# Moving forward responsibly

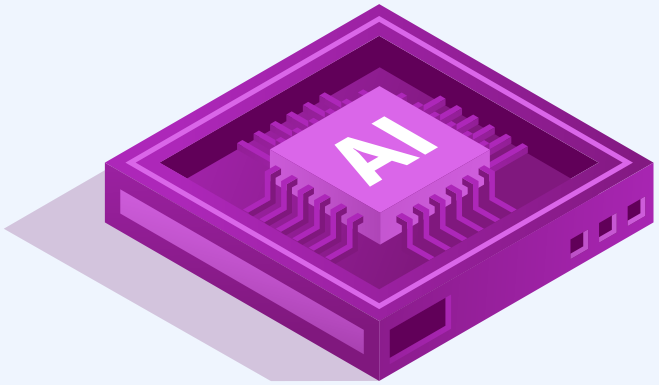
## GDPR preconditions for generative AI

AP | bescherming in een digitale wereld

Consultation version May 2025

# Table of contents






# Consultation

The Autoriteit Persoonsgegevens (AP) (Dutch Data Protection Authority) invites you to respond to this document by email to **genai-loket@autoriteitpersoonsgegevens.nl**. The consultation is open until 27 June 2025. We will summarise the responses we receive in one document, without including any names of persons and organisations or contact details. We will publish the summary on the AP's website and use it to further improve this document. The final version of this document will follow later in 2025.

# Management summary

1. **While generative AI has so far fallen short of complete legitimacy, the AP does see a possibility to work towards legitimate development and deployment of generative AI.** A legal analysis by the AP has uncovered some irregularities and uncertainties with regard to the GDPR. Subsequently, the development and deployment of this technology can be organised in way that prevents these irregularities and uncertainties. In this document, we will discuss the preconditions and possibilities in the light of the GDPR.

2. **According to the AP, it is plausible that irregularities occurred during the development of foundation models.** The overarching estimate is that, based on current practice, the vast majority of all generative AI models currently fall short in terms of legitimacy. Generative AI applications build on so-called 'foundation models'. To train these models, almost all publicly accessible data on the internet has been used (scraped). Special categories of personal data have been placed on the internet, which have not been made public by a data subject themselves. It is, therefore, plausible that a foundation model contains unlawfully collected special categories of personal data. The AP notes that these special categories of personal data form a small part of all the data collected to train foundation models. However, this does not alter the fact that these special categories of personal data have been obtained unlawfully. Nevertheless, the continued use of these foundation models by Dutch and European parties is, therefore, not inherently unlawful, as follows from an analysis by the EDPB.

3. **At application level, the AP has identified a number of preconditions, challenges and opportunities for responsible deployment of generative AI.** These applications have a clear purpose and are used for that purpose after a thorough risk assessment with appropriate safeguards. The AP sees further opportunities and challenges in the field of generative AI, which are separate from the framework conditions for data protection set out below. For example, creating awareness among users about sharing sensitive data in generative AI applications. The 'Forward responsibly: the AP's vision on generative AI' vision paper goes into these opportunities and challenges in greater detail.

4. **Preconditions for further processing of already collected personal data will be elaborated later.** There may be situations where a party wishes to further process its own personal data already collected in a generative AI model. Usually, this already collected data will, in itself, not be sufficient for training a foundation model. In that situation, the further processing deviates from the original purpose of the processing. It will then have to be assessed whether that further processing is compatible with the original purpose (under Article 6(4) of the GDPR). This analysis falls outside the scope of this document and may be addressed in a subsequent iteration of this document.

# 1. Introduction

The consultative vision paper entitled 'Forward responsibly: the AP's vision on generative AI' describes the AP's understanding of generative AI, the opportunities it offers to society and the risks it poses. In addition, the vision paper describes what needs to be done to embrace generative AI responsibly. For this, developers and users of generative AI will have to comply with the obligations under the GDPR when processing personal data in these models and/or applications. This document is particularly pertinent for professionals who develop generative AI or want to use it in their own business operations.

This document can be read on its own or as a complement to the 'Forward responsibly: the AP's vision on generative AI' vision paper. In this document, we will first provide a brief definition of generative AI. We will then set out the generative AI chain and discuss some non-exhaustive GDPR preconditions for generative AI. For each precondition, we will indicate to which parties it applies and where there are opportunities for responsible development and deployment.

The AP is aware that this technology is subject to rapid change and is currently much discussed in society. These GDPR preconditions therefore reflect the AP's view of the state of play at the time of publication (May 2025). The European Data Protection Board (hereinafter: EDPB) is currently drafting guidelines for generative AI, to which the AP is actively contributing. As a result, the following preconditions may be subject to change. The AP also stresses that, in addition to the preconditions set out below, other GDPR obligations may apply to the development or use of generative AI models and applications.

# 2. Basic understanding of generative AI

## Definition of Generative AI

Generative artificial intelligence (Generative AI) is a form of AI that is capable of generating new data. The most popular generative AI applications create texts and images that are all but impossible to identify as having been generated by AI. This guidance is about AI models that are able to generate realistic data and about the systems and applications that these models are part of.

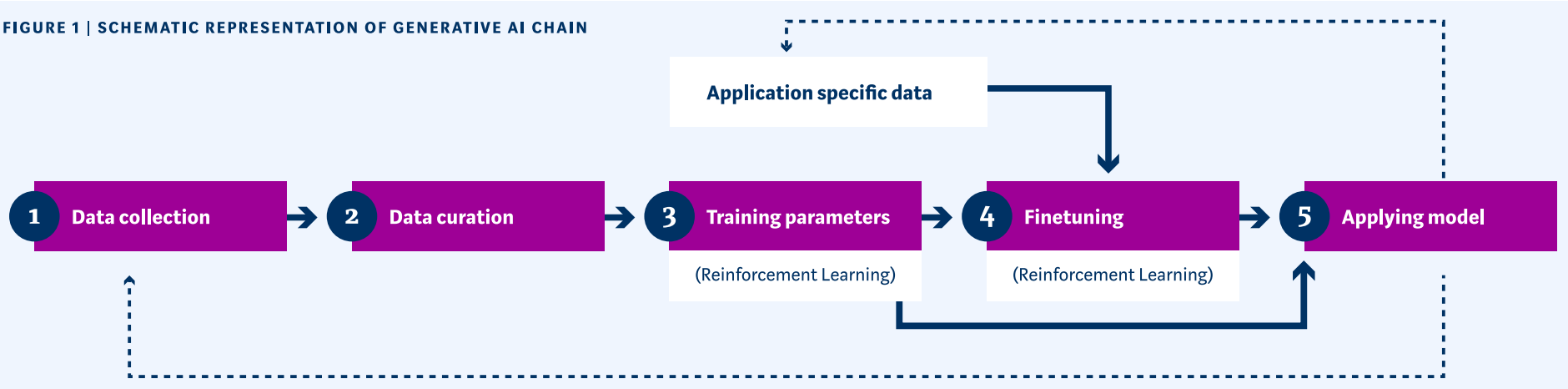
Generative AI models can generate and control all kinds of different forms of output. These models can serve as a foundation for many different specialised applications and can be used for all kinds of purposes. These types of models are often referred to as 'foundation models' or 'general-purpose AI models'. In our chosen approach, models with those denominators all fall under generative AI. Beyond the scope of this guidance, there are numerous other forms of AI. AI technology has been in development for a long time and is, for example, widely used for data classification. However, such models cannot generate data, which is the main difference with generative AI.

## Training a generative AI model

Simply put, training a generative AI model can be seen as a set of sequential steps:

1. **Data collection:** sample data is collected, often through scraping.
2. **Data curation:** unwanted examples (such as personal data or hateful content) are removed in a curation step.
3. **Training parameters:** the parameters are trained on the basis of patterns in sample data and possibly with reinforcement learning.
4. **Fine-tuning:** fine-tuning adapts the model to a specific application or limitation.
5. **Deployment of the model:** the outcome of these steps is a trained generative AI model that is deployed in an application.

FIGURE 1 | SCHEMATIC REPRESENTATION OF GENERATIVE AI CHAIN



Seeing as the generated data is often collected again for training, there is a feedback loop.

## Schematic representation of lawfulness

In the diagram below, we show different situations for the development and deployment of generative AI models. The question of whether the development and/or deployment of a generative AI model or application took place lawfully is not unequivocal. It depends greatly on the contextual factors in the development or deployment of such models and applications, as evidenced by the preconditions.

In the following situation, controller A develops a foundation model. This foundation model is then further fine-tuned and deployed by controller B.

Controller A has developed a foundation model in this scenario which contains personal data that has been unlawfully collected through scraping. The personal data was then curated. This unlawful collection of data may result from the lack of a legal basis under Article 6 GDPR. Or from the lack of the possibility to invoke an exception under Article 9 GDPR. Controller A then makes this model available to Controller B. The latter uses the foundation model to further fine-tune and/or build a generative AI application. Two scenarios can then be distinguished:

**Scenario 1: Controller A has verifiably anonymised all personal data in the foundation model**
According to the EDPB and the Court of Justice of the European Union (hereinafter: CJEU), the verifiable anonymity of a generative AI model should be assessed, inter alia, on the possibility of identifying the data subjects.[1] If controller B does not process personal data in the anonymised foundation model during the fine-tuning or deployment of its generative AI application, the GDPR no longer applies as established by the EDPB.[2] In that case, no personal data was used in further training the foundation model and controller B does not process personal data.

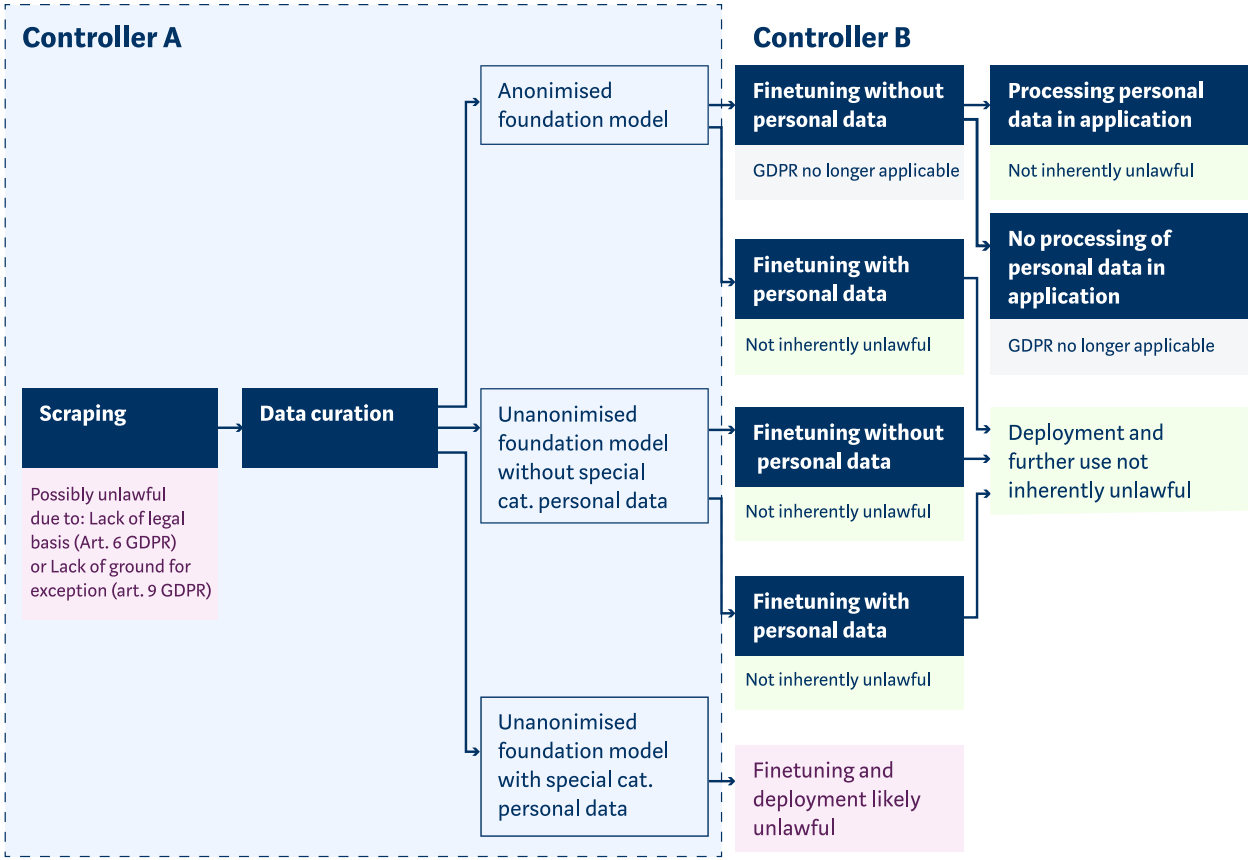
If controller B uses the anonymised fine-tuned model with a data set containing personal data or processes personal data when using it, the fine-tuning and use is not inherently unlawful. In that case, the unlawfulness of developing the foundation model does not affect controller B.[3] In that context, the EDPB considered that controller B itself needs a legal basis for the processing of personal data.[4] However, this is without prejudice to the fact that controller A may be held responsible by the competent authority for the unlawful acts committed during data collection, data curation and development of the foundation model.

**Scenario 2: Controller A has not (verifiably) anonymised the foundation model**
If controller A has not (verifiably) anonymised the foundation model, the GDPR will continue to apply. However, the unlawfulness of the processing by controller A does not mean that the processing carried out by controller B is inherently also unlawful. Controller B will have to carry out an adequate assessment of its accountability for the development of the model of unlawfully obtained personal data by Controller A.[5] The depth of that assessment depends on several factors. For example, what risks the use of the foundation model entails for the data subjects.[6] If it follows from this assessment that special categories of personal data were part of the data sets used to train the foundation model, controller B cannot use this model because of the processing ban under Article 9 GDPR.[7] After all, controller B will not be able to rely on an exception either.

If the assessment shows that no special categories of personal data were trained into the foundation model, fine-tuning and the further use of that foundation model is not inherently unlawful. Controller B will have to have its own legal basis for fine-tuning and deploying its generative AI models and applications. As in the first scenario, controller A remains responsible for the irregularities that occurred during data collection, data curation and development of the foundation model.

1. For a more detailed explanation of the anonymity of AI models, see: EDPB Opinion 28/2024 on certain data protection aspects related to the processing of personal data in the context of AI models, Chapter 3.2. See also CJEU *Breyer v Bundesrepublik Deutschland*, ECLI:EU:C:2016:779, para. 42.
2. EDPB Opinion 28/2024 on certain data protection aspects related to the processing of personal data in the context of AI models, para. 134.
3. Ibid., para. 135.
4. Ibid., para. 126.
5. Ibid., para. 129.
6. Ibid., para. 130.
7. For a more detailed explanation of 'special categories of personal data', see: What are personal data? Autoriteit Persoonsgegevens (AP)

**FIGURE 2 | SCHEMATIC REPRESENTATION LAWFULNESS GENERATIVE AI CHAIN**



## GDPR preconditions for generative AI

As the diagram above illustrates, the GDPR applies in several steps of the generative AI chain, as soon as personal data is processed. During collection, personal data is processed and even after the curation of data sets, these data sets may still contain personal data. For example, data sets on which generative AI models are trained or fine-tuned may contain personal data. Furthermore, personal data can be entered while using a generative AI application, as is the case with a generative AI application that is used as support in the health care sector. A generative AI application can, for example, create draft answers to patients' questions, which health-care providers then review and send.[8] Furthermore, even the increasingly popular 'AI agents' can process personal data,[9] such as an AI agent that assists in ordering products online. Typically, the AI agent will consult, among other things, location data and financial data for this purpose.

The GDPR imposes obligations not only on developers of generative AI models ('providers'),[10] but also on parties that further fine-tune these models and/or use them for their own use ('deployers').[11] Below we will set out some of the preconditions for the various steps in the generative AI chain: data collection, data curation, training and fine-tuning of the foundation model and deployment of

**FIGURE 3 | GDPR REQUIREMENTS FOR GENERATIVE AI**


8. See for example: TNO (April 2024), 'Generatieve AI in de Nederlandse Zorg,' <https://publications.tno.nl/publication/34643183/U5tb8oyL/TNO-2024-R10662.pdf>.
9. See for example: OpenAI (January 2025), 'Introducing Operator,' <https://openai.com/index/introducing-operator/>.
10. For a definition of 'provider', see Article 4(3) of the AI Act.
11. For a definition of 'deployer', see Article 4(4) of the AI Act.

generative AI applications. For each precondition, we will suggest how it might be relevant to certain parties, but it is up to providers and deployers themselves to determine which preconditions are relevant to them. The following preconditions therefore assume that personal data are processed in the various steps of the generative AI chain. The preconditions are shown schematically in the figure below.

**Preconditions for data collection**

*1. Data used to train and fine-tune a model must be lawfully obtained*

Relevant to: developers and fine-tuners

To train generative AI models, large amounts of data are needed. At the moment, it seems that the necessary extensive data sets for foundation models can only be gathered by using (untargeted) scraping. As a result, personal data almost certainly ends up in such data sets. In the event that proprietary previously gathered data sets containing personal data are used for the training of a foundation model, that data alone will usually not be sufficient to train a foundation model. For the collection of such personal data, the controller needs a legal basis, as described in Article 6 GDPR. There is usually no direct relationship between data subjects whose data is scraped from the internet and the generative AI developers who collect the data. Therefore, it seems that the 'legitimate interest' legal basis[12] can currently be used as the only possible legal basis for collecting data sets to train those foundation models. Whether a legitimate interest can be invoked depends on the EDPB Guidelines on legitimate interest.[13] In order for a legitimate interest to be successfully invoked, there must be a legitimate interest. In addition, the controller must pass the necessity test and carry out a balancing of interests. As previously stated in the AP's Scraping Guidance, it is not a given that reliance on the basis of legitimate interest will always be successful in the context of scraping. The conditions for a successful invocation of legitimate interest as a legal basis for processing must always be assessed on a case-by-case basis.

For the development of generative AI models and the deployment of generative AI applications, a number of conditions from the EDPB guidelines are important. In general, under the necessity test, stricter collection criteria for data collection are more likely to lead to a successful test result with regards to necessity. More targeted collection contributes to the proportionality of the processing of personal data. Data controllers will therefore have to check whether, and if so, how much, personal data should be scraped during data collection. When personal data is anonymised before being used in training a generative AI model, this makes a significant contribution to the success of the necessity test.

Stricter collection criteria may also play a role in the balancing of interests. Stricter collection criteria will generally reduce the impact on data subjects' rights by reducing the scope of personal data collected. In addition, the processing of personal data is usually an additional consequence and, therefore, a small part of training a generative AI model. Finally, anonymisation of personal data, before being used for training in a generative AI model, weighs heavily in favour of the controller.

After training a foundation model, the model can be further fine-tuned to smaller data sets. The developer of the foundation model can fine-tune the model. But a party that ultimately wants to integrate a fine-tuned model into a generative AI application can also do this. Typically, these data sets for fine-tuning are a lot smaller and not always obtained by scraping. If the data sets for fine-tuning contain personal data, the controller must have a legal basis, as described in Article 6 GDPR. However, the smaller size of these data sets makes it possible that, for example, data subjects' consent can also be used as a legal basis, or that there can be a compatible further processing of own data under Article 6(4) of the GDPR. For these smaller data sets, there may be a direct relationship between the data subjects and the controller. However, controllers will have to assess the applicable legal basis on a case-by-case basis and prior to initial collection of personal data.

**Stricter framework conditions for the collection of special categories of personal data**

Where special categories of personal data are collected during data collection, the controller must be able to rely on an exception to the prohibition on processing special categories of personal data.[14] This assessment will have to take place before the special categories of personal data are scraped along with all the other data. Fine-tuners of

12. Article 6(1)(f) of the GDPR.

13. European Data Protection Board (October 2024), 'Guidelines 1/2024 on processing of personal data based on article 6(1)(f) GDPR,' <https://www.edpb.europa.eu/system/files/2024-10/edpb_guidelines_202401_legitimateinterest_en.pdf>.

14. Article 9 GDPR.

foundation models may be able to rely on consent as a ground for an exception, because of the aforementioned direct relationship.[15] With (untargeted) scraping for training foundation models, this direct relationship usually does not exist and so an appeal to consent as a ground for an exception is nearly impossible. If the data subject has manifestly made the special categories of personal data public, this may apply as an exception to the processing ban in the case of untargeted scraping.[16] This will have to be assessed by the controller prior to data collection. If the controller cannot rely on an exception, the processing of those special categories of personal data cannot be considered lawful.

The situation where special categories of personal data are scraped without having been disclosed by data subjects themselves first, probably constitutes a small part of the entire data collection. However, that does not make the collection any less unlawful. With regard to the collection of special categories of personal data, the CJEU has ruled that the prohibition on processing may not be applied to an operator of a search engine as if that operator had published the special categories of personal data itself.[17] In such a case, the CJEU has stated that a search engine operator does not have to assess, prior to the data collection, whether an exception applies.[18]

15. Article 9(2)(a) GDPR.
16. Article 9(2)(e) GDPR. For a more detailed explanation of the apparent disclosure in scraping, see the AP's Guide to Scraping by Individuals and Private Organisations.
17. See the AG's Opinion in GC & Others (ECLI:EU:C:2019:14), recital 55.
18. See the AG's Opinion in GC & Others (ECLI:EU:C:2019:14), recitals 55 and 56, as well as recital 48 of the present judgment (ECLI:EU:C:2019:773).

This assessment by the operator should only take place after a removal request has been submitted by a data subject. In addition, the CJEU has ruled that there is an interest in search engines for the general public having access to information.[19] This can contribute to the right to information set out in Article 11 of the Charter of Fundamental Rights of the European Union.

Collecting data to build a training set for generative AI shows similarities to collecting data to provide a search engine. In both cases, as much data as possible is retrieved in an untargeted manner. Furthermore, in both cases, a bit of analysis will also take place on the data. On the other hand, there are also important differences:

- **Storage of data**. For a search engine, it is important to analyse data from a web page so that it is properly indexed. This means that the search engine does not have to store the source data completely after this processing, while storing data is the goal when building a training data set for generative AI. The differences between search engines and data collection for generative AI lie in the way data is stored and further used. A search engine provider has structured all data and made it searchable. This allows data to be easily viewed and deleted. With generative AI, data sets are curated, so that certain unwanted data and personal data can be removed. However, it can happen that pieces of training data that are left behind during the training of the model end up in the model, which can be reproduced verbatim. However, specific pieces of data cannot yet be 'untrained' from a trained model. In the future, the technique of 'machine unlearning' may offer a solution, but at the time of writing this technique does not yet offer a theoretical guarantee that the special categories of personal data will be removed from a trained model.
- **Purpose and frequency**. A search engine provider will scrape anything that is technically accessible (and not rejected according to Robots.txt). By definition, search engines try to provide the best and most up-to-date search results possible. This means that they constantly search the internet and, in principle, consult all available information. For scraping in the context of generative AI, data is scraped from the internet at some point and a training set is determined on which a model is then trained for weeks or months. While iterative, this process is still based on a snapshot of the scraped data. Rules can also be used to exclude certain websites directly from the data collection criteria. With scraping in the context of generative AI, it is not a goal in itself to return all websites in search results. Low quality or unwanted data (think of extremist forums, or health forums where a lot of special categories of personal data can be found) will have to be excluded in order to create a well-functioning model.

19. See ECLI:EU:C:2019:773, para. 53.

The AP therefore finds that search engines and scraping for generative AI show significant similarities. The snapshot of the processing of special categories of personal data in scraping may have a lower impact on data subjects than the continuous processing thereof by search engines. However, search engines' ability to comply with a removal request is an important difference to the extent to which trained generative AI models can comply with removal requests. The 'machine unlearning' technique may offer a solution for trained models in the future. The AP sees similarities and differences between scraping for generative AI and search engines when it comes to the right to information. Indeed, some generative AI applications could contribute to the provision of information to the wider public. But due to the wide availability of foundation models, it can also be stated that the function of these models is not only to provide information to a wider audience.

The AP therefore does not rule out the possibility that the CJEU's assessment of whether there is an exception for the processing of special categories of personal data may also apply to trained generative AI models in the future. However, the general nature of foundation models represents a significant difference in disclosures to a wider audience. Therefore, the aforementioned statement cannot be applied directly to scraping for training generative AI models. This means that, prior to scraping data for generative AI, controllers will have to assess whether they can rely on an exception for the processing of special categories of personal data.

**Preconditions for data curation**

*2. Training data must be lawfully and carefully curated and stripped of (unwanted) personal data as much as possible*

Relevant for: developers and fine-tuners

Typically, data sets are cleaned up before they are used to train or fine-tune a generative AI model. This is also known as 'data curation'. During data curation unwanted data, such as offensive language or (unwanted) personal data, are extracted from data sets. In line with the previous requirement, it is important that data sets curated by a controller be lawfully collected. This obligation applies both to developers of generative AI models and to fine-tuners of those models, who train or fine-tune their models on a curated data set. Data curation is an important step for the protection of personal data because it is the last step where it can be explicitly checked which personal data are included in the data set used to train the model. As soon as the personal data are implicitly trained into the model, the control and execution of additional rights is of a more complex nature. The curation of data sets can even be considered as necessary under the requirements for the legitimate interest asessment. This applies when certain personal data are not necessary for the development of a generative AI model. Curating data sets on certain (unwanted) personal data contributes to the principle of data minimisation.[20]

Since data curation qualifies as processing under the GDPR[21], the controller must have a legal basis, as described in Article 6 of the GDPR. If the same controller carries out both data collection and data curation, both processing operations may be carried out under the same legal basis, provided that the transparency requirements of the GDPR are met.[22] However, the controllers will have to assess this on a case-by-case basis and prior to processing. This obligation also applies to developers of generative AI models, as well as to fine-tuners of these models, in case they (further) train their models on a curated data set.

Finally, data curation usually takes place automatically, which raises the question whether curation offers a 100% guarantee that unwanted personal data will no longer return in the cleaned data set. However, there are also exceptional cases where generative AI models have reproduced personal data despite data sets having been curated.[23] This is also known as 'regurgitation'. Such regurgitation is then the result of personal data having been trained into the model, because the automatic curation of those personal data has not (completely)

20. This is the second step of the test that controllers must carry out when they rely on the legitimate interest basis, also known as the 'necessity test'. See Guidelines 1/2024 on processing of personal data based on Article 6(1)(f) GDPR, p. 12.

21. The processing or deletion of data is regarded as processing under Article 4(2) of the GDPR.

22. In line with the principle of purpose limitation and the transparency obligations towards data subjects, data subjects should be aware of the purposes for such processing prior to data collection and data curation. For more information, see Article 5(1)(b) GDPR and Article 12-14 GDPR.

23. See for example: Carlini et al. (2021), 'Extracting Training Data from Large Language Models,' <https://www.usenix.org/conference/usenixsecurity21/presentation/carlini-extracting>. See also: Lukas et al. (2023), 'Analyzing Leakage of Personally Identifiable Information in Language Models,' <https://arxiv.org/abs/2302.00539>.

succeeded. It is, therefore, important that controllers use state-of-the-art techniques to prevent such regurgitation.

Regurgitation can possibly be suppressed by doing an extra training on specific examples after the training on raw data, which people have indicated to be unwanted information.[24] Teaching the model that it is not appropriate to produce personal data will make it less likely to cough up that personal data. If a model can be further fine-tuned, it may be possible to undo this. In addition to suppression in the model, it is possible to check the output (filter) for personal data. If detected, a default response could be returned or a new response generated instead of the generated output. This is a solution that works at system level but does not directly modify the model. As a result, the personal data are still in the model. This actually works but does not replace the exercise of rights by data subjects.

If a generative AI model has **verifiably** been anonymised and still produces personal data, it can be assumed that such reproduction of personal data constitutes a 'hallucination'. A generative AI model produces plausible content based on common patterns in the data. However, the model has no actual knowledge of the world and may, therefore, produce plausible sounding falsehoods. These are called 'hallucinations'. A hallucination is not a result of training a personal data, but an inherent failure of the technology. The difference between hallucinations (not personal data) and regurgitation (personal data) by a generative AI model is important for the fulfilment of data subjects' rights. Personal data that are not included in the data set or are trained in the model cannot be rectified or deleted. The following precondition goes further into the rights of data subjects, including in the event of regurgitation.

**Preconditions for training a foundation model & fine-tuning a model**

*3. Controllers have put in place a system to facilitate data subject rights*

Relevant for: developers, fine-tuners and deployers

Based on the above, an anonymised model does not involve the processing of personal data. This precondition is, therefore, about the situation where a model has not verifiably been anonymised. Data subjects have various rights with regard to their personal data under the GDPR, including the right of access, the right to rectification and the right to erasure.[25] The obligation to facilitate those rights lies with the controller receiving the request. Data subjects can exercise their rights at different steps in the generative AI chain. Providers and deployers of generative AI models must establish a system to comply with the rights of data subjects when they receive requests from data subjects who want to exercise their rights.

The patterns that generative AI models have learnt are embedded in numbers, also known as 'weights'. The information stored herein is no longer explicitly represented but is implicitly part of the collection of weights. Any personal data contained therein are, therefore, not immediately identifiable. However, the parameters form a statistical relationship between the training data, making it possible to extract personal data from the model by, for example, questioning the model or by directly extracting relationships between the data in the model. Thus, the information contained in the model may relate to a natural person, despite the fact that the model is technically organised or encoded. As a result, a relationship with a natural person is not immediately clear.[26] This means that generative AI models can indeed contain personal data.

The AP is aware of the difficulties in facilitating data subjects' rights with respect to already trained models that have not verifiably been anonymised. In order to comply with the right of access, a controller could provide access (using a search function) to the data set used for training or fine-tuning a model. Data subjects can also exercise their right to erasure and rectification regarding these data sets. By rapidly phasing out old models, controllers can comply with the right to erasure and rectification of an already trained or fine-tuned model. The new model must then be trained or fine-tuned on a data set without the personal data. In the future, certain data may be 'untrained' from these models using the 'machine unlearning' technique. However, at the time of writing, this technique does not yet guarantee that the personal data will be unlearnt from the model. Retraining the model is, therefore, currently the only solution.

24. This is also known as 'reinforcement learning from human feedback'.

25. See, inter alia, Articles 15, 16 and 17 of the GDPR.

26. EDPB Opinion 28/2024 on certain data protection aspects related to the processing of personal data in the context of AI models, para. 37.

In practice, providers often make foundation models available to fine-tuners or deployers without training sets. Any personal data trained into that foundation model are not explicitly represented. Without access to the training set, those parties have virtually no means of identifying data subjects.[27] In other words, the personal data have been processed in the generative AI model of the provider, but identification by the fine-tuner or deployer is not a purpose in itself. In that case, the fine-tuner or deployer does not have to comply with the rights of data subjects, according to Article 11(2) of the GDPR. Only when a data subject themselves demonstrates to the fine-tuner or deployer that their personal data are in the model does the obligation to facilitate the rights of data subjects come back to life.[28]

In that case, access to the training data will not be possible, because fine-tuners and deployers do not have those training data. For erasure and rectification, retraining the model seems to be the only solution at the moment. However, since they are not the ones training the foundation model, this does not offer a solution. In the future, 'machine unlearning' may offer a solution to this problem. These difficulties in upholding the rights of data subjects can be avoided. Providers and fine-tuners can contractually agree that providers provide fine-tuners or deployers with access to recent training sets as and when requested. It can also be contractually agreed that providers provide a new model without the concerning personal data, if a fine-tuner or deployer receives a rectification or erasure request. Furthermore, when making foundation models available, providers could provide the most recently used training set to fine-tuners or deployers. In that case, the right of access can be facilitated by providing access to those training data. Without such contractual agreements, it seems that a fine-tuner or deployer will not be able to fulfil the rights of data subjects. As a result, the provision of foundation models cannot be considered lawful.

*4. The purposes for training personal data in generative AI models and for processing personal data in generative AI applications must be determined and communicated to data subjects in advance*
Relevant for: developers, fine-tuners and deployers

According to the purpose limitation principle, personal data may only be processed if controllers establish and communicate specific processing purposes to data subjects prior to such processing.[29] These purposes must be lawful and formulated specifically and explicitly enough so that data subjects know what their personal data are processed for. If the inclusion of personal data in training a model is necessary, controllers must clearly communicate this to data subjects.

Generative AI models can be used for a wide range of tasks, with fine-tuning often taking place at a later stage. Developers and fine-tuners of these generative AI models must formulate processing purposes specifically enough before they use personal data for training a foundation model. It is important that the processing purposes provide at least some context considering the use of a foundation model by, for example, describing the functionalities of the model or by describing whether the model has been developed for internal purposes or is intended for subsequent distribution or sale.[30] Fine-tuners can usually formulate goals more specifically when fine-tuning foundation models than developers when training foundation models.

When developers or fine-tuners of generative AI models collect personal data by scraping, the exception of Article 14(5)(b) GDPR may apply. It follows that, where the provision of information to data subjects proves impossible or would involve a disproportionate effort, that provision of information may be omitted. In that case, the controller must still provide information about the processing purposes on its website or in its application.

Deployers of generative AI applications must also comply with the purpose limitation principle. If the deployer of a generative AI application processes personal data, it must determine these purposes prior to the processing and communicate them to data subjects. The purposes depend on the type of generative AI application deployed. Deployers of generative AI applications will usually have a direct relationship with data subjects, which means that the exception in Article 14(5) GDPR does not apply.

27. See CJEU *Breyer v Bundesrepublik Deutschland,* ECLI:EU:C:2016:779, para. 42.
28. See also recital 57 of the GDPR.
29. See Article 5(1)(b) GDPR.
30. For more information, see: EDPB Opinion 28/2024 on certain data protection aspects related to the processing of personal data in the context of AI models, para. 64.

**Preconditions for the deployment of generative AI applications**

*5. Generative AI applications generate as little erroneous or unwanted personal data as possible*

Relevant for: deployers

Trained generative AI models do not work with an explicit knowledge model, but instead produce smooth text and accessible images or videos based on probability. Personal data may, therefore, be produced incorrectly, because the model makes an incorrect association. In addition, generative AI models may accidentally cough up personal data from the training set (also known as 'regurgitation'). As a result, outcomes in generative AI applications may be unwanted.

Under the GDPR, personal data must be accurate taking into account the purposes for which they are processed.[31] If personal data are incorrect, they must be erased or rectified, taking into account all reasonable possibilities.[32] The principle of accuracy should be seen in relation to the risks and consequences for data subjects when their personal data are processed.[33]

For deployers of generative AI applications, this means that they must take all reasonable measures to prevent the reproduction of erroneous or unwanted personal data. New technologies such as retrieval-augmented generation (RAG) and chains of thought (CoT) can offer a solution to reduce the reproduction of incorrect and unwanted personal data.[34] Through retrieval-augmented generation, the system generates content based on searched sources. With chains of thought, decision-making is based on several steps or sequences.

It is also important that the controller informs data subjects in a clear and comprehensible manner of the possibility of erroneous and unwanted outcomes in generative AI applications. As a result, the controller contributes to the increase in 'AI literacy' in society. People therefore understand that errors can occur in the (output of) generative AI applications. If a request for access shows that the personal data was not included in the training set, but output containing this personal data was produced, it can be assumed that this personal data was incorrectly produced (hallucinated). The controller must take as many mitigating measures as possible to avoid those erroneous or unwanted outputs. If the model nevertheless hallucinates personal data, the controller must be able to demonstrate that that hallucination is not the result of the personal data having been trained into the model.

31. EDPB Guidelines 4/2019 on Article 25 Data protection by design and by default, para. 77.
32. Article 5(1)(d) GDPR.
33. EDPB Guidelines 4/2019 on Article 25 Data protection by design and by default, para. 78.
34. Wei et al. (2022), 'Chain-of-thought prompting elicits reasoning in large language models,' <https://arxiv.org/abs/2201.11903>.

# 3. Conclusion: actions by the AP

In the coming period, the AP will make an effort to contribute to the desired 'Values at work' vision of the future, as detailed in the vision paper. We will do this on the one hand by paying extra attention to generative AI within our existing activities and on the other hand by starting a number of new activities that make a positive contribution to the responsible use of generative AI in our society.

As part of our regular work, the AP will contribute to clear and realistic working methods, including by actively writing standards and opinions, such as the December EDPB Opinion and the standards for high-risk systems in the AI Act. In addition, the AP is working on digital resilience, including by investing in the level of knowledge in society. We are, for example, providing guidance on AI literacy, publishing comprehensible information on our website and organizing seminars. Risk identification is also an integral part of the work of the AP. For example, via this overview of GDPR risks for generative AI and the biannual risk reports on algorithms and AI. Finally, the AP is available for prior consultation and is setting up a sandbox process for high-risk applications under the AI Act.

The AP will also take a number of additional steps to make responsible progress with generative AI. We will start by identifying the questions and challenges surrounding the responsible development and deployment of generative AI. To this end, we will organise a number of meetings and set up an online platform for questions and ideas about generative AI. The information we collect will be the basis for further periodic dialogue on the responsible development and deployment of generative AI.

We will also work with concrete instruments and tools that contribute to protecting fundamental values in the development and deployment of generative AI. These include EU directives, efforts to encourage the use of methods to anonymise or remove personal data from models, guidelines for AI literacy and principles for transparency. And together with other regulators in the digital domain, we are also committed to jointly explaining standards, so that we create as much clarity as possible and thus legal certainty for organisations that want to work with generative AI. We are giving positive examples and use cases a platform to show and inspire what responsible generative AI can look like in practice. And of course, the AP will continue to monitor illegalities that occur in the playing field of generative AI.

This is how, together, we will make the responsible use of generative AI possible.