Orientation guide

of the Conference of Independent Federal and State Data Protection Supervisory Authorities of May 6, 2024

Artificial intelligence and data protection

version 1.0

Many companies, authorities and other organizations are currently asking themselves under what conditions they can use AI applications in compliance with data protection regulations. From 2023, the focus will be on so-called Large Language Models (LLM), which are often offered as chatbots, but can also serve as the basis for other applications. The focus of the following guidance is therefore currently on these AI applications. Beyond LLMs, however, there are numerous other AI models and AI applications that can be considered for use and for which many of the following considerations are also likely to be relevant.

This guide provides an overview of data protection criteria that must be taken into account for the data protection-compliant use of AI applications. It can serve as a guide for selecting, implementing and using AI applications.

The guidance is likely to be adapted in the future to include current developments and other relevant aspects. It offers a guideline, but does not represent an exhaustive list of requirements. In some cases, additional resources will need to be consulted in order to implement the points addressed in this guide.

The guidance is primarily aimed at those responsible who wish to use AI applications. It is indirectly aimed at developers, manufacturers and providers of AI systems, as it contains information on selecting AI applications that comply with data protection regulations. However, the development of AI applications and the training of AI models are not the focus of this guidance.

Table of contents

1. Conception of the use and selection of AI applications 3
1.1 fields of application and purposes? 3
1.2 Fields of application lawful? 3
1.3 Fields of application without personal data? 3
1.4 Data protection-compliant training of AI applications 4
1.5 Legal basis for data processing? 4
1.6 No automated final decision 5
1.7 Closed or open system? 5
1.8 Transparency 6
1.9 Transparency and choice regarding AI training 7
1.10 Transparency and choice regarding input history 7
1.11 Rectification, erasure and other data subject rights 7
1.12 Involving data protection officers and employee representatives 8
2. Implementation of AI applications 8
2.1 Defining and bindingly regulating responsibilities 8
2.2 Making internal regulations 9
2.3 Data protection impact assessment 10
2.4 Protect employees, set up company accounts 10
2.5 Data protection through technology design and data protection-friendly default settings 11
2.6 Data security 11
2.7 Sensitize employees 12
2.8 Follow further developments 12
3. Use of AI applications 12
3.1 Caution when entering and outputting personal data 12
3.2 Special care with special categories of personal data 14
3.3 Check results for correctness 14
3.4 Checking results and procedures for discrimination 15

1. Conception of the use and selection of AI applications

1.1 Fields of application and purposes determined?

Before using an AI application, data controllers should explicitly define1 which fields of application are intended for the AI application and what specific purpose it serves. With regard to the processing of personal data, this purpose definition is fundamental for data protection-compliant operation, as it is only possible to check whether the processing of personal data is necessary to achieve the purpose on the basis of specific pre-defined purposes.

In this regard, it is also important for public bodies to ensure that

t

he2 field of application is within the scope of the public tasks assigned to them by law and that the processing of personal data within this scope is necessary for the fulfillment of the task.

1.2 Fields of application lawful?

Certain fields of use for AI applications may be prohibited from the outset. 3
For example, according to the European AI Regulation, "social scoring" and biometric real-time monitoring of public spaces are considered artificial intelligence practices that are either completely prohibited or only permitted u n d e r very narrow exceptions.

1.3 Fields of application without personal data?

When defining fields of application, it may turn out that there are definable 4
There are areas of use in which no personal data is used, neither as input data or output data of an AI application, nor in the registration and processing process of the AI application. Such areas of application are not subject to data protection law. However, it should be noted that a personal reference can arise from many characteristics, not just names and address data. The check as to whether or not personal data is present in a field of application must therefore be carried out thoroughly and throughout the life cycle of the data.

Example 1: A state geological office would like to use an AI application to re evaluate. Here are only geological maps are used, which have no personal reference and do not refer to areas with residential buildings.

Example 2: A company's development team uses an LLM chatbot to 6
error in a code sequence that does not contain a personal reference. It is

However, it must be checked whether personal data can be processed due to a personal reference in the AI model.

1.4 Data protection-compliant training of AI applications

7 With regard to the selection of AI applications, it may play a role whether and to what extent AI applications have been trained in compliance with data protection regulations.

- Was personal data used for the training?
- If yes, was there a legal basis for using the data for the training?
- To what extent does the AI application itself relate to people at the time of use?

8 As a rule, controllers who use AI applications have no influence over these parameters. However, controllers who use AI under their own responsibility must ensure that errors in the training of an AI application do not affect the data processing under their responsibility.

1.5 Legal basis for data processing?

9 A legal basis under data protection law is required for each processing step in which personal data is processed with the help of an AI application. Various legal bases can be considered, depending on whether it is a public or non-public body and whether it is, for example, an application in the field of human resources, healthcare or processing in the area of a consumer or service contract.

10 Furthermore, the use of AI applications that process citizens' personal data by public authorities may require a specific legal basis that specifically addresses the risks to the rights and freedoms of data subjects arising from the processing, depending on the intensity of the processing.

11 An overview of far-reaching issues relating to the selection of and compliance with data protection legislation in connection with the use of AI applications can be found, for example, in the discussion paper
"Legal bases in data protection in the use of artificial intelligence" of the

State Commissioner for Data Protection and Freedom of Information of Baden-Württemberg.[1]

1.6 No automated final decision

According to Art. 22 (1) GDPR, decisions with legal effect may

o

nly be taken by humans12 . Exceptions are only permitted in certain cases, such as the consent of the data s u b j e c t . If an AI application develops proposals that have legal effect for a data subject, it must
the procedure must be designed in such a way that the person making the decision has real leeway in the decision-making process and that decisions are not based primarily on the AI proposal. Insufficient personnel resources, time pressure and a lack of transparency regarding the decision-making process of the AI-supported preliminary work must not lead to results being accepted without review. The merely formal involvement of a human being in the decision-making process is not sufficient.

Example:

An AI application evaluates all applications received for an advertised position and automatically sends out invitations to interviews. This constitutes a violation of Art. 22 para. 1 GDPR.

The following also applies to public bodies: The fully automated decree [14]
of an administrative act is regulated in Section 35a VwVfG. If the requirements are met, Art. 22 para. 1 GDPR does not apply in accordance with Art. 22 para. 2 lit. b GDPR. A fully automated issuance of an administrative act is only permissible if it is a binding decision and there is an explicit basis for authorization. If the public authority has a margin of appreciation or exercises discretion, fully automated enactment is ruled out.

1.7 Closed or open system?

In AI applications, a distinction can be made between closed and open systems. [15]
be differentiated.

In closed systems, data processing takes place in a restricted

an

d16 technically closed environment. In addition, only a specific, strictly limited group of users has access to the AI application. Control over the
The user is responsible for input and output data in closed systems. The system does not provide for the data entered or output during the application to be

[1] https://www.baden-wuerttemberg.datenschutz.de/rechtsgrundlagen-datenschutz-ki/.

The resulting data can be used by the system provider for further training.

17 The situation is different with open systems. Such AI applications are operated by the provider as a cloud solution, for example, and are accessible to an undefined group of users via the internet. The input data thus leaves the protected area of the user and, depending on the design of the AI application, can also be used by the user to answer queries from other users. In this case, there is a risk that personal data may be further processed for other purposes or may also be accessible to unauthorized third parties and disclosed to them. In this context, reference should also be made to possible transfers of data to third countries, as they are frequently encountered in such constellations. The regulations in Chapter V of the GDPR must be taken into account for transfers to third countries.

18 There is also a risk with regard to official information that is not intended for the public or that is classified.

19 Open systems can also have access to other data sources such as the open Internet and thus establish personal references to data or expand the information on a person.

20 Technically closed systems are therefore preferable from a data protection perspective.

1.8 Transparency

21 The use of AI applications poses particular challenges for controllers with regard to their information and transparency obligations in several respects. If controllers do not develop an AI application themselves, they must ensure that they are provided with sufficient information by the provider in order to be able to implement the transparency requirements of Art. 12 et seq. GDPR can be implemented. To this end, manufacturers must provide AI users with appropriate documentation. If the AI application is used as a cloud solution, for example, the processor is obliged to support the controller in complying with the rights of the data subject in accordance with Art. 28 para. 3 sentence 2 lit. e GDPR.

22 The information about which the data controllers must inform and provide information also includes information about the data subjects involved in automated decision-making, including profiling in accordance with Art. 22 (1) GDPR.

logic and the scope and possible effects for the data subject. The term "automated decision" will cover many AI applications that make automated decisions themselves or whose outcome significantly influences decisions.

From the term "logic" it can at least be concluded that an[23] explanation of the method of data processing must be provided in relation to the functioning of the program sequence in connection with the specific application.
Visualizations and interactive techniques can help to break down the complexity of the logic to an understandable level.

1.9 Transparency and choice regarding AI training

It must be checked whether input and output data is

u

sed for training, [24]
sufficient information is provided in this regard and the possibility i s given to exclude the use of the data for training. If an exclusion of the use
If this is not possible for training purposes and personal data is involved, a legal basis is required for this purpose. Applications that do not use input and output data for training purposes are therefore preferable under data protection law.

1.10 Transparency and choice regarding input history

Many services controlled by text input (prompts) offer to save the input[25] so that, for example, the dialogue on a topic can be resumed at a later time or to work on further optimization of the prompt.
This creates a history of a person's entries. This must be communicated transparently, especially when shared by several employees, and users must be able to decide for themselves whether their own input history is saved.

1.11 Rectification, erasure and other data subject rights

Controllers must ensure that data subjects can exercise their rights to[26] rectification in accordance with Art. 16 GDPR and erasure in accordance with Art. 17 GDPR. Organizational and technical procedures must be designed for both rights so that they can be exercised effectively. To this end, the requirements of data protection-compliant technology design must be implemented.

When using AI applications, incorrect personal data may be processed27 for various reasons. Many providers of AI applications (in particular LLM chatbots) even point out that

expressly points out that users cannot rely on the accuracy of the results, but must check them. With regard to personal data, however, data subjects have a right to rectification in the event of inaccuracy. It must be possible to implement this rectification in an AI application, for example by correcting data or through retraining/fine tuning.

28 If data subjects exercise their right to erasure in accordance with Art. 17 (1) GDPR, it should be noted that some AI applications may be able to establish a personal reference by linking different data. It is therefore particularly important that when deleting personal data, care is taken to ensure that it is permanently impossible to restore the personal reference. Depending on the AI application, this can be implemented in various ways.

29 The suppression of unwanted output by means of downstream filters does not generally constitute erasure within the meaning of Art. 17 GDPR. This is because the data that leads to a certain output after a certain input could still be available for the AI model in a personally identifiable form. However, filter technologies can help to avoid certain outputs and thus serve the rights and freedoms of the persons affected by a particular output.

30 The other data subject rights to restriction of processing and data portability as well as the right to object must also be taken into account when designing the AI application.

1.12 Involve data protection officers and employee representatives

31 Company and official data protection officers should always be involved when decisions on AI applications are prepared or made. The involvement of works and staff councils should also be considered and examined.

2. Implementation of AI applications

2.1 Define responsibilities and make binding arrangements

32 The controller within the meaning of the GDPR is the entity that decides on the purposes and means of processing personal data. If the AI application is operated by one entity exclusively for its own purposes on its own servers, this entity is generally also to be regarded as the sole controller.

If a body uses an AI application from an external provider for its own purposes, for example as a cloud solution, the external provider acts as an extended arm on behalf of the controller. In this case, there is often an order processing relationship between the provider of the application and the controller in accordance with Art. 28 et seq. GDPR with the consequence that an agreement must be concluded with the provider in accordance with Art. 28 para. 3 GDPR.

Joint controllership pursuant to Art. 26 GDPR can b e as sumed[33] if two entities jointly decide on the purposes and means of processing, i.e. make a joint decision on this. However, joint controllership may also exist if the entities involved make complementary decisions and these are necessary for the processing in such a way that they have a significant influence on the determination of the purposes and means of the processing. An important criterion for the assumption of joint controllership in the case of converging decisions is, in particular, whether the processing would not be possible without the involvement of both parties in the purposes and means in the sense that the processing operations of both parties are inextricably linked. This can be considered in the case of cooperation between several entities, for example, if an AI application is fed or trained with different data sets or if its AI application is further developed into new AI applications by other entities on the platform of one entity. It is not necessary for the controller to actually have access to the processed data in order to be classified as a joint controller.

Pursuant to Art. 26 para. 1 sentence 2 GDPR, the bodies involved must s pecify in a tr ansparent manner in an ag reement[34] which of them complies with which obligations of the GDPR, in particular the fulfillment of the rights of data subjects and the information obligations pursuant to Art. 13 and 14 GDPR.

Art. 26 GDPR does not constitute a legal basis for the processing of personal data· data, so that each controller requires a separate legal basis for processing within the scope of joint controllership. In addition, the transfer of personal data between joint controllers is a separate processing operation and as such requires a legal basis.

2.2 Make internal regulations

Without clear regulations on whether and how AI applications can be used in day-to-day work [36]

If employees are not allowed to use AI applications, there is a risk that they will use them in an unauthorized and uncontrolled manner. It can be assumed that this is currently a reality in many companies and authorities. This can lead to data protection violations or even

to other damage for the respective organization. Clear internal instructions should therefore be issued and documented as to whether, under what conditions and for what specific purposes which AI applications may be used. Specific examples of permitted and prohibited use scenarios can be helpful in clarifying this and are therefore recommended.

37 Regardless of whether personal data (including employee usage data) is processed by an AI application, it is advisable to issue a service/operating directive or conclude a service/works agreement between management and the staff/works council. In any case, a clear framework for the use of AI applications should be specified. This applies all the more if personal data is processed. In some cases, the introduction of an AI application will also fulfill an operational co-determination requirement.

2.3 Data protection impact assessment

38 Before processing personal data, a general assessment (prior check) of the risk must be carried out with regard to the type, scope, purpose and circumstances of the processing.

39 If it is determined that the processing is likely to result in a high risk to the rights and freedoms of natural persons, a data protection impact assessment (DPIA) is required in accordance with Art. 35 GDPR. This will often be the case when using AI applications. The data protection supervisory authorities also publish so-called "must lists"[2] for processing operations for which a DPIA must be carried out, as well as information on the cases in which the DPIA can be dispensed with.

40 If the controller is not also the provider of the AI system, it is dependent on information from the provider, in particular on how the system works, in order to carry out a risk assessment or DPIA. Therefore, when selecting and purchasing an AI application, care must be taken to ensure that this information is provided by the provider.

2.4 Protect employees, set up company accounts

41 Employers should provide devices and accounts for the professional use of AI applications by employees. Employees should

[2] "Must list" for the non-public sector: https://www.datenschutzkonferenz-online.de/media/ah/20181017_ah_DSK_DSFA_Muss-Liste_Version_1.1_Deutsch.pdf. The data protection supervisory authorities have published their own "must" lists for the public sector.

do not have to work with AI applications independently and using private accounts and devices, as this can create profiles for the respective employees.

The accounts should not contain the names of individual employees

u

nless the KI-

[42] application is operated on the company's own servers. If the e-mail address is requested, the specification of a functional e-mail address of the company or
of the office. In some cases, mobile phone numbers are also required for registration. The employer must also provide a telephone for this purpose.

2.5 Data protection through technology design and data protection-friendly default settings

According to Art. 25 GDPR, controllers of systems in which personal data is processed must take technical and organizational measures to implement the data protection principles. Data protection-compliant design in the sense of "data protection by design" and "data protection by default", which takes into account the special features of AI systems, can be considered. These requirements must already be taken into account when designing the AI system. This can be
Depending on the AI application, this may affect various aspects. For example, for accounts that employees are to use, the functions for using input for AI training and the input history can be selected when an account is set up so that no input data is processed for training purposes and no input history is saved beyond the session. Output data belonging to the account must also not be published automatically.

2.6 Data security

As information technology systems, AI applications must comply with data protection [44]
In principle, the technical and organizational measures required by law (in particular in accordance with Articles 25 and 32 GDPR) must also meet the requirements that generally apply to IT systems. These include, in particular, criteria such as reliability and usability as well as security (confidentiality, integrity, availability and resilience).

If attackers manage to gain unauthorized access to the AI applications, [45]
they may be able to find out about previous activities, personal information and business secrets.

The German Federal Office for Information Security (BSI), for example, provides extensive information on information security specifically for the use of AI applications.[3]

2.7 Sensitize employees

46 Employees should be sensitized through training, guidelines and discussions as to whether and how they should and may use AI applications.

2.8 Follow further developments

47 In both the legal and technical areas, controllers must monitor current developments that have an impact on the processing of personal data and the necessary control of the risk to the rights and freedoms of natural persons. This includes examining the extent to which additional requirements from the European AI Regulation must be met. In relation to technical developments that may have an impact on the risk, such as advancing technical solutions and updates, it should also be regularly reviewed whether the internal requirements need to be adapted. To this end, a routine should be established as part of operational or official data protection management (Art. 24 GDPR).

3. Use of AI applications

3.1 Caution when entering and outputting personal data

48 If the input data is personal data, data subjects must be informed transparently about the use of their data. A legal basis must also be fulfilled for the processing of personal data in AI applications and the possible transfer of this data to providers of AI applications.

49 To avoid entering personal data, it is generally not sufficient to remove names and addresses from an entry. If necessary, a personal reference to data subjects can be established from the context. This risk is particularly high for AI applications, which are designed to create cross-references from unstructured data.

[3] https://www.bsi.bund.de/DE/Themen/Unternehmen-und-Organisationen/Informationen-und-Empfehlungen/Kuenstliche-Intelligenz/kuenstliche-intelligenz_node.html.

Example of an entry with personal reference:

"Draft a reference in the satisfactory range for a customer advisor [50]
at car dealership X."

The input can be personal if it is recognizable from which [51]
company at which time it was made.

Even if an entry does not contain any personal data, the AI- [52]
application generates an output with personal data. In such cases, a legal basis may be required. It must also be checked whether the data subjects must be informed about the processing in accordance with Art. 14 GDPR. This is necessary at least if the request was aimed at a personal reference or the personal output is to be further processed.

Example 1 of an input that produces an output without personal reference:

Input: [53]
"Name two novels by Hermann Hesse."

Issue: [54]
"Two novels by Hermann Hesse are:
1. *"Siddhartha" (1922): A philosophical novel describing the spiritual quest of a young Brahmin named Siddhartha in ancient India.*
2. *"The Steppenwolf (1927): A psychological novel that explores the inner turmoil and search for identity of a man named Harry Haller in the modern world."*

The edition does not contain any data on living/real persons. [55]

Example 2 of an input that produces an output without personal reference:

Input: [56]
"Write a haiku about data protection."

Issue: [57]
"Data protected finely, privacy
preserved, security on the net."

Example 3 of an input that produces an output without personal reference:

58 Input:
"Name the first three elements of the periodic table."

60 Issue:
"The first three elements of the periodic table are hydrogen (H), helium (He) and lithium (Li)."

Example of an input that produces an output with a personal reference:

61 *The soccer coach asks an AI: "Suggest the best players for the line-up."*

3.2 Special care with special categories of personal data

62 The Union legislator considers special categories of personal data to be particularly worthy of protection. These include, for example, personal data revealing religious or philosophical beliefs, trade union membership or a preference for a particular political party, as well as health data, genetic or biometric data. This includes, for example, the intake of certain medication or regular attendance at a particular church. The processing of such data is generally prohibited under Art. 9 para. 1 GDPR and is only permitted in exceptional cases under the conditions of Art. 9 para. 2 to 4 GDPR. In this respect, it must therefore be checked both with regard to the input and with regard to the processing and output of specially protected data whether one of the exceptions of Art. 9 para. 2 GDPR is fulfilled.

Example:

63 The use of AI-based systems plays an important role in skin diagnostics, particularly in the early detection of cancer in medical practices, and complements medical diagnostics. If an AI application meets the professional standard and is approved as a medical device, Art. 9 para. 2 lit. h GDPR in conjunction with the treatment contract may be suitable for the processing of patient data. Otherwise, informed explicit consent pursuant to Art. 9 para. 2 lit. a GDPR may be considered, which must be preceded by information and explanations about the specific functioning of the AI application.

3.3 Check results for correctness

64 The results of AI applications with personal references must be critically scrutinized. The providers of LLM in particular often make it clear that the data collected with their

application do not claim to be correct and should always be questioned. Furthermore, AI applications can have different levels of information.

With regard to personal results or a personal application of the results, however, incorrect results can lead to unauthorized processing, so that a check must be carried out before further processing.

3.4 Check results and procedures for discrimination

Irrespective of their factual accuracy or their own personal reference[66], the results of AI applications can also lead to unlawful processing of personal data, e.g. if they have a discriminatory effect. Data processing based on this may be unlawful if, for example, it is intended to violate the General Equal Treatment Act (AGG) and therefore does not fulfill the legal basis of Art. 6 para. 1 lit. f GDPR. Data controllers must therefore check whether the results of an AI application are suitable for further use.
are acceptable within the legal framework.

Even if discrimination is not recognizable in individual results[67]
these have a discriminatory effect.

Example:

> An AI application could make the following recommendation based on previous successful applications[68] a company: *"Male applicants should be preferred for the vacancy."*

Applying this result to applicants in an application procedure would constitute unlawful processing of the applicants' gender data, as it would violate Section 7 (1) in conjunction with Section 1 (1) AGG by discriminating against non-male persons. in conjunction with Section 1 (1) AGG by discriminating against non-male persons.