petersberg Declaration' can be applied to the processing of personal data in preparation for the training of AI, provided that the high level of protection provided for by law in accordance with Art. 32 para. 1 and Art. 89 para. 1 GDPR is maintained: "The higher the level of protection of data subjects through appropriate safeguards and measures, the more extensive and specific the data can be used."

> **Brief:**
> The processing of large data sets of personal data in connection with AI involves the risk that these will become special categories of personal data in the course of the processing lifecycle. Therefore, controllers should consider the requirements for processing special categories of personal data from the outset.

KEY QUESTIONS:

- Can processed personal data lead to information and findings that then allow special categories of personal data to become the subject of processing?
- Does the life cycle of processed personal data mean that a legal basis in accordance with Art. 6 para. 1 in conjunction with Art. 9 para. 2 GDPR must also be included?

1. Data processing for scientific or historical research purposes and for statistical purposes, Section 27 BDSG

According to Section 27 BDSG, particularly protected special categories of personal data within the meaning of Art. 9 para. 1 GDPR can be used for scientific research in the context of AI under certain circumstances. It is controversial whether Section 27 BDSG is a separate legal basis or requires an additional legal basis.[73] In most constellations, the application of the law will depend on balancing and proving that the interests of the controller override the interests of the data subject.

The provision of Section 27 para. 1 sentence 1 BDSG is based on the opening clause of Art. 9 para. 2 point (j) GDPR and requires a balancing of interests with a proven, significantly overriding interest of the controller for the processing in purpose of scientific research. The processing of personal data for structuring and the use of the data for machine learning, and ultimately the use of the results of the machine learning would each need to demonstrably serve scientific research purposes and be necessary for these.

2. Data processing for scientific or historical research purposes and for statistical purposes, Section 13 LDSG BW

Section 13 LDSG BW applies to the privileged processing of personal data by public authorities in Baden-Württemberg in connection with the use of artificial intelligence for scientific and historical research purposes and for statistical purposes. This provision supersedes the general provisions.[74]

The purpose of scientific research is to be understood as the intention of determining the truth in the sense of a serious, planned attempt or the intention of methodically generating new knowledge.[75] A weighing of the fundamental right to freedom of research in comparison to the right to informational self-determination presupposes that the processing of personal data is necessary for the research project. This must be thoroughly evaluated in advance of the research project.

Higher requirements apply to the balancing of interests in the sense of a more intensive justification if the controller processes special categories of personal data in accordance with Art. 9 para. 1 GDPR.[76] The existing legal basis covers preparatory measures necessary for the processing operationst.[77] The legal basis can therefore include the collection and structuring of training data, the process of machine learning with this data, and the use of AI insofar these processing operations are covered by the originally defined purposes and are necessary to achieve them.

In addition to the legal basis to be evaluated, it is necessary that the technical and organisational measures are implemented, particularly to implement the principles of data minimisation, anonymisation, encryption, and pseudonymisation of the data, see Section 13 para. 2 LDSG BW. This applies to the processing phases in connection with AI, whereby the preliminary question arises as to whether Section 13 LDSG BW is applicable as a legal basis, since special regulations, e.g., from the State Hospitals Act, may apply. However, consent is required for the publication of research results, unless the presentation of research results on events in contemporary history is essential, see Section 13 para. 3 LDSG BW.

KEY QUESTIONS:

- Is the purpose pursued by scientific research in the public interest?
- Is the processing of personal data absolutely necessary for the research project described?
- What are the conflicting interests of the data subjects and the controller?

IX. Conclusion

The trust of the citizens in the innovative capacity, security, and responsible use of AI requires that data protection regulations be implemented. There are numerous options for the controller to bring their data processing in connection with artificial intelligence in compliance with data protection law. In some cases, the controller has the choice of which legal basis is preferable for its processing steps. Once they have chosen a legal basis, it is essential that they also fully comply with its requirements. Among other things, this concerns the necessity test that often must be carried out. This paper includes a clear and concise checklist below to facilitate the checking of the steps involved in determining a legal basis for data controllers.

X. Concise checklist for processing

1. Which phase of data processing in connection with AI is subject to legal evaluation? (See part III.)

2. Is personal data processed within the scope of the GDPR (see Art. 2, 3 GDPR)? Or is anonymous data processed that could become personal data?

   - There is a broad definition of personal data, see Art. 4 no. 1 GDPR, i.e., all information relating to an identified or identifiable person. Only permanent anonymisation removes the personal identifiability.

   - The term “processing” is defined in Art. 4 no. 2 GDPR. Please note that each individual phase of processing requires a legal basis (collection, storage, modification, etc.). In particular, anonymisation also requires a legal basis.

   The GDPR, BDSG, and LDSG BW do not apply if no personal data is processed on an ongoing basis.

3. Who is the controller of the data processing? (See also part IV.) The controller according to Art. 4 no. 8 GDPR is the natural or legal person who decides on the means and purposes of processing.

4. Is there a legal basis in data protection law? Are special categories of personal data according to Art. 9 para. 1 GDPR processed that require a legal basis according to Art. 9 para. 2 GDPR? (For the latter, see part VIII.)

5. Legal consequence: In principle, processing is possible. However, the other obligations must be complied with, e.g., the principles of the GDPR (Art. 5 GDPR), compliance with the rights of data subjects (Art. 12 et seq. GDPR), the implementation of technical and organisational measures and guarantees (Art. 24 et seq. and Art. 89 para. 1 GDPR) and, if necessary, the performance of a data protection impact assessment (Art. 35 GDPR).

XI. Collection of materials

Judgements

Judgement of the European Court of Justice of 19 October 2018, Breyer v Bundesrepublik Deutschland, C-582/14, ECLI:EU:C:2016:779 (Breyer decision on the personal nature of data). Available online at: https://curia.europa.eu/juris/document/document.jsf?text=&docid=184668&pageIndex=0&doclang=DE&mode=lst&dir=&occ=first&part=1&cid=2858251 (last accessed on 03.11.2023).

Judgment of the European General Court of 26 April 2023, SRB v. EDPS, T-557/20, (not final, on Regulation (EU) 2018/1725), para. 68.of 26 April 2023, T557/20-, ECLI:EU:T:2023:219.Available online at: https://curia.europa.eu/juris/document/document.jsf?text=&docid=272910&pageIndex=0&doclang=DE&mode=lst&dir=&occ=first&part=1&cid=2858552 (last accessed on 03.11.2023).

Judgement of the European Court of Justice of 21 June 2022, C-817/19, ECLI:EU:C:2022:65 (Artificial intelligence for passenger data), available online at: https://curia.europa.eu/juris/document/document.jsf?text=&docid=252841&pageIndex=0&doclang=DE&mode=lst&dir=&occ=first&part=1&cid=2859153 (last accessed on 06.11.2023).

Papers of the European Data Protection Board (EDPB)

Article 29 Data Protection Working Party, Opinion 06/2014 on the notion of legitimate interests of the data controller under Article 7 of Directive 95/46/EC, WP 217, available online at: https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2014/wp217_en.pdf (last accessed on 03.11.2023).

EDPB, Guidelines 2/2019 on the processing of personal data under Article 6(1)(b) GDPR in the context of the provision of online services to data subjects. Available online at: https://edpb.europa.eu/sites/default/files/files/file1/edpb_guidelines-art_6-1-b-adopted_after_public_consultation_de_0.pdf (last accessed on 06.11.2023).

EDPB, Guidelines 3/2019 on processing of personal data through video devices, available online at: https://edpb.europa.eu/sites/default/files/files/file1/edpb_guidelines_201903_video_devices_de.pdf (last accessed on 03.11.2023).

EDPB, Guidelines 07/2020 on the concepts of controller and processor in the GDPR, available online at: https://edpb.europa.eu/system/files/2022-02/eppb_guidelines_202007_controllerprocessor_final_de.pdfb (last accessed on 03.11.2023).

Papers of the German conferende of independent federal and laender data protection authorities (German Data Protection Conference)

Kurzpapier Nr. 20, Einwilligung nach der DS-GVO [Short Paper No. 20, Consent under the GDPR], 22 February 2019, available online at: https://www.datenschutzkonferenz-online.de/media/kp/dsk_kpnr_20.pdf (last accessed on 06.11.2023).

Hambacher Erklärung zur Künstlichen Intelligenz [Hambach Declaration on Artificial Intelligence], 3 April 2019, available online at: https://www.datenschutzkonferenz-online.de/media/en/20190405_hambacher_erklaerung.pdf (last accessed on 03.11.2023).

Positionspapier der DSK zu empfohlenen technischen und organisatorischen Maßnahmen bei der Entwicklung und dem Betrieb von KI-Systeme [Position paper on recommended technical and organisational measures for the development and operation of AI], 06 November 2019. Available online at: https://www.datenschutzkonferenz-online.de/media/en/20191106_positionspapier_kuenstliche_intelligenz.pdf (last accessed on 03.11.2023).

Petersberger Erklärung zur datenschutzkonformen Verarbeitung von Gesundheitsdaten in der wissenschaftlichen Forschung [Petersberg Declaration on the data protection-compliant processing of health data in scientific research], 24 November 2022, p. 5. Available online at: https://www.datenschutzkonferenz-online.de/media/en/20221124_en_06_Entschliessung_Petersberger_Erklaerung.pdf (last accessed on 06.11.2023).

Essays

Joos, Daniel/Meding, Kristof: Anforderungen bei der Einführung und Entwicklung von KI zur Gewährleistung von Fairness und Diskriminierungsfreiheit [Requirements for the introduction and development of AI to ensure fairness and non-discrimination], DUD 46/2022, 376-380.

Keber, Tobias/Maslewski, Daniel: Rechtsgrundlagen für das Training von Systemen Künstlicher Intelligenz nach der DS-GVO [Legal bases for the training of artificial intelligence systems under the GDPR], RDV 5/2023, pp. 273-280.

Merkert, Pina: Aufmerksamkeit reicht. So funktionieren Sprach-KIs vom Typ „Transformer“ [Attention is enough. How the “Transformer” type language AI function], c’t 11.2022, pp. 136-142.

Wacke, Jan/Nägele, Peter: AI and data protection, BvD-NEWS 02.2023.

Miscellaneous

Arbeitsgruppe „Einsatz von KI und algorithmischen Systemen in der Justiz“, Grundlagenpapier zur 74. Jahrestagung der Präsidentinnen und Präsidenten der Oberlandesgerichte, des Kammergerichts, des Bayerischen Obersten Landesgerichts und des Bundesgerichtshofs vom 23. bis 25. Mai 2022 in Rostock [Working group “Use of AI and algorithmic systems in the judiciary”, basic paper for the 74th Annual Conference of the Presidents of the Higher Regional Courts, the Court of Appeal, the Bavarian Supreme Court and the Federal Court of Justice from 23 to 25 May 2022 in Rostock]. Available online at: https://oberlandesgericht-celle.niedersachsen.de/download/184478/Grundlagenpapier_der_Arbeitsgruppe_zum_Einsatz_von_KI_und_algorithmischen_Systemen_in_der_Justiz.pdf (last accessed on 06.11.2023).

Bundesministerium für Arbeit und Soziales (BMAS)/Bundesministerium des Innern und für Heimat (BMI), Eckpunktepapier: Vorschläge für einen modernen Beschäftigtendatenschutz Innovation ermöglichen – Persönlichkeitsrechte schützen – Rechtsklarheit schaffen [Federal Ministry of Labour and Social Affairs (BMAS)/Federal Ministry of the Interior and Community (BMI), Key issues paper: Proposals for modern employee data protection Enabling innovation – protecting personal rights – creating legal clarity], May 2023. Available online at: https://fragdenstaat.de/anfrage/aktueller-stand-beschaeftigtendatenschutz/804753/anhang/vorschlge-beschftigtendatenschutz.pdf (last accessed on 06.11.2023).

Bundesamt für Sicherheit in der Informationstechnik (BSI), Große KI-Sprachmodelle – Chancen und Risiken für Industrie und Behörden [Federal Office for Information Security (BSI), Large AI language models – opportunities and risks for industry and authorities], 2021. Available online at: https://www.bsi.bund.de/SharedDocs/Downloads/DE/BSI/KI/Grosse_KI_Sprachmodelle.pdf?__blob=publicationFile&v=2 (last accessed on 03.11.2023).

Commission Nationale de l'Informatique et des Libertés (CNIL), Extensive materials from the French supervisory authority on the subject of AI. Available online at: https://www.cnil.fr/en/topics/artificial-intelligence-ai" (last accessed on 06.11.2023).

Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legistlative acts (Version 21 April 2021). Available online at: https://eur-lex.europa.eu/resource.html?uri=cellar:e0649735-a372-11eb-9585-01aa75ed71a1.0019.02/DOC_1&format=PDF (last accessed on 03.11.2023).

Gesetz zur elektronischen Verwaltung für Schleswig-Holstein (E-Government-Gesetz – EGovG) vom 08. Juli 2009 [Act on Electronic Administration for Schleswig-Holstein (E-Government Act – EGovG) of 8 July 2009]. Available online at: https://www.gesetze-rechtsprechung.sh.juris.de/bssh/document/jlr-EGovGSH2009rahmen/part/X" (last accessed on 06.11.2023).

Information Commissioner's Office (ICO), Extensive materials from the UK supervisory authority. Available online at: https://ico.org.uk/for-organisations/uk-gdpr-guidance-and-resources/artificial-intelligence/" (last accessed on 06.11.2023).

LfDI BW, Letter with list of questions regarding OpenAI on ChatGPT of 21 April 2023, available online at: https://fragdenstaat.de/anfrage/schreiben-an-openai/811596/anhang/openai-chatgpt.pdf" (last accessed on 06.11.2023).

Stiftung Datenschutz, Praxisleitfaden für die Anonymisierung personenbezogener Daten [Practical guide for the anonymization of personal data], 2022. Available online at: https://stiftungdatenschutz.org/fileadmin/Redaktion/Dokumente/Anonymisierung_personenbezogener_Daten/SDS_Studie_Praxisleitfaden-Anonymisieren-Web_01.pdf (last accessed on 03.11.2023).

Footnotes

[1] Data protection law does not define the term artificial intelligence. In this paper, we understand the term – which is controversial in science and practice – to mean all machine learning systems in the sense of a broad working definition. See also Art. 3 para. 1 no. 1 AI Regulation-E.

[2] Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legistlative acts (Version 21 April 2021). Available online at: https://eur-lex.europa.eu/resource.html?uri=cellar:e0649735-a372-11eb-9585-01aa75ed71a1.0019.02/DOC_1&format=PDF ↗ (last accessed on 03.11.2023).

[3] The term *AI* encompasses the following: 1.) the preparation of an AI, 2.) the model set up, 3.) the training of the AI, 4.) the application and 5.) the use of the AI.

[4] See Paal, in: Kaulartz/Braegelmann, Rechtshandbuch Artificial Intelligence und Machine Learning [Legal Guide Artificial Intelligence and Machine Learning], 2020, p. 427 para. 1 f.

[5] See Ernst, in: Paal/Pauly, DS-GVO Kommentar [GDPR Commentary], 3rd ed. 2021, Art. 4 para. 3; Judgment of the European General Court of 26 April 2023, SRB v. EDPS, T-557/20, (not final, on Regulation (EU) 2018/1725), para. 68.

[6] See Kaulartz, in: Kaulartz/Braegelmann, Rechtshandbuch Artificial Intelligence und Machine Learning [Legal Guide Artificial Intelligence and Machine Learning], 2020, p. 463 para. 4.

[7] See Judgement of the European Court of Justice of 19 October 2018, Breyer, C-582/14, ECLI:EU:C:2016:779.

[8] See ibid., para. 43.

[9] An AI model is a component of the AI, which is made up of parameters, learned data, and an architecture.

[10] See Kaulartz, in: Kaulartz/Braegelmann, Rechtshandbuch Artificial Intelligence und Machine Learning [Legal Guide Artificial Intelligence and Machine Learning], 2020, pp. 468 et seq.

[11] See Rigaki/Garcia, A Survey of Privacy Attacks in Machine Learning, 2020, p. 1.

[12] See Veale/Binns/Edwards, Phil. Trans. R. Soc. A 2018, 376, 376 et seq.

[13] See Kaulartz, in: Kaulartz/Braegelmann, Rechtshandbuch Artificial Intelligence und Machine Learning [Legal Guide Artificial Intelligence and Machine Learning], 2020, p. 467 para. 13; on the basic problem, see also: Stiftung Datenschutz, Praxisleitfaden für die Anonymisierung personenbezogener Daten [Practical guide for the anonymization of personal data], 2022, p. 28, point 6.2. Available online at: https://stiftungdatenschutz.org/fileadmin/Redaktion/Dokumente/Anonymisierung_personenbezogener_Daten/SDS_Studie_Praxisleitfaden-Anonymisieren-Web_01.pdf (last accessed on 03.11.2023); Bundesamt für Sicherheit in der Informationstechnik, Große KI-Sprachmodelle – Chancen und Risiken für Industrie und Behörden [Federal Office for Information Security, Large AI language models – opportunities and risks for industry and authorities], 2021, p. 16, point 3.4.2. Available online at: https://www.bsi.bund.de/SharedDocs/Downloads/DE/BSI/KI/Grosse_KI_Sprachmodelle.pdf?__blob=publicationFile&v=2 (last accessed on 03.11.2023).

[14] See also: Blanco-Justicia, Alberto et al, A critical review on the use (and misuse) of differential privacy in machine learning, ACM Computing Surveys 2022, 1, pp. 1 et seq.

[15] Bourtoule, Lucas et al, Machine Unlearning, in: 2021 IEEE Symposium on Security and Privacy (SP), 2021.

[16] For the problem of personal identifiability, see part II.

[17] As a rule, the users cannot invoke the household exception Art. 2 para. 2 point (c) GDPR, since the processing is not exclusively in a private context.

[18] See EDPB, Guidelines 07/2020 on the terms “controller” and “processor” GDPR, 07.07.2021, p. 3.

[19] See ibid.

[20] See LfDI BW, Mustervereinbarung nach Art. 26 DS-GVO [Model agreement in accordance with Art. 26 GDPR]. Available online at: https://www.baden-wuerttemberg.datenschutz.de/praxishilfen/#gemeinsame_verantwortlichkeit (last accessed on 03.11.2023).

[21] Irrespective of the agreement, data subjects may assert their rights in accordance with Art. 26 para. 3 GDPR with and against each of the joint controllers.

[22] See EDPB, Guidelines 07/2020 on the terms “controller” and “processor” GDPR, p. 4. Building on this, see: LfDI BW, FAQ zur Abgrenzung der Verantwortlichkeiten und des Begriffs der Auftragsverarbeitung [FAQ on the delimitation of responsibilities and the concept of a controller processor relationship]. Available online at: https://www.baden-wuerttemberg.datenschutz.de/faq-zur-abgrenzung-der-verantwortlichkeiten-und-des-begriffs-der-auftragsverarbeitung/ (last accessed on 03.11.2023).

[23] See LfDI BW, Auftragsverarbeitungsvertrag nach Art. 28 Abs. 3 DS-GVO [Data processing agreement in accordance with Art. 28 para. 3 GDPR]. Available online at:

https://www.baden-wuerttemberg.datenschutz.de/wp-content/uploads/2020/12/200429_AVV-Muster_DE_neu.pdf (last accessed on 03.11.2023).

[24] EDPB, Guidelines 05/2020 on consent in accordance with Regulation 2016/679, p. 7. Available online at: https://edpb.europa.eu/our-work-tools/our-documents/guidelines/guidelines-052020-consent-under-regulation-2016679_de (last accessed on 03.11.2023).

[25] See Arning/Rothkegel, in: Taeger/Gabel, DS-GVO – BDSG – TTDSG [GDPR – Federal Data Protection Act – Federal Telecommunications-Telemedia Data Protection Act], 4th ed. 2022, Art. 4 para. 329.

[26] See ibid., Art. 4 para. 330.

[27] See Albers/Veit, in: Wolff/Brink/v. Ungern-Sternberg, BeckOK Datenschutzrecht [Data protection law], 45th ed. 2023, Art. 6 para. 36.

[28] See Keber/Maslewski, RDV 2023, 273, 277.

[29] See Niemann/Kevekordes, CR 2020, 17, 23.

[30] See Kloos/Schmidt-Bens, in: Hartmann, KI & Recht kompakt [AI & law], 2020, p. 174.

[31] See Datenethikkommission, Gutachten [Data Ethics Commission, Opinion], 23 October 2019, p. 169. Available online at: https://www.bmi.bund.de/SharedDocs/downloads/DE/publikationen/themen/it-digitalpolitik/gutachten-datenethikkommission.pdf?__blob=publicationFile&v=6 (last accessed on 03.11.2023).

[32] See Skistims, in: Kaulartz/Braegelmann, Rechtshandbuch Artificial Intelligence und Machine Learning [Legal Guide Artificial Intelligence and Machine Learning], 2020, p. 358 para. 19 f.

[33] See Judgement of the European Court of Justice of 1 October 2019, Planet 49 GmbH, C-673/17, ECLI:EU:C:2019:801, ZD 2019, 556, 560.

[34] See Recital 42 sentence 4 GDPR.

[35] See the detailed discussion of Art. 6 para. 1 point (f) GDPR (part VI.1.) on the concept of necessity

[36] See Judgement of the European Court of Justice of 4 July 2023, Meta Platforms, C-252/211, ECLI:EU:C:2023:537, para. 98: “it [processing personal data] must be objectively indispensable for a purpose that is integral to the contractual obligation intended for the data subject. The controller must therefore be able to demonstrate how the main subject matter of the contract cannot be achieved if the processing in question does not occur.”

[37] See Judgement of the European Court of Justice of 4 July 2023, Meta Platforms, C-252/211, ECLI:EU:C:2023:537, para. 99.

[38] This is the case if it is "actually necessary for compliance with a legal obligation to which the controller is subject, pursuant to a provision of EU law or the law of the Member State concerned, where that legal basis meets an objective of public interest and is proportionate to the legitimate aim pursued and where that processing is carried out only in so far as is strictly necessary." Judgement of the European Court of Justice of 4 July 2023, Meta Platforms, C-252/211, ECLI:EU:C:2023:537, para. 138.

[39] See Kaulartz, in: Kaulartz/Braegelmann, Rechtshandbuch Artificial Intelligence und Machine Learning [Legal Guide Artificial Intelligence and Machine Learning], 2020, p. 472 para. 30.

[40] On the controversial question of whether a further (separate) legal basis in accordance with Art. 6 para. 1 GDPR is required for the lawfulness of data processing in the context of further processing, see: Albers/Veit, in: Wolff/Brink/v. Ungern-Sternberg, BeckOK Datenschutzrecht [Data protection law], 45th Ed., Art. 6 para. 107 f.

[41] See in this regard: Keber/Maslewski, RDV 2023, 273.

[42] The questions of the extent to which the exclusion of Art. 6 para. 1 point (f) GDPR through Art. 6 para. 1 point (2) GDPR applies to competitive public companies and the extent to which processing by public authorities outside the fulfilment of tasks can be based on Art. 6 para. 1 point (f) GDPR will not be addressed here.

[43] See Schantz, in: Simitis/Hornung/Spiecker gen. Döhmann, Datenschutzrecht [Data protection law], 2019, GDPR Art. 6 para. 1 para. 86.

[44] See Judgement of the Court of Justice of 4 July 2023, Meta Platforms, C-252/211, ECLI:EU:C:2023:537, para. 106 with further references.

[45] See Buchner/Petri, in: Kühling/Buchner, DS-GVO/BDSG [GDPR/BDSG, 3rd ed. 2020, Art. 6 para. 146.

[46] See EDPB, Guidelines 3/2019 on the processing of personal data through video devices, para. 18.

[47] See Recital 47 sentence 6 GDPR.

[48] Judgement of the European Court of Justice of 4 July 2023, Meta Platforms, C-252/211, ECLI:EU:C:2023:537, para. 108.

[49] See ibid., para. 109.

[50] See ibid., para. 106.

[51] See ibid., para. 110.

[52] See ibid., para. 116.

[53] See EDPB, Guidelines 3/2019 on processing of personal data through video devices, para. 33; see also Article 29 Working Party, Opinion 06/2014 on the notion of legitimate interest of the data controller under Article 7 of Directive 95/46/EC, WP 217, pp. 43 et seq.

[54] See Recital 38 GDPR.

[55] See Judgement of the European Court of Justice of 4 July 2023, Meta Platforms, C-252/211, ECLI:EU:C:2023:537, para. 112.

[56] See Kaulartz, in Kaulartz/Braegelmann, Rechtshandbuch Artificial Intelligence und Machine Learning [Legal Guide Artificial Intelligence and Machine Learning], 2020, p. 473 para. 35.

[57] The extent to which the controller has implemented “data protection by design” and “data protection by default” may also have an impact on the assessment.

[58] See Judgement of the European Court of Justice of 30 March 2023, C-34/21, ECLI:EU:C:2023:270. See also: https://www.baden-wuerttemberg.datenschutz.de/faq-rechtsgrundlagen-bei-beschaeftigtendaten/ (last accessed on 03.11.2023).

[59] For more information on consent as a legal basis, see part V.1.

[60] See also detailed information on the problem of AI in the employment context: Joos, NZA 2020, 1216, 1216 et seq.

[61] Directive (EU) 2016/680 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data by competent authorities for the purposes of the prevention, investigation, detection, or prosecution of criminal offences or the execution of criminal penalties, and on the free movement of such data, and repealing Council Framework Decision 2008/977/JHA. Available online at: https://eur-lex.europa.eu/legal-content/DE/TXT/?uri=uriserv%3AOJ.L_.2016.119.01.0089.01.DEU ↗ (last accessed on 03.11.2023).

[62] See also the state government’s draft law on the law to adapt the general data protection law and other provisions to Regulation (EU) 2016/679, state parliament printed paper 16/3930, p. 93; for more information, see: LfDI BW, Datenschutz bei Gemeinden [Data protection in municipalities]. Available online at: 🗎 https://www.baden-wuerttemberg.datenschutz.de/wp-content/uploads/2019/11/Brosch%C3%BCre-Gemeinden-November-2019.pdf (last accessed on 03.11.2023).

[63] See also Holz/Stich, in: Brinktrine/Hug, BeckOK Beamtenrecht [Civil service law] Baden-Württemberg, 2020, Section 84 LDSG, para. 7 et seq.

[64] Debus, in: Debus/Sicko, Landesdatenschutzgesetz Baden-Württemberg [Baden-Württemberg State Data Protection Act], 2020, Section 18 para. 24 et seq.

[65] For basic information on the concept of necessity, see the above comments on Art. 6 para. 1 point (f) GDPR, part VI.1.

[66] See also Osterried, in: Debus/Sicko, LDSG [Baden-Württemberg State Data Protection Act] , Section 4 para. 36.

[67] See also ibid, LDSG [Baden-Württemberg State Data Protection Act], Section 4 para. 12 et seq.

[68] See Recital 51 sentence 1 GDPR.

[69] Lock-in effect: technical-functional customer loyalty, e.g., through the fact that a service can only be used with the same company's device or through a monopoly position; nudging: unconscious influencing of behaviour for the user; deceptive design patterns: interface design that uses targeted psychological effects to unconsciously entice users to behave in a desired way.

[70] See Schantz, in: Simitis/Hornung/Spiecker gen. Döhmann, DS-GVO [GDPR], 2019, Art. 6 para. 4.

[71] See regarding this: German Data Protection Conference (DSK), Petersberger Erklärung zur datenschutzkonformen Verarbeitung von Gesundheitsdaten in der wissenschaftlichen Forschung [Petersberg Declaration on the data protection-compliant processing of health data in scientific research], 24 November 2022, p. 5. Available online at: https://www.datenschutzkonferenz-online.de/media/en/20221124_en_06_Entschliessung_Petersberger_Erklaerung.pdf (last accessed on 06.11.2023).

[72] German Data Protection Conference (DSK), Petersberger Erklärung zur datenschutzkonformen Verarbeitung von Gesundheitsdaten in der wissenschaftlichen Forschung [Petersberg Declaration on the data protection-compliant processing of health data in scientific research], 24 November 2022, p. 7. Available online at: https://www.datenschutzkonferenz-online.de/media/en/20221124_en_06_Entschliessung_Petersberger_Erklaerung.pdf (last accessed on 06.11.2023).

[73] Koch, in: Wolff/Brink/v. Ungern-Sternberg, BeckOK Datenschutzrecht [Data protection law], 45th ed., BDSG, Section 27 para. 5.

[74] Landtagsdrucksache 16/3930 [Baden-Württemberg's State parliament printed paper 16/3930], p. 100.

[75] See Keber, in: Debus/Sicko, LDSG [Baden-Württemberg State Data Protection Act], 2022, Section 13 para. 13.

[76] See ibid., para. 23.

[77] See ibid., para. 14.