GDPR when using by generative AI

The Privacy Protection Authority's part of the government's mission to develop guidelines for the use of generative AI in public administration

Diary number
IMY-2024-9162

Date
2025-02-05

Diary number:
IMY-2024-9162

Date:
2025-02-05

GDPR when using generative AI

Table of contents

Summary

Public administrations can reconcile the use of generative AI with the General Data Protection Regulation (GDPR). In essence, the following needs to be done. What is stated here is the same as the boxes below at the beginning of each section.

Consider data protection regulations as a starting point. Businesses that intend to use generative AI often need to ensure that the use is compatible with data protection regulations, that is, mainly the General Data Protection Regulation (GDPR) and supplementary legislation.

Use generative AI according to data protection principles. The use of generative AI can be done in a way that is consistent with the fundamental principles of data protection. This requires that measures are taken, among other things, to minimize the use of personal data and that the processing of personal data is carried out in a transparent manner in relation to the individual. An organization must ensure that the principles are complied with before the processing of personal data begins and also apply them on an ongoing basis for as long as the processing is ongoing.

Clarify the division of roles between data controller and data processor. The business that uses generative AI to perform its mission is the data controller for the processing of personal data. A provider of an AI system that processes personal data on behalf of the business may be a data processor. In such cases, a data processor agreement is required between the controller and the processor.

Ensure that there is a legal basis and other legal support for the use. When generative AI is used to fulfill the mission of the business, there may be a legal basis for processing personal data. The requirement of necessity may be met if the use is to make the business more efficient. The greater the risks associated with the processing of personal data that occurs when using generative AI, the higher the requirements for the legal basis.

Consider whether automated decision-making and transfer to third countries occur. When using generative AI, provisions in the Data Protection Regulation regarding automated decision-making and the transfer of personal data to third countries may be relevant. In order for authorities to be able to use generative AI for automated decision-making, the requirements of the Administrative Procedure Act on proportionality, objectivity and legality need to be followed. If the use of AI means that personal data is transferred to a country outside the EU/EEA, specific conditions under the Data Protection Regulation must be met.

Ensure individual rights in the use of generative AI. Businesses need to take measures to ensure the rights of individuals under GDPR when using generative AI. For example, the business needs to consider whether updates to the business's privacy policy or similar documents are required to inform individuals about the processing of their personal data. Furthermore, the business should strive to use generative AI with built-in protections for individuals' rights.

Assess the risks of use and ensure appropriate safety. It is possible to reconcile the use of generative AI with the security provisions of the General Data Protection Regulation. The security measures that need to be taken depend on of the risk of the processing. Activities that intend to use generative AI therefore need to carry out risk assessments before starting a processing with generative AI. In addition, the use needs to be continuously monitored from a security perspective. An impact assessment under the Data Protection Regulation needs to be carried out if the processing of personal data is likely to entail high risks for individuals. Examples of risks with the use of generative AI include overconfidence in the capabilities of the AI system, a lack of understanding of its limitations and the risk of data leakage. These risks may require measures such as a clear governance structure, human control and technical security measures.

Especially about the use of widely available and integrated generative AI. The use of publicly available generative AI services or integrated generative AI services should be under the management and oversight of the business. It should be clear to employees what the service may be used for and what information may be processed in the service. When using integrated generative AI services, which means that the service has access to business data, good information management and access control are required. It is important to have knowledge of the tools used and their capabilities and limitations.

Introduction

Background

In the summer of 2024, the Swedish Data Protection Authority (IMY) together with the Swedish Digital Governance Agency (Digg) was commissioned by the government to develop guidelines to promote the efficient and appropriate use of generative artificial intelligence (AI) in public administration. According to the assignment, IMY, as an independent data protection authority, would develop and decide on the part of the guidance that concerned data protection. Digg would coordinate the work of the authorities also in the part that concerned data protection.[1]In January 2025, the results of the assignment were published in the form of guiding guidelines from both Digg and IMY on Digg's website: *digg.se/ai*.[2]

The report

This report contains the part of the guidelines that IMY has developed and decided on, that is, the part that concerns data protection. The reasons for publishing the content in this report are that IMY wants to meet a demand for the report format, among other things because it facilitates reference to the content, and to be able to publish the content in IMY's own channels and thereby increase accessibility.

Issues involved

Businesses that intend to use generative AI need to consider data protection and often ensure that the use is compliant with data protection regulations. Questions that then need to be answered are provided in this report. These questions are briefly outlined as follows.

- Does data protection regulation need to be considered when using generative AI?
- Is the use compatible with the fundamental principles of data protection law?
- Who is the data controller? Are there data processors?
- Is there a legal basis and other legal support for the processing of personal data?
- Does automated decision-making or transfer of personal data to third countries occur during use and, if so, are there conditions for this?
- How should individual rights be met during use?
- How should appropriate safety be achieved during use?
- What risks from a data protection perspective does the use of generative AI entail? What measures can be taken to limit the risks?

Especially about widely available and integrated generative AI
A separate section provides more detailed guidance for the use of generative AI that is widely available via the internet and for generative AI that may be integrated into existing office and enterprise software.

[1] Government decision 2024-07-04, Fi2024/01535,available here .
[2] Seehttps://www.digg.se/ai-for-offentlig-forvaltning/riktlinjer-for-generativ-ai .

Addresses public administration

The guidelines are intended for operations within public administration. This does not prevent the guidelines from being useful for others who use generative AI in their operations.

Development and fine-tuning not included

Developing generative AI systems, that is, creating a basic model that can be used to solve different types of tasks, is a resource-intensive process that requires significant investments, large amounts of data and an extensive technical infrastructure. Public administration operations generally have limited opportunities to develop such systems themselves from scratch.

An alternative to developing a new basic model is to fine-tune an existing model to adapt it to the specific needs of the business. Even fine-tuning existing basic models is an advanced process that requires specialist expertise, access to larger data sets and significant technical resources. This is something that most public administration operations lack today.

Against this background, the guidelines neither address the development of new basic models nor the fine-tuning of existing basic models. Instead, the guidelines focus on how public organizations can use existing generative AI systems in a responsible and legally secure manner.

Consider data protection regulations as a starting point

Businesses that intend to use generative AI often need to ensure that the use is compatible with data protection regulations, that is, mainly the General Data Protection Regulation (GDPR) and supplementary legislation.

Data protection regulations must be taken into account if personal data is processed

A large part of the public administration's use of generative AI can be assumed to involve the processing of personal data. For example, personal data is processed when generative AI is provided with data that contains personal data. If personal data is processed when using generative AI, this means that data protection regulations must be taken into account.

The data protection regulations should also be taken into account as a starting point for other uses.

Even in situations where businesses do not intend to process personal data with generative AI, data protection regulations may need to be taken into account when using it. This is because a basic model (also called a general-purpose AI model or AI model) that is present in generative AI systems often has personal data learned since the development of the model. Such personal data can be processed, for example, by an employee intentionally or unintentionally instructing the AI system in a way that causes it to enter personal data that has been learned. It can also happen by an external party attacking the system and then obtaining personal data that has been learned.

As a starting point, public administration activities can therefore assume that data protection regulations should be taken into account when using generative AI. However, in some cases, measures may have been taken to ensure the anonymization of AI systems, which means that data protection regulations do not apply. To achieve anonymization, it is necessary that the risk is negligible that both users of the system and external parties can obtain personal data from the AI system.[3]

What is meant by data protection regulations?

The data protection regulations refer here to rules that concern the processing of personal data. These rules are primarily the EU Data Protection Regulation, commonly referred to as GDPR, and supplementary national legislation. Examples of such supplementary national legislation are the so-called Data Protection Act and sector-specific regulations on the processing of personal data, so-called register statutes. More information about the data protection regulations can be found on the IMY website, for example onwhat generally applies according to the Data Protection Regulation and how data protection laws are interconnected .

[3] See the European Data Protection Board (EDPB)*Opinion 28/2024 on certain data protection aspects related to the processing of personal data in the context of AI models*, adopted on 17 December 2024, https://www.edpb.europa.eu/system/files/2024-12/edpb_opinion_202428_ai-models_en.pdf , sections 3.1 – 3.2, especially paragraphs 31 and 43.

Use generative AI according to data protection principles

The use of generative AI can be done in a way that is consistent with the fundamental principles of data protection. This requires that measures are taken, among other things, to minimize the use of personal data and that the processing of personal data is carried out in a transparent manner in relation to the individual. An organization must ensure that the principles are complied with before the processing of personal data begins and also apply them on an ongoing basis for as long as the processing is ongoing.

The principle of accountability

The business is responsible for the use

The principle of accountability means that the business is responsible for ensuring and being able to demonstrate that the GDPR is being complied with. The use of generative AI should be approved by the business in advance to ensure this. It must also be clear to employees what generative AI may be used for and what information may be processed in generative AI systems.

Develop guidance documents for the use of generative AI

An organization that intends to use generative AI should develop and implement a policy or similar governing document that regulates its use. The documents should clearly describe how generative AI may be used within the organization and what requirements are set to ensure responsible handling. The documents should, among other things, clarify whether and under what conditions personal data may be processed when the organization uses generative AI. The organization should describe the considerations and measures taken to ensure that the use is in accordance with the data protection regulations. To ensure compliance, everyone involved within the organization should be provided with information about and have access to these documents.

The business should also develop procedures or similar control documents to manage risks systematically and ensure regular audits of the AI system's functioning in an appropriate manner.

Document the use

An important part of compliance is documenting the processing of personal data. If the processing is likely to result in a high risk to the rights and freedoms of individuals, the business needs to carry out an impact assessment. Read more about this in the section *Assess the risks of use and ensure appropriate safety,*especially under the heading*Impact assessment for generative AI.*

Businesses using generative AI should also implement appropriate traceability mechanisms when processing personal data, making it possible to verify and deduce how the AI system has been used.

The principle of legality, regularity and transparency

Personal data must be processed lawfully

The principle of lawfulness means, among other things, that the processing of personal data must have a legal basis. More information about this can be found in the section on*Ensure that there is a legal basis and other legal support for the use.*

Personal data must be processed correctly

The principle of fairness means that the processing should be fair, reasonable and reasonable and that the data is processed in a way that individuals can reasonably expect. The principle also requires proportionality. This means that the privacy infringement that the processing entails needs to be reasonable in relation to the benefit of the processing.

Bias

For example, in order to process personal data correctly, measures are required to minimize bias.,that can occur in the output of an AI system. Bias can arise for a number of reasons and can, for example, stem from the quality of the data on which the base model has been trained. For example, if the training data of the base model has not been a representative sample of data for the purpose, the model's output will reflect these biases. This in turn can lead to discrimination or unfair treatment of individuals or groups of individuals. Businesses should therefore evaluate the risks of a treatment and whether it means that the model can generate distorted or biased information.

In the section*Assess the risks of use and ensure appropriate safety*There is more information about bias and measures that can be taken to limit the risks in this regard.

Personal data shall be processed in a transparent manner

The principle of openness requires transparency about how the AI system is used to process personal data and what consequences it may have for individuals. Informing individuals when their personal data is being processed is necessary for them to be able to take steps to protect their rights. At the same time, it can be challenging, for example, to explain in an easy-to-understand way how an AI system works or has arrived at a certain answer. As technology develops rapidly, it is important that individuals receive up-to-date information about its use and how it affects them.

In the section*Ensure individual rights in the use of generative AI*more guidance is provided on how personal data can be processed in a transparent manner.

Purpose limitation principle

Specify the purpose

The purpose limitation principle means that personal data should be collected for specified, explicit and legitimate purposes and not further processed in a manner incompatible with those purposes.

Generative AI is a type of technology that can be used for a variety of purposes. However, using generative AI is not an end in itself, but a means to process information. It is important that the organization specifies the purposes for which the AI system will be used and that this is made clear to the organization's employees.

Consider technical limitations

The business may also consider introducing measures to delimit the scope of use of the AI system, for example by technically limiting what instructions and information users can provide to the system through so-called prompts. In some cases, the functionality of the AI system may need to be limited so that it is only technically possible to perform tasks that are relevant to the purpose of the processing.

The principle of data minimization

Limit the processing of personal data to what is necessary for the use The business must be able to demonstrate that the extent of the personal data processing carried out using generative AI is necessary in relation to the purpose and that the same result cannot be achieved with a smaller amount of personal data. By providing only the most relevant personal data in a prompt,
personal data processing is limited to what is necessary to perform the task. If the business uses techniques that mean that an AI system is given other access to personal data, the business should evaluate and, if appropriate, limit the amount of personal data that the system can access when it is used.

Ensure access control
Permission controls that ensure which data specific users have access to should be reflected in the use of AI systems. The same applies when the business uses generative AI systems that are integrated into the business's existing software, for example, additional services that can be activated in the software. This ensures that personal data is only retrieved from relevant sources and that access is limited to sources to which the user has permission.

The principle of correctness

Risk of hallucinations
Businesses that process personal data must ensure that the data is accurate and, if necessary, updated. Generative AI creates its results based on statistical probabilities, which means that it can create content that is factually incorrect or fabricated. This is usually referred to as hallucinations. There are several reasons why an AI system can hallucinate. For example, it can depend on how the basic model has been trained, what parameters have been used and the quality and amount of data used during training. Hallucinations can also occur due to how questions are asked to the AI system, as questions can intentionally or unintentionally be asked in a way that leads the AI system to generate answers that contain hallucinations. Anyone using a generative AI system should therefore review and check its answers and not assume that the system's output is accurate and reliable.

To generate more accurate and precise answers, techniques such as*retrievalaugmented generation*(RAG) can be helpful. RAG combines the generative capabilities of the basic model with additional information from selected sources, giving the model access to more up-to-date information than is available in its original training data. RAG can thus improve the responses generated by the AI system. However, it should be emphasized that it is still necessary to carefully review and control the responses of the AI system, as hallucinations occur even when RAG is used.

Read more about hallucinations and measures to ensure accuracy in the section *Assess the risks of use and ensure appropriate safety.*

The principle of storage minimization

Avoid saving irrelevant personal data
The principle of retention minimization means that businesses need to ensure that personal data is not processed for longer than necessary. In order to effectively use generative AI on business data, it may be necessary in some cases to

data is processed (for example, vectorized) to make it searchable and accessible to the AI system. It can therefore be challenging to strike a balance between, on the one hand, the need to store data that may be relevant to the purpose of use, and, on the other hand, the requirement for storage minimization. The business needs to ensure that the AI system does not process personal data that is no longer relevant in a way that poses risks to individuals. When it is no longer necessary to process the personal data based on the purpose, it should be deleted or anonymized. This also applies to data that, for example, a generative AI system uses, as well as any copies of data sets that are stored somewhere other than internally in the business.

The principle of privacy and confidentiality

Risk-based approach

There is no universally applicable standard for how appropriate security should be achieved when personal data is processed in connection with the use of generative AI systems. The level of security and the security measures that need to be taken depend on the risk in the individual case. The business must identify the risks that the planned use of generative AI entails and, where applicable, take appropriate security measures to mitigate those risks.

Since AI systems are often integrated into complex chains of other systems, data flows and processes, organizations should take a holistic approach to ensuring the security of personal data. Generative AI technology is evolving rapidly, as are methods for securely processing personal data. As part of complying with the principle of privacy and confidentiality, organizations should actively monitor and address the latest technological developments, including methods that increase the resilience of generative AI models to various types of attacks.

Read more about the security of personal data in the section*Assess the risks of use and ensure appropriate safety.*

Further reading

The IMY website provides general guidance onthe basic principles.

Clarify the division of roles between data controller and data processor

The business that uses generative AI to perform its mission is the data controller for the processing of personal data. A provider of an AI system that processes personal data on behalf of the business may be a data processor. In such cases, a data processor agreement is required between the controller and the processor.

Personal data responsibility

To ensure effective protection of personal data, there is always one or more data controllers, i.e. the person responsible for compliance with the data protection regulations. The data controller is the person who determines the purposes and means of the processing of personal data, i.e. how and why the personal data is processed. In public administration, it is normally an authority that is the data controller.

If an organization processes personal data on behalf of a data controller, that organization is called a data processor. The Data Protection Regulation requires that there is a data processor agreement between the controller and the processor.

The business is the data controller for its use of generative AI

The public administration activity that uses generative AI to carry out its mission is, as a starting point, the data controller for the personal data processing that occurs in connection with the use. With this responsibility comes an obligation to choose generative AI systems that enable the business to ensure that the processing is carried out in accordance with the General Data Protection Regulation.

Suppliers of the AI system are, as a starting point, personal data processors

If a business engages a supplier to provide an AI system for a specific task, the supplier is a data processor for the personal data processing that the supplier performs on behalf of the business. A prerequisite for the supplier not also being considered a data controller is that it does not have its own purpose in processing personal data, for example, using the data to further develop the basic model of the AI system. The supplier's sole interest in selling an AI system is not a purpose that in itself entails data responsibility.

Responsibility for unforeseen treatments

It may be difficult to predict how a generative AI system may function in practice. However, as a starting point, an organization is responsible for the processing of personal data that occurs with generative AI even when personal data is processed in an unforeseen way. Exceptions may apply to functions that the organization has explicitly objected to and where clear instructions have been given to the provider that such functions should not be included in the AI system. It is therefore important to consider what limitations should be built into the system.

Further reading

The IMY website provides general guidance ondata protection officer and data processors.

Ensure that there is a legal basis and other legal support for the use

When generative AI is used to fulfill the mission of the business, there may be a legal basis for processing personal data. The requirement of necessity may be met if the use is to make the business more efficient. The greater the risks associated with the processing of personal data that occurs when using generative AI, the higher the requirements for the legal basis.

AI is a means, not an end

Any processing of personal data must be for one or more legitimate purposes. The purposes need to be specific and explicitly stated. Using AI is not such a purpose in itself, but a means of processing information. The processing of personal data that occurs when using generative AI, however, can be carried out for several different purposes at the same time.

Use generative AI to meet business needs

The starting point is the mission

For any processing of personal data, there needs to be a legal basis. The legal basis for processing personal data in public administration is normally that the activity is to perform a task of public interest. Tasks given to authorities and other public bodies are tasks of public interest. Such tasks can, for example, be given by law or regulation or a government decision. To the extent that generative AI contributes to the activity performing its mission, there may be support for processing personal data with such technology.

Efficiency is key

In order for personal data to be processed for a task of public interest, the processing must be *necessary*to perform the task. The requirement of necessity may be met if the use of generative AI contributes to making the operation more efficient. Efficiency is a requirement that applies to large parts of public administration, for example as a result of provisions in the Authority Ordinance, the Administrative Procedure Act and provisions on financial management in the Local Government Act.

Consider how risky the treatment is

In order for the legal basis*task of public interest*To be used, it is required that the information is of public interest, for example in a law or other regulation or a government decision. The greater the risks of the intended processing, the higher the requirements for the clarity, precision and predictability of this legal support.

Less risky treatments

Processing with lower risks, such as processing a smaller amount of non-sensitive personal data, can be carried out with the support of a more general legal basis. This could be the case, for example, if generative AI is used to summarise the content of a public report where personal data only appears in the form of the names of the report's authors and responsible decision-makers. In these cases, it can constitute an efficiency gain to summarize the report with generative AI and no further considerations about the legal basis are then required.

Riskier treatments

For processing operations involving higher risks, such as processing large amounts of sensitive personal data, higher requirements are placed on the legal support in terms of clarity, precision and foreseeability. In such cases, special regulation may be required that makes it foreseeable that the processing will take place.

Sensitive personal data

Sensitive personal data, such as health data, are only permitted to be processed if there is specific support for it. This also applies when using generative AI systems.

The European Court of Justice has made a broad interpretation of what should be considered sensitive personal data in relation to technology that has the ability to process large amounts of information.[4]When using generative AI, it may mean that the business needs to assess whether there is support for processing sensitive personal data, even if the intention is not to process such data.[5]

Example

In the reportDisclosure of public documents using AI IMY analyzed the use of generative AI at a municipality to facilitate the handling of requests for public documents according to the principle of public access, including in the form of a so-called RAG solution (eng. *retrieval-augmented generation*). The legal basis for processing ordinary personal data and sensitive personal data was analysed in the report. Two different cases were addressed: a more comprehensive case called the comprehensive service and a narrower case called the masking service.

Overall, IMY assessed that there was much to support the processing of personal data, including sensitive data, in the narrower masking service. However, it was considered that there were overriding reasons to argue against the existence of legal support for the processing of personal data in the more comprehensive comprehensive service. The legal support in the two cases was essentially the same. The difference in assessment was determined by the different risks that the two cases entailed, which is an example of how the greater the risks of the intended processing, the higher the demands placed on the legal support.

Further reading

The IMY website provides general guidance onlegal basis and sensitive personal data.

[4] See judgment of the Court of Justice of the European Union 2023-07-04, Bundeskartellamt, C-252/21, ECLI:EU:C:2023:537, question 2 a in particular p. 73. See also judgment of the Court of Justice of the European Union 2024-10-04, Lindenapotheke, C-21/23, ECLI:EU:C:2024:846, question 2 paragraphs 86-88.

[5] See e.g. Claudio Novelli et al., Computer Law & Security Review, Volume 55, November 2024, https://doi.org/10.1016/j.clsr.2024.106066 , section 3.1.2 on p. 5 f.

Consider whether automated decision-making and transfer to third countries occur

When using generative AI, provisions in the Data Protection Regulation regarding automated decision-making and the transfer of personal data to third countries may be relevant. In order for authorities to be able to use generative AI for automated decision-making, the requirements of the Administrative Procedure Act on proportionality, objectivity and legality need to be followed. If the use of AI means that personal data is transferred to a country outside the EU/EEA, specific conditions under the Data Protection Regulation must be met.

Automated decision-making

Generative AI can be used to make decisions about individuals. There are provisions in the General Data Protection Regulation that regulate the situations in which automated decision-making is permitted. However, the provisions only apply to decisions that have legal consequences for or significantly affect the individual, for example in the exercise of official authority.

What is automated decision-making?

Automated decision-making means that decisions are made without human intervention. For something to be considered automated decision-making under GDPR, the decision-making itself must be done through automated processing of personal data. If generative AI is used to support decision-making but a human being makes the actual decision, this is not automated decision-making under the GDPR. However, if a human being relies in practice on the basis generated by an AI system in a decisive way, this can be considered automated decision-making under the GDPR even if the formal decision is made by the person in question.[6]

Is automated decision-making allowed?

The main rule in the Data Protection Regulation is that automated decision-making is prohibited. However, there are a number of exceptions to the prohibition. One exception is that there is regulation that allows automated decision-making and that this regulation contains appropriate safeguards. The Government has assessed that the Swedish regulation in the Administrative Procedure Act, the Local Government Act and certain special legislation provides the conditions for automated decision-making under the Data Protection Regulation.[7]

In order for generative AI to be able to form the basis for automated decision-making within authorities, it is necessary that the requirements of the Administrative Procedure Act for proportionality, objectivity and legality are complied with. It is therefore necessary to ensure that decision-making follows applicable rules, is predictable and that irrelevant considerations are not taken. This means, among other things, that it must be ensured that partiality, so-called bias, and hallucinations do not affect individuals in a negative way. If this cannot be ensured to a sufficient extent, other solutions need to be chosen, for example a rules-based system.

Further reading

More information about IMY's innovation portal is available.automated decision-making .

[6] See the judgment of the European Court of Justice 2023-12-07, SCHUFA Holding (Scoring), C-634/21, ECLI:EU:C:2023:957, p. 73.

[7] See, for example, Bill 2017/18:95 p. 100 and Bill 2021/22:125 p. 58.

Transfer of personal data to third countries

The use of generative AI can often mean that personal data is transferred to an external service provider. If the personal data is then transferred to a country outside the EU/EEA, i.e. a third country, the transfer is only permitted if specific conditions in the Data Protection Regulation are met. For transfers of personal data to the USA, such conditions may exist if the provider is covered by the so-called EU-US Data Privacy Framework. An organization considering using generative AI that involves the transfer of personal data to a third country needs to check whether there are conditions for this.

More information about IMY is available on the IMY website.transfer of personal data to third countries .

Ensure individual rights in the use of generative AI

Businesses need to take steps to ensure that individuals' rights under the GDPR are protected when using generative AI. For example, businesses need to consider whether updates to their privacy policy or similar documents are required to inform individuals about the processing of their personal data. Businesses should also strive to use generative AI with built-in protections for individuals' rights.

Introduction

According to the General Data Protection Regulation, an organisation must, on its own initiative, provide certain information to individuals about the organisation's processing of personal data. This is often referred to as the right to information. Individuals also have certain rights that they can exercise on their own initiative, such as the right to access, rectification and erasure.

The IMY website contains general guidance information about these so-called data subject rights , for example, on the deadlines for responding to a request from an individual. The IMY website also provides more specific guidance on, for example,The black box and the right to information , which includes information about automated decision-making.

Personal data in input and output data

Below are guidelines for personal data in input to and output from generative AI systems. Input refers primarily to instructions and context that are supplied to the AI system. Output refers primarily to what the AI system generates, i.e. the result.

Examples regarding the right to information

According to the right to information, businesses must inform individuals about, among other things, which recipients or categories of recipients have access to their personal data. For example, if a business uses generative AI with a RAG solution (eng.*retrievalaugmented generation)*which involves the processing of personal data, this may mean that the data is transferred to an external supplier. This supplier then processes the personal data on behalf of the business and acts as
personal data processor in this regard. In such cases, the supplier is a type of recipient about whom the business is obliged to inform individuals.

The right to information also means that individuals must be informed about, among other things, transfers of personal data to a third country, i.e. a country outside the EU/EEA, and the legal basis for the transfer. This may be relevant when using generative AI.

These requirements for clear information may mean that a business using generative AI needs to review and update its privacy policies and similar documents with information to individuals about how their personal data is processed when using generative AI.

Examples for the right of access

The right of access means that individuals, as a general rule, have the right to know, upon request, whether an organization is processing personal data concerning them. If so, they have the right to also the right to receive a copy of this data and information about how it is used. This right generally also includes the personal data used as input data and generated as output data by a generative AI system.

However, there are exceptions to the right of access that may apply to both input and output data. The right of access does not apply to personal data that may not be disclosed to the individual due to confidentiality or professional secrecy. Exceptions also apply to personal data in continuous text that has not yet reached its final form at the time of the request, as well as to personal data contained in memos or similar documents. However, the exceptions for continuous text and memos only apply if the documents have not been disclosed to a third party. In this context, personal data processors are not considered third parties.

It is possible that prompts containing personal data could be considered as running text that has not yet been finalized or as a memo. This would mean that the prompt could be exempted from the right of access, provided that it has not been disclosed to a third party. Output data generated based on the prompt could similarly be considered as running text or a memo. However, if the output data is recorded or forwarded, it is generally covered by the right of access, but may still be exempt if the information is subject to confidentiality or professional secrecy.

Personal data that may be contained in AI models

It can be challenging to meet individual rights Generative AI systems may have personal data learned from the development of the basic model. This may involve data on a very large number of people. It is currently almost impossible for a business using a generative AI system to be able to determine whose personal data has been processed in connection with the development of the basic model and which may be learned in the model. This makes it challenging for businesses planning to use generative AI to be able to satisfy individuals' rights under the Data Protection Regulation. Against this background, the following guidelines are provided.

The right to information

Provide general information about how generative AI is used in business Businesses do not have to voluntarily disclose information about their use of a generative AI system to all individuals whose personal data may have been used as training data to develop the system's base model. This would entail a disproportionate effort under the provisions of the General Data Protection Regulation. Instead, businesses should provide information to the public about which generative AI systems or base models are used, for example by including such information in a privacy policy or AI policy that is publicly available.

The right to access, rectification and deletion

Choose AI systems with built-in data protection as standard

Businesses should prioritize using generative AI systems that enable them to satisfy a request for the exercise of individuals' rights under
When choosing an AI system, the business should consider the latest technical developments in the field, the implementation costs and the risks that may arise for personal data in connection with the use of the AI system.

Facilitate the exercise of individuals' rights and explain any refusal If an organisation considers that it cannot comply with a request from an individual to exercise their rights under the GDPR, the organisation should still endeavour to facilitate this for the individual. This could be done, for example, by referring them to appropriate channels for making a corresponding request with the provider of the AI system that the organisation uses. The organisation should also explain the reasons why a request cannot be complied with, but the organisation is not obliged to respond to requests that are manifestly unfounded or unreasonable.

Assess the risks of use and ensure appropriate safety

It is possible to reconcile the use of generative AI with the GDPR's provisions on security. The security measures that need to be taken depend on the risk of the processing. Businesses that intend to use generative AI therefore need to carry out risk assessments before starting processing with generative AI. The use also needs to be continuously monitored from a security perspective. An impact assessment under the GDPR needs to be carried out if the processing of personal data is likely to entail high risks for individuals. Examples of risks with the use of generative AI include overconfidence in the capabilities of the AI system, a lack of understanding of its limitations and the risk of data leakage. These risks may require measures such as a clear governance structure, human control and technical security measures.

Method for appropriate security

The General Data Protection Regulation requires that security measures be taken that are appropriate in relation to the risk of the processing of personal data. Higher risks require more extensive and far-reaching measures, and lower requirements for less risky processing. Important steps in ensuring appropriate security are:

- *Analyze risks in advance:*Analyze and assess risks associated with use before introducing the AI system into the business. The purpose of this is to take appropriate measures in advance to reduce risks and not carry out treatments that pose too high a risk.
- *Monitor and manage risks during use:*Continuously monitor risks and take appropriate measures to reduce risks. Where appropriate, discontinue treatments that prove to be too risky.
- *Manage incidents:*Identify, analyze and resolve incidents. Inform affected persons and IMY as necessary.

Risks of using generative AI

The use of generative AI can entail several different risks. Examples are given below.

Overconfidence and inappropriate use

One risk with generative AI is that users in the business lack an understanding of the technology, especially when it comes to its limitations. This can lead to employees using the technology in their work relying too heavily on the results generated by the AI system. It can also lead to the technology being used in ways that are inappropriate from an information security perspective.

Data leak

Another risk is that personal data transferred to an external generative AI service may end up in the wrong hands or be used in an unauthorized manner outside the business's control.

The black box problem

One challenge with generative AI is the so-called black box problem.*the black box problem*), which means that it is not always possible to fully understand how the technology works or on what basis a response is generated. This can lead to security risks that we do not yet know or understand. Public organizations should therefore consider whether the use of generative AI is necessary or whether a more predictable, rule-based IT solution can fulfill the same function. Simpler AI solutions that can provide better control over and understanding of the processing should also be considered, for example a small language model (eng.*small language model, SLM*) that can be managed closer to the business.

Hallucinations

Hallucinations are when generative AI creates output that does not match reality. Examples of hallucinations include responses or actions that contain misleading or incorrect information. Some hallucinations can be very difficult to distinguish from facts. The more difficult it is to identify the incorrect, the greater the risk and consequences of the hallucination. Hallucinations can make it difficult to comply with the GDPR's principle of accuracy.

Bias

Bias means that a generative AI system produces results that are discriminatory or distorted. This is often because biases in the training data have affected the learning of the base model. For example, if the base model's training data shows that women generally earn less than men, the model may incorrectly learn that this is the norm and reproduce or reinforce such patterns in its results. Bias can make it difficult to comply with the GDPR's principle of correctness.

Appropriate measures to limit risks

Several different measures may be appropriate to limit risks associated with the use of generative AI. Examples are provided below.

Governance structure

To mitigate the risks of generative AI, it is usually appropriate to establish a clear governance structure, both internally and in relation to suppliers of generative AI systems. Such a governance structure may include guidelines and procedures regarding the use, security and handling of personal data. For example, there should be clear governance around which personal data may be used as input data and which personal data a basic model may be given access to in other ways.

Training

Businesses should train employees who will use generative AI in their work. Employees should learn to identify risks and it should be emphasized that the technology is a support tool that does not replace human judgment.

Human control

Another appropriate measure is to maintain human control.*human-in-the-loop*) by ensuring human involvement in key steps when using generative AI. Human control means that the results should be able to be reviewed and assessed by a human, which can be a challenge in some cases when the AI system generates complex or difficult-to-interpret responses. For example, the risk of hallucinations can be limited by having employees review the content that generative AI produces. Review can be facilitated if the AI system's basic model uses the method*chain of thought*(CoT) which means that the model presents its reasoning step by step, making its processes more transparent and easy to understand.

Technical security measures

To protect personal data, both at rest and in transit, it is important to use strong encryption. Furthermore, access control and authorization management solutions should be implemented to ensure that only authorized persons have access to the personal data transferred to the AI system, especially when using an external service. In addition, it is important to implement security measures that can detect and prevent unauthorized use of personal data, such as monitoring system activity and alarms in case of suspected anomalies.

To further strengthen security, businesses should regularly evaluate the performance and functionality of generative AI systems to identify and address any deficiencies.

Probability thresholds and temperature

A generative AI system is often designed to generate answers regardless of whether the answer is factually correct or not. Unlike a human, who can express uncertainty or admit that they do not know or understand, generative AI tends to always generate an answer, even in the face of uncertainty. One way to counteract this problem is to use so-called probability thresholds (Eng. *confidence thresholds*). The AI system can then be configured to refrain from generating an answer if the probability of the answer being correct is below a certain threshold. The limit for when an answer should be considered to be below the threshold cannot be determined generally and depends, among other things, on how high the so-called temperature is allowed to be.

Temperature is a parameter that controls how “creative” an AI system is allowed to be when generating output. A low temperature makes the system more predictable by prioritizing the most likely responses, while a high temperature creates more varied and creative results. Configuring the temperature can therefore also be an appropriate measure in this context.

RAG

To generate more accurate and precise answers, the technology can*retrieval-augmented generation*(RAG) can be used. RAG combines the generative capabilities of the basic model with external or business-specific information from selected sources. This allows the system to not only generate more qualitative answers, but also to answer questions based on the business's own data. It is important to note that RAG is no guarantee against bias or hallucinations, but this technology can still help reduce the risks of these problems.

It should also be noted that RAG carries its own risks. For example, advanced processes such as vectorization and indexing of data are required for the system to be able to use the information. This in itself can involve complex processing of personal data, which requires analyses to be carried out to ensure that the processing is necessary and proportionate in accordance with the General Data Protection Regulation. In addition, measures such as authorization control are required, where careful consideration should be given to which information the RAG solution should have access to.

PET

Various privacy-enhancing technologies, often called*privacy-enhancing technologies* (PET), can be an important part of the work to reduce the risks of generative AI. PET can ensure compliance with fundamental data protection principles, such as data minimization. PET can also increase security, especially in advanced personal data processing in connection with the use of generative AI.

Impact assessment for generative AI

For processing personal data that is likely to pose a high risk to individuals, the business is required to conduct a data protection impact assessment. This can be described as a process for identifying risks associated with the processing of personal data and developing measures to manage those risks.

The analysis of whether an impact assessment needs to be carried out before an organization starts using a generative AI system is no different from the corresponding analysis for other personal data processing. However, conducting a risk analysis before using technology that is new to the organization may require special expertise. Examples of factors that may arise when using generative AI and that indicate that an impact assessment needs to be carried out include the use of new technology, processing of personal data on a large scale, processing of sensitive personal data and automated decision-making.

The use of generative AI may therefore require an impact assessment, but this is not always the case. An assessment must be made on a case-by-case basis.

Further reading

The IMY website contains guiding information on, among other things:built-in data protection and data protection as standard ,information security ,impact assessments and prior consultations andpersonal data incidents More in-depth information can also be found in, for example, IMY's report.Disclosure of public documents using AI , see especially section 6.

Especially about the use of widely available and integrated generative AI

The use of publicly available generative AI services or integrated generative AI services should be under the management and oversight of the business. It should be clear to employees what the service may be used for and what information may be processed in the service. When using integrated generative AI services that mean that the service has access to business data, good information management and access control are required. It is important to have knowledge of the tools used and their capabilities and limitations.

Publicly available generative AI

The development of machine learning and computing power has led to the launch of a number of new services where users can access generative AI via user-friendly interfaces in cloud services and open APIs that can be used for a variety of purposes. Through this type of service, which is currently mainly developed and provided by American providers, generative AI has become available to the public. To start using these types of open services, the user usually only needs to create an account or take out a subscription with the provider.

Using an AI system based on generative AI may involve the processing of personal data that requires compliance with the requirements of the General Data Protection Regulation and other data protection legislation.

Use of generative AI needs to be controlled

For publicly available AI services, it is rarely possible to set requirements for the system's functions, security and information management. If employees are to be allowed to use such types of services at work, the organization must first decide how this should be done. The use may involve sensitive information being transferred to the AI service and a risk that it will become part of the training data of a basic model. The organization should develop internal guidelines that regulate the use in order to protect the organization's information and to ensure that the use is done in a secure manner that is also compatible with data protection legislation.

Processing of personal data in publicly available generative AI services Using publicly available generative AI services in a way that involves entering personal data into prompts or giving the service instructions that lead to the creation of personal data constitutes processing of personal data. Such processing must be carried out in accordance with the General Data Protection Regulation. An analysis of whether and how the use meets the requirements of the General Data Protection Regulation needs to be carried out before the services are put into use. There needs to be clear internal guidelines on and under what conditions personal data may be used as input data or generated as output data.

Data protection for the use of publicly available generative AI services

Generative AI services that are publicly available are often offered in different versions, from free options to subscription-based services or those governed by specific licensing agreements. Businesses using these services should always ensure that personal data is processed lawfully and is adequately protected, regardless of the option used.

Publicly available generative AI services without a data processing agreement When creating an account with a publicly available generative AI service, it may happen that there is no personal data processing agreement for the use of the service. This is often because the terms and conditions of some versions of these services require that the account be registered by a user as a private individual, with the aim of using the service for their own use. If an employee uses such a service in their work and in connection with it processes personal data, the employer may be

data controller for the processing. Since the employer in these cases has not entered into a data processing agreement with the supplier, the use may be in breach of the GDPR. Employers should therefore make it clear that these services may not be used by employees at work without prior approval and not without a data processing agreement that meets the requirements of the GDPR.

Publicly available generative AI services with data processing agreements When creating an account with a publicly available generative AI service where the terms of the agreement are aimed at non-private use, the provider in many cases acts as a data processor in relation to the user business. The provider of such services often offers data processing agreements that state how the provider processes personal data on behalf of the user business. For example, it can be agreed that the provider may not process personal data that the provider gains access to for its own purposes. A common problem, however, is that the provider often uses standardized data processing agreements and terms of use where there are limited or no opportunities to negotiate and adapt the terms to the needs of the business. In these situations, the business needs to assess whether it is possible to comply with the Data Protection Regulation with the provider's terms and conditions and otherwise refrain from using the service.

General advice when using publicly available generative AI services The use of generative AI services by employees for work-related purposes on their own initiative may result in the employer becoming the data controller for the processing. The use of these services should therefore instead be under the employer's control. The service should be procured by the business based on the intended use and the contractual terms need to ensure an adequate level of protection for the personal data.

Many providers of publicly available generative AI services place the responsibility on users to ensure that the service is used lawfully. This means that the business must ensure that the service is properly set up and configured to comply with the requirements of the General Data Protection Regulation.

Consider the appropriateness of use

An organization should consider whether generative AI services are appropriate to implement in the workplace based on the purpose of use. It should also consider whether the service should be used throughout the organization or only in certain parts. These considerations should be documented in internal governance documents that regulate how the service may be used and what type of information may be processed in it.

Integrated generative AI

In addition to generative AI being developed in new services, it has also become more common for generative AI to be integrated into existing office and enterprise software. This can range from simpler AI services that automatically generate text and content suggestions, to more advanced solutions where the AI service acts as a personal, digital assistant.

assistant that can be used for everything from analyzing data to suggesting more relevant and personalized search results based on the user's work context and preferences.

These services function as an addition to existing office software that can be activated by the user. Many public administration businesses today use office software with the ability to activate this type of additional services. Such tools usually involve connecting an AI service to the user's data and giving it access to information that the business processes. In simpler cases of integrated generative AI services, this may involve integrating a chatbot that answers general questions without having direct access to the business's internal information.

Many of these additional features of generative AI services are in an early stage of development and often offer limited opportunities for local and flexible adaptations. For example, this may involve limited possibilities to determine which data the AI service should have access to. If the business
If information management and access control are inadequate, it can lead to a user gaining access to information that they should not have access to. This in turn increases the risk that the AI service can also access and process data that the service should not be handling.

Processing of personal data in integrated generative AI solutions
Activating a generative AI service that is already integrated into existing software may mean that the AI service has access to information processed in the software. If this information contains personal data, the activation means that personal data is processed by the AI service. This circumstance does not necessarily mean that a new processing of personal data occurs, as the AI service may be intended to fulfill the same function as the software already used in the business. Whether the use of the AI service constitutes a new processing of personal data must be assessed on a case-by-case basis.

A generative AI service, such as a digital assistant with access to the business's data, can be used to process personal data in many different ways and for various purposes. An organization that intends to activate an integrated generative AI service should clearly define the purpose of its use. If the AI service is built as an additional function to the existing software, the purpose of using the AI service can often be linked to the same purpose as the use of the software in question.

New processing of personal data may arise The activation of an integrated generative AI service may lead to new personal data processing operations. This may be the case if the AI service being activated is able to collect user data by recording interactions with the service, including the instructions the user creates and the responses the service generates.

An organization planning to activate an integrated generative AI service should first analyze whether the use entails new processing of personal data and document this. If more personal data is processed or combined in new ways through the use of the AI service, it is important to map these new processing operations. The organization may also need to update its privacy policy or similar document.

Data protection for integrated generative AI services Using an integrated generative AI service in existing software often means that the business becomes the data controller for the processing that occurs through the AI service. The responsibility for ensuring that employees use integrated AI services in a legal and privacy-friendly manner usually lies with the business. Therefore, the business must have sufficient knowledge of the functionality of the service to ensure correct use.

General advice when using an integrated generative AI service Generative AI services integrated into existing software differ from traditional digital assistants in that they combine the functionality of basic models with access to the business's own data. The use of these AI services requires that the business has good overview and control over its information. The AI service should not be given more access than the data to which the individual user has permission. The business needs to have adequate information management in place before the AI service is given access to the business's data. This is a fundamental prerequisite for being able to comply with the Data Protection Regulation.

Information mapping may be necessary

To reduce risk and to be able to comply with the principle of accountability, the business should conduct an information mapping exercise. The purpose of such an exercise is, among other things, to identify what information exists, where it is stored, and who within the business has access to it. A carefully conducted information mapping exercise is also crucial to ensuring correct permissions when activating an AI service that accesses the business’s data.

The business should establish routines to continuously classify information and comply with any legal requirements for information management. This requires a concerted and careful effort from the entire business.

Consider new structures to adapt the business

Since the use of the AI service may involve handling the business's information in new ways, existing processes may need to be adapted. It is important for the business to gain knowledge about the functionality of the service in order to be able to integrate it in a privacy-friendly manner.

Enabling a generative AI service in existing software often requires organizations to implement new procedures, guidelines, and training to create a deeper understanding of the AI service's capabilities while also raising awareness of generative AI's potential, limitations, and risks.