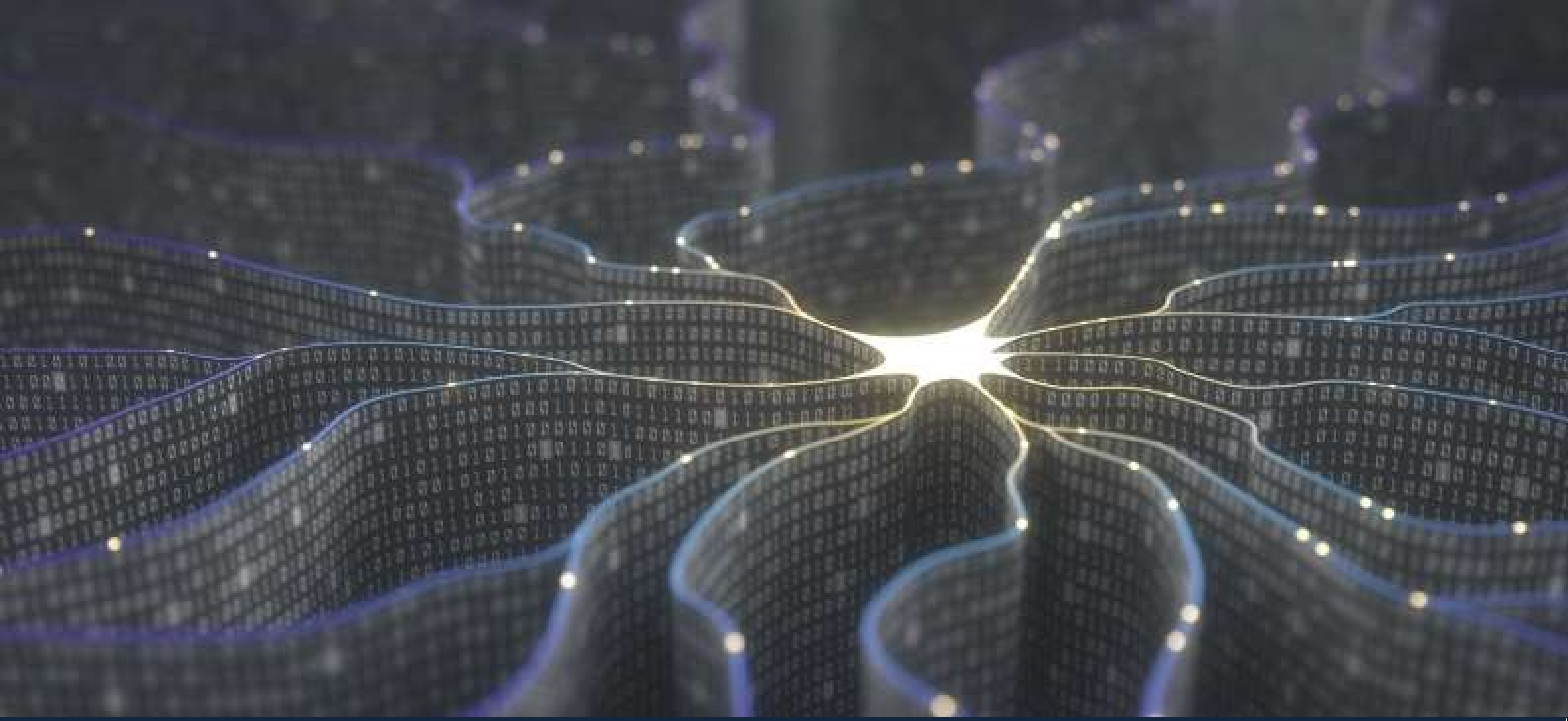


The use of artificial intelligence by public authorities

Before you get started

October 2023

DATATILSYNET

Content

Foreword

Data protection rules are sometimes perceived as an obstacle to the development of effective, data-driven solutions based on new technologies such as artificial intelligence ("AI"). However, the development of robust, long-lasting AI solutions in the public sector requires that the enormous potential of data is harnessed while respecting the fundamental rights of citizens. This is precisely what data protection rules aim to ensure.

The Danish Data Protection Agency believes that innovation and data protection are not mutually exclusive. Compliance with data protection rules, including the key principles of data minimization and purpose limitation, is a prerequisite for a democratic and appropriate technological development of our society. The preface to the Data Protection Regulation emphasizes that one of the primary purposes of the rules is to create trust in the processing of personal data by authorities and companies.[1] Without citizens' trust that new technology is used responsibly and with respect for their rights, otherwise promising solutions to important societal challenges risk meeting resistance and not finding a foothold. By demonstrating compliance with data protection rules, you send a clear signal to citizens that their fundamental rights are protected and that they can trust the technological solutions you provide.

Data protection rules apply regardless of the chosen technology and must therefore also be complied with when personal data is processed using AI. When developing AI solutions, it is important to consider data protection already in the early stages of the project. It can be very costly and technically demanding to adapt or modify an AI solution, for example to take into account an issue of discrimination or lack of legal basis, when the solution is fully developed or close to it. Data protection rules should therefore always be handled as an integral part of the project process. This applies both before and during the development process and when using the solution.

The purpose of this guide is to enable authorities to make the initial data protection law considerations that are a prerequisite for initiating an AI project. The guide is primarily aimed at the employees responsible for the project and the employees who advise and guide on data protection in connection with such projects.

The guide is about authorities' development and use of AI solutions that primarily involve the processing of personal data about citizens and possibly incidentally about the authorities' employees.

Finally, the guide only concerns data protection rules and does not relate to, among other things, the rules in the EU's upcoming regulation on artificial intelligence. The guide also does not affect other legislation such as the Medical Devices Regulation, the Health Act, etc.

[1] Preamble recitals 6 and 7 of the GDPR.

1. What is artificial intelligence?

There is no precise and universally accepted definition of AI yet.[2] However, a number of international actors have developed their own definitions of AI. For example, in 2019, the OECD adopted a set of principles for artificial intelligence.[3] Here, an AI system is defined as follows:

> "*An AI system is a machine-based system that can, for a given set of human-defined ob- jectives, make predictions, recommendations, or decisions influencing real or virtual envi- ronments. AI systems are designed to operate with varying levels of autonomy*."

It is also expected that the upcoming EU Artificial Intelligence Regulation, which is still under negotiation, will include a definition of artificial intelligence. In the European Commission's proposal for the regulation from April 21, 2021[4] , an AI system is defined as follows (Article 3(1) of the draft):

> 'software developed using one or more of the techniques and approaches listed in Annex I that, for a given set of human-defined goals, can generate outputs such as content, predictions, recommendations or decisions that affect the environments they interact with'

The European Council, in its general approach of December 6, 2022 to the Commission proposal[5] , proposed this definition instead:

> "a system that is designed to operate with elements of autonomy and, based on data and inputs from machines and/or humans, derives how a given set of goals can be achieved using machine learning and/or logical and knowledge-based approaches, and produces system-generated output such as content (generative AI systems), predictions, recommendations or decisions that affect the environments with which the AI system interacts".

Most recently, on June 14, 2023, the European Parliament adopted its amendment to the regulation.[6] Here AI is defined as:

> "a machine-based system that is designed to operate with varying degrees of autonomy and that, with explicit or implicit goals, can generate outputs such as predictions, recommendations or decisions that affect the physical or virtual environments."

In simple terms, systems based on AI, such as machine learning, are systems that, by recognizing patterns and relationships in data sets, can derive conclusions and apply them in future analyses.

In the development phase, an AI system is trained using selected data sets ("training data") to identify certain patterns. The system is then able to identify the same patterns when it receives input in the form of new data during the operational phase. By analyzing this data, the system can

[2] A report prepared for the European Commission has examined different definitions of artificial intelligence across 55 different documents including national and international strategies and reports: AI WATCH. Defining Artificial In- telligence, Publications Office of the European Union: https://publications.jrc.ec.europa.eu/repository/handle/JRC118163

[3] OECD, Recommendation of the Council on Artificial Intelligence, C/MIN(2019)3/FINAL.

[4] Proposal for a Regulation of the European Parliament and of the Council laying down harmonized rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts, COM(2021)206 final: https://eur-lex.eu- ropa.eu/legal-content/EN/TXT/?uri=CELEX%3A52021PC0206

[5] Council of the European Union, Interinstitutional file 2021/0106(COD), Proposal for a Regulation of the European Parliament and of the Council laying down harmonized rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts - General approach (6 December 2022): https://data.consilium.europa.eu/doc/document/ST-15698-2022-INIT/en/pdf

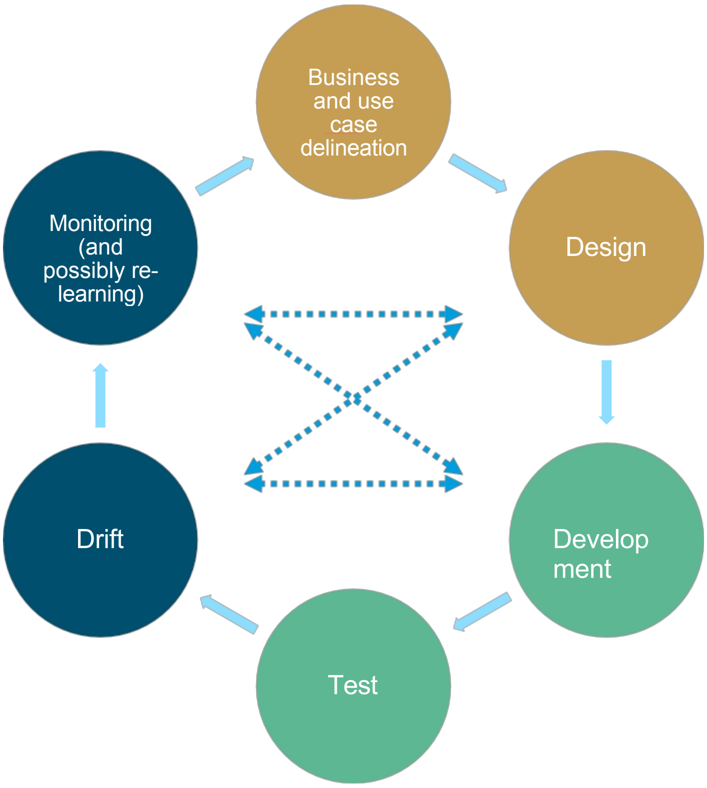
[6] Amendments adopted by the European Parliament on 14 June 2023 on the proposal for a regulation of the European Parliament and of the Council laying down harmonized rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts, amendment 165: https://www.europarl.europa.eu/doceo/document/TA-9-2023-

0236_DA.html

The system generates output in the form of content, predictions, recommendations or decisions based on the probability of a known pattern occurring in the new datasets.

In the context of data protection law, whether a system should be considered AI or not is relevant, but not decisive. Data protection rules are technology-neutral, so you must comply with the rules regardless of whether you process personal data using AI or using a traditional IT system. The reason why it is relevant to consider whether a system should be considered AI is that the development and use of AI can involve a particularly extensive processing of personal data with resulting risks to citizens that you must identify and manage.

2. The lifecycle of AI solutions


Business and use case delineation
Design
Develop ment
Test
Drift
Monitoring (and possibly re-learning)

Developing and operating an AI solution is typically an iterative process consisting of a number of phases that do not necessarily happen in a specific order. A (very) simplified description of the process can be seen above.

1) **Business and use case delineation**

If you as an authority want to develop an AI solution yourself, you will usually start by identifying a problem to solve or a hypothesis to investigate. This phase includes considerations on how you can achieve the identified purpose, including what types of personal data you need and how many. The Danish Data Protection Agency recommends that in this early phase you also involve professional competencies, e.g. health, social or similar, as well as legal competencies that can help ensure that the system is fit for purpose, effective, accurate and legal.

2) **Design**

A large part of the design phase of an AI solution is to provide adequate training data. This can be done by collecting new information or reusing existing information from internal and external data sources. You'll usually need to make a number of crucial legal judgments when selecting training data. For example, there will be questions about whether data can be used for the intended (new) purpose of developing an AI solution, whether you can obtain data from other authorities for this purpose, whether citizens must be informed that their data is being used for this purpose, etc. Already at this stage, you should also consider how to develop the solution using as little personal data as possible.

3) **Development and testing**

Developing the solution will often be done using the training data provided, but there are also some AI models that can be developed without training data. Regardless of the type of model, you will need to make a number of additional assessments and decisions. These may include the choice of

AI model, adjusting the model to ensure its statistical correctness and generalizability[7] as well as managing risk of bias and lack of transparency.

This is typically also the phase where the developed solution will be tested. For example, you will use the developed solution on a limited data set in a controlled environment to test whether the solution has flaws or errors.

4) Drift

Subsequently, the developed solution is put into use and becomes part of the authority's daily operations. This typically means that the solution is used to generate content to help with operations. It may also be that the solution generates predictions or recommendations and is used as decision support as part of the case processing. Finally, it may be that the solution generates and makes fully automated decisions for citizens.

The solution's output is based on the correlations and patterns that the solution has identified in the training data and that the solution can (possibly) find in the datasets that are introduced into the solution after it has gone live.

Is the AI model personal data in itself?

The Danish Data Protection Agency assumes that an AI model as a clear starting point does not in itself constitute personal data, but is only the result of the processing of personal data. This means that a statistical report will also not be considered personal data if the report only contains conclusions and aggregated data that are the results of the statistical analysis.

However, some machine learning models can be attacked in different ways (so-called *model inversion attacks* and *membership inference attacks*) that make it possible to re-identify the citizens whose information has been included in the model's training data. A successful attack that results in the re-identification of citizen information in the training data can be a personal data breach and must be handled accordingly.

The risk of a malicious actor re-identifying citizens by deliberately carrying out an attack to derive data that has been included in the training data does not, in the opinion of the Danish Data Protection Agency, mean that the model should be considered personal data in itself.

5) Monitoring and possibly re-learning

Once an AI solution is deployed, it must be continuously monitored to ensure that its output remains accurate. The obligation to ensure that the solution (continues to) process correct personal data and provides accurate predictions, recommendations, etc. follows from the principle of accuracy and the requirement of data protection by design and by default settings.[8]

In machine learning, a general distinction is made between static and dynamic models. A static model is developed and trained on selected data sets until it is deemed ready for use. During the operational phase, there will be a need for regular monitoring of input data to ensure accurate output, but the model does not change during use. However, as the model is not continuously updated, its predictions will gradually become less accurate as input data changes over time, for example due to demographic trends. The model must therefore be re-trained periodically to ensure that the model's output remains accurate.

[7] The ability of the model to handle new input data and generate correct predictions, recommendations, etc. in the same way as it did with the training data.

[8] Article 5(1)(d) and Article 25 of the GDPR.

A dynamic model, on the other hand, is continuously (re)trained on the new data it processes while in use. The model adapts itself continuously and takes into account any changes reflected in the input data. This requires more extensive monitoring to prevent the model from developing in an inappropriate direction. On the other hand, the model can continuously improve and adapt to changes in the underlying input data.

The development and operational phases of a static machine learning model will be more clearly separated, while in a dynamic model these phases will flow together. You need to pay special attention to this, as it has implications for data protection considerations before and during an AI project.

3. The scoping and design phases

When considering developing or procuring and deploying an AI solution, you should start by answering two basic questions:

1) What purpose(s) will the solution be used for?
2) What personal data will be processed through the solution?

Answering these questions is a fundamental prerequisite for compliance with data protection rules. Even in cases where the answer to one or both questions seems obvious, the Danish Data Protection Agency recommends that, as part of the scoping and design phase, you conduct a mapping that systematically answers these two questions. It is the Danish Data Protection Agency's experience that there often turn out to be several purposes for processing, or that more types of data will be processed than originally assessed.

If, based on this initial mapping, you assess that your AI solution can be legally developed and operated, you must also be aware of the other requirements that follow from the data protection rules. These include the requirement to ensure proportionality throughout the solution, to ensure data protection by design and by default settings, and to ensure the necessary security of processing. You must also comply with these requirements throughout the lifecycle of the AI solution.

3.1 Purpose

The data protection rules contain a general requirement that personal data may only be processed for an explicitly stated and legitimate purpose.

In the development phase of an AI solution, the purpose, including the purpose of processing personal data, is to develop one or more solutions. Typically, historical data is used that was originally collected for a purpose other than developing an AI solution, such as specific case processing.

In the opinion of the Danish Data Protection Agency, the development of an AI solution must be considered a purpose in itself in the context of the data protection rules. The processing of personal data for the purpose of developing new technological solutions inherently serves a different purpose than the processing of personal data as part of the authority's daily operations, e.g. as part of the municipality's case management or the region's healthcare initiatives for specific citizens. This also applies even if the long-term purpose of developing the solution is to use it in the authority's daily operations.[9]

Processing personal data in connection with the development and operation of an AI solution also involves different conditions and risks for citizens.

Citizens will rarely experience direct consequences of their data being used to develop an AI solution. However, any processing of personal data involves risks for the citizens whose data is involved. When developing AI solutions, there may be a risk of unnecessary data accumulation, as separate training datasets will often be generated based on the authority's existing registers etc. There may also be a risk that the authority does not provide the same equivalent level of processing security for the training data as is the case for production data.

Finally, the processing is unlikely to be within citizens' reasonable expectations of what authorities will use their data for.

[9] In this regard, see paragraphs 40-43 of the CJEU judgment of October 20, 2022 in case C-77/21, where the Court seems to predict that testing and error correction of an IT system constitutes a separate purpose from the fulfillment of subscription agreements from customers that the IT system originally supported.

Processing citizens' data in an AI solution as part of the authority's operations, on the other hand, will usually involve greater risks for the individual citizen. This may be the case if the solution's output is important in an administrative decision or a decision to initiate healthcare treatment. Therefore, greater requirements apply to this type of treatment.

3.2 Proportionality

Once you have determined the purpose or purposes of your processing of citizens' personal data for the development of an AI solution, you need to assess whether the processing will be proportionate - that is, *suitable, necessary and proportionate* - in relation to the purpose or purposes. This follows from the principle of data minimization.

At first glance, it may seem difficult to reconcile this principle with the development of AI solutions, which generally require the processing of large amounts of data, including personal data. However, it is important to keep in mind that the data minimization principle does not mean that you cannot use personal data at all. However, you are under an obligation to carefully consider how you can best achieve your purpose - the development of an AI solution - using only the data that is necessary.

In this connection, you should first and foremost compare the overall consideration for citizens with the considerations that speak in favor of developing and using the AI solution as part of the exercise of authority. In other words, you need to consider the benefits for both the authority and the citizens that may be associated with the use of the technology. For example, shorter case processing times, new treatment options in the healthcare sector and a more efficient use of available resources versus the risks to citizens' rights that the use of technology may entail.

When the AI solution is developed, citizens will typically not be directly affected by the processing of personal data that occurs in this context. When the AI solution is then put into operation, the solution, through its predictions, recommendations, decisions, etc. will have a greater impact on the individual citizen's social, economic, educational or other types of circumstances. You must therefore consider the proportionality of the processing of personal data in both the development and operation of the AI solution, which may involve different risks for citizens.

The proportionality assessment then requires you to consider how the AI solution can be developed, trained and operated using as little personal data as possible - and if possible without any personal data at all. As mentioned, the data protection rules do not contain an actual prohibition against processing personal data in connection with the development and testing of new technological solutions, but the starting point is that anonymized data should be used as far as possible.

In an AI context, there are several techniques that can be used to process less personal data. These include the use of synthetic data and federated learning.[10] When developing an AI solution, consider the use of such techniques already in the design phase. You should make efforts to ensure that you process as little data as possible when designing and developing the solution. There may be legitimate reasons to deviate from this principle, but you must describe why the principle is deviated from. The justification must describe why it is not possible to use synthetic or anonymized data. One reason could be that the construction of suitable synthetic test data is impossible or that without the use of personal data there is a risk that the future solution will subsequently generate incorrect output. On the other hand, any costs associated with developing e.g. synthetic data cannot in itself justify deviating from the starting point. If you find that this is not possible because personal data is necessary for the development of the AI solution, you must, as a clear starting point, only use pseudonymized data.

[10] The Norwegian Data Protection Authority has published a report on federated learning as part of their regulatory sandbox for AI: Finterai, final report: Machine learning without data sharing | Data Protection Authority

3.3 Use and reuse of datasets, including from external data sources

The development, training and operation of AI solutions typically requires the processing of large datasets. These can be the authority's own datasets, such as historical cases, own registers, etc. It can also be datasets from other authorities, such as BBR, CVR or patient records from other governments.

Whether you want to use your own data or data from other authorities, it is important to be aware of the purpose for which the data was originally collected. This is because data protection rules require that data cannot be reprocessed for a purpose that is incompatible with the original purpose.

The rule means that any data you collect cannot be freely reused, disclosed, etc. You can only reuse data or receive data from other authorities if your new purpose for processing is compatible with your original purpose. Likewise, you may only receive data from other authorities if your purpose for processing the data is not incompatible with the purpose for which the originating authority originally collected the data.

It is up to the transmitting authority to assess whether the data in question will be used for a compatible purpose. This is because the disclosure in itself constitutes processing of personal data, and even this processing must not be for an incompatible purpose. Therefore, if another authority or company requests data from you, you must assess the purpose for which the authority or company will use the data.

There are generally two options for authorities to further process data. Firstly, it can be done if the further processing is not incompatible with the original collection purpose. Secondly, it can be done if it is stipulated in EU or Danish legislation.

Incompatible with the original purpose

When you - or the authority that will disclose data to you - assess whether your (re)use of the identified datasets is compatible with the purpose for which the data was originally collected, you must take into account, among other things:

a) any connection between the purpose for which the data was collected and the purpose of your intended use
b) the context in which the personal data has been collected, in particular with regard to the relationship between you and citizens
c) the nature of the personal data, in particular whether it concerns special categories of data or data relating to criminal offenses
d) The possible consequences for citizens of your intended use
e) the presence of so-called necessary safeguards such as pseudonymization.

In general, in practice, there is a relatively broad framework for public authorities to process information for other purposes, as opposed to private actors, where the framework is narrower in practice. There will usually be nothing to prevent information from being passed on to other authorities that need the information in their case processing.

If you want to use external data or allow others to use your data, you also need to be aware of the issue of legal basis. An authority that wants to give access to its data to, for example, a company for the development of an AI solution, must have a legal basis for disclosing the data. At the same time, the authority must, to some extent, ensure that the recipient, e.g. the company to which the data is disclosed, has a legal basis for processing the data. If you as an authority become aware that it is unlikely that the recipient has a legal basis for processing the information, it will not be legal for you to disclose the information.

Examples from the Danish Data Protection Agency's practice include:

The Danish Data Protection Agency's practice (j.nr. 2006-321-0486)

The MFA asked the Oversight Board for an advance statement on the question of whether the Ministry, based on information in a register (received from the police) in which the names and social security numbers of evacuated persons from Lebanon were listed, could either confirm or deny whether a specific person was listed in the register to, among others, municipalities that wanted to check whether the person in question had committed social fraud.

The Danish Data Protection Agency ended up accepting that the municipalities' reuse of personal data to check for social fraud was compatible with the purpose of the collection, which was to evacuate the Danish citizens in question.

The Danish Data Protection Agency thus stated that the MFA could legally check or confirm whether a specific person was listed in the register as long as the requesting municipality could prove that it had the legal authority to carry out such a check of individuals. The MFA could also disclose the entire register to a requesting municipality if the municipality itself fulfilled the conditions for being able to compile and compare personal data for control purposes. This means that the receiving authority had to have a clear and unambiguous legal basis that provides the legal authority to carry out compilation or interconnection for control purposes. - and if the authority had previously informed the groups of persons affected by the control about the possibility of conducting a general control.

However, this does not mean that authorities are free to reuse their own or other authorities' data. There are limits that can also be found in the Danish Data Protection Agency's practice:

The Danish Data Protection Agency's practice (j.nr. 2008-632-0034)

The Danish Data Protection Agency requested a statement from the Danish Defense Personnel Service, as the agency had become aware through media coverage that the service had passed on personal data on 15,000 employees to the insurance company Topdanmark.

The Danish Defense Personnel Service stated that the service, as part of an agreement with Topdanmark on providing discounts on insurance premiums to employees in the Armed Forces, had temporarily passed on an address extract to Topdanmark for the purpose of sending offers to Armed Forces employees. The address extract included all employees under the authority of the Defense Command and contained names, job titles and addresses.

The Danish Data Protection Agency did not agree with the Armed Forces Personnel Service that the disclosure could take place within the framework of, among other things, the purpose limitation principle. In this connection, the Danish Data Protection Agency emphasized that the disclosed information had been collected and processed to administer an employment relationship, and that disclosure to a private company for marketing purposes could not be considered compatible with this purpose. In addition, the Danish Data Protection Agency emphasized that it could not be assumed to be clear to the employees of the Armed Forces that information provided in connection with an employment relationship could be disclosed to a private company for marketing purposes.

In the opinion of the Danish Data Protection Agency, authorities - subject to the administrative law principles of objectivity and equal treatment - have a wide discretion regarding the extent to which

The authority can reuse its own or other authorities' data or obtain data from other authorities as part of the exercise of authority.[11]

However, as an authority, you must pay special attention to cases where you want to develop an AI solution for the purpose of linking data for control purposes. Although - unlike in the past - it is no longer a prerequisite that the interconnection of data for control purposes must have a separate legal basis in a law, the data protection rules set a framework for the extent to which interconnection can take place.[12]

Example 1

A government authority is tasked with paying out a wide range of public benefits. The authority must also monitor and combat benefit mispayments and fraud.

It follows from the legislation under which the authority operates that the authority can compare information from its own registers and information obtained from other authorities for control purposes.

The authority now also wants to collect and compare information about citizens' electricity consumption for control purposes. The information can be obtained from Energinet, which is a public company responsible for the so-called Data Hub. The information available in the Data Hub comes from the electricity trading companies, which have collected the information for billing purposes and to ensure security of supply and quality and capacity in the electricity grid.

In the opinion of the Danish Data Protection Agency, the authority's desire to receive information on electricity consumption for the purpose of interconnection for control purposes is incompatible with the purpose for which the information was originally collected, as the authority does not have a clear legal basis for obtaining information from companies.

The authority must have a clear and unambiguous legal basis that allows it to perform the interconnection for control purposes, which is not the case in the authority's current legislation.

Example 2

Under the Social Services Act, a municipality is obliged to provide support for body-worn aids such as corsets, prostheses, orthopedic footwear, etc. to citizens who have permanent physical or mental disabilities.

Every year, the municipality receives many applications from citizens, and citizens experience long case processing times. In order to alleviate the long processing times, the municipality decides to develop an AI solution that can search for previous similar cases that can support case processing.

The municipality's processing of citizens' personal data as stated in the application takes place for the purpose of carrying out the municipality's official duties under the Service Act.

Since the development of an AI solution must be considered a purpose in itself, the municipality must assess whether the processing of citizens' data for the purpose of developing an AI solution

[11] However, it is unclear - and disputed in the legal literature - exactly to what extent the rules on purpose limitation set limits for authorities' (re)use of their own data and for data obtained from other authorities. See Niels Fenger, Forvaltningsloven med kommentarer (2013), p. 782ff.

[12] Articles 5 and 6(4) of the General Data Protection Regulation. See also section 2.3.2.3.3.4 in the general comments to the proposed Data Protection Act (L 68), FT 2017-18, and the Minister of Justice's answer to question no. 52 from the Danish Parliament's Legal Affairs Committee of February 9, 2018.

solution is compatible with the municipality's original purpose of the processing, which is to receive and process applications for body-worn assistive devices.

It is the Danish Data Protection Agency's assessment that the purposes in this case will be compatible. This is partly due to the coherence between the purposes, as the AI solution will be used to assist in the processing of the same type of cases for which the data was originally collected. Likewise, using the historical information to develop the solution has no direct consequences for the citizens who have already received a decision on assistive technology.

Established in EU or Danish law

You can also (re)use data for a new purpose if it follows from legislation. This can be both EU law and Danish law. In that case, you do not need to make an independent assessment of whether the new purpose of using the data is incompatible with the original purpose.

Example 3

A government authority is tasked with assisting the regions in coordinating the distribution of students in upper secondary education. The authority will do this by using an AI solution. The distribution of students must be based on the parents' income, among other things. The authority therefore obtains information about the students' parents' income and assets from the tax administration.

It appears from the legislation that the authority has the possibility of obtaining this information from the customs and tax administration. As the (re)use of information about the parents' financial circumstances follows from legislation, the authority does not have to specifically assess whether the processing is compatible with the original purpose that justified the processing of the information about the parents (tax assessment).[13]

Scientific or statistical purposes

If you assess that the development of your AI solution is for scientific or statistical purposes, the processing of the necessary data will not be incompatible with the original purpose. See more about when the development of AI can be considered to be for scientific or statistical purposes below in section 4.1.2. In addition, you should be aware that this principle of compatibility does not necessarily apply to the operational phase, as operations can generally no longer be considered to be for statistical or scientific purposes.

Example 4

One authority wants to gain more knowledge about the income level of 21-35-year-olds over the last three decades to analyze the correlation between education and income level and possibly predict future patterns. In this connection, the authority wants to obtain, among other things, employment information from the job centers in the municipalities. Since the study requires an analysis of large amounts of data, the authority chooses to develop an AI solution that will perform the statistical analysis and make predictions of future patterns.

[13] The Danish Data Protection Agency notes that the example is fictitious. The Danish Data Protection Agency is not aware of such an AI solution being used in practice.

The information was originally collected by the job centers for use in citizens' specific cases in the employment area and was thus collected for a different purpose than for statistical purposes, which is what the authority now wants to use the information for.

As the processing of the data is done for statistical purposes, this further processing shall not be considered incompatible with the original purpose.

Whether you reuse data yourself for scientific or statistical purposes or disclose data to other authorities for the same purpose, you must remain aware of the requirement that data must be processed for a legitimate and specific purpose.

The Danish Data Protection Agency's practice (j.nr. 2022-32-2939)

The Danish Data Protection Agency received a number of inquiries from citizens who were unhappy that the Danish National Police had passed on information that they had received a speeding ticket to Aalborg University for use in a specific research project.

The research project was about preventing speeding in traffic and was designed to show whether drivers get fewer speeding tickets if, after receiving a speeding ticket, they undergo online learning about road safety.

One of the inquiries to the Danish Data Protection Agency was from a citizen who had objected to the proposed fine and was awaiting the courts' processing of the case.

The Danish Data Protection Agency found that the Danish National Police could generally disclose information about motorists' traffic violations to Aalborg University for use in the research project pursuant to section 10(1) of the Data Protection Act.

However, the Danish Data Protection Agency also found that the Danish National Police's disclosure of the information in the specific case was in violation of the principles of purpose limitation and data minimization, as the Danish National Police's disclosure was not for an objective and relevant purpose.

In this connection, the Danish Data Protection Agency emphasized that at the time of the disclosure by the Danish National Police, it must still be considered to have had the presumption against it that the citizen in question had violated the Danish Road Traffic Act until the case had been decided by the courts, and that the research project's target group, in the opinion of the Danish Data Protection Agency, was persons who have violated the Road Traffic Act.

As part of the development and operation of AI solutions, additional purposes for processing personal data may arise as a result of the results delivered by the solution. If you as an authority wish to pursue these purposes, you must make new assessments of whether these new purposes are incompatible and whether you have the legal basis to pursue these purposes.

Example 5

A municipality decides to develop an AI solution to improve its handling of access requests in terms of response time, quality and consistency. The solution must be able to efficiently search files and documents and identify information to be anonymized and support case management.

The AI solution is trained using data derived from historical access cases.

This means that the municipality processes personal data that was originally collected for one purpose (processing access requests) for a new purpose (developing an AI solution).

The municipality must therefore assess whether the new purpose is compatible with the original purpose. In the opinion of the Danish Data Protection Agency, there is a natural connection between the processing of personal data as part of case processing of access requests and the development of a tool to support the same case processing. In addition, further processing is carried out by the same authority that originally collected the information.

The new purpose - the development of the AI solution - can therefore be considered compatible with the original purpose.

You should also be aware that processing data for purposes other than the original ones means that you must inform citizens about the new purpose(s). See sections 4.2 and 5.4 for more information on this.

3.4 General about treatment basis

It is a fundamental prerequisite for you to legally process personal data for one or more purposes that you have a so-called basis for processing for each purpose. Initially, you should therefore conduct an overall assessment of the entire lifecycle of your future AI solution to ensure that you have identified the necessary processing basis(s) and can thereby lawfully develop the AI solution and subsequently deploy it.

When assessing possible processing bases in an AI context, you should distinguish between development and training and the subsequent operation of the solution. This is because development and training in a data protection context should be considered a separate purpose from the subsequent operation.

In the development phase, the aim of the treatment is to develop one or more AI solutions. In the operational phase, the AI solutions are used to solve one or more specific tasks in practice. In this case, the purpose of the processing is more closely related to the task to be solved. For certain types of AI solutions, there may also be a post-learning phase, where the AI solution is further developed and improved while it is in operation. Here, the solution is continuously developed based on information gathered as part of the operational phase.

Your processing of personal data in these different phases cannot necessarily be based on the same processing basis. This is because, among other things, the specific risks to citizens' rights can vary greatly between the different stages of processing.

This would be the case, for example, if you as an authority purchase an AI solution from a supplier. The supplier has probably processed personal data based on its legitimate interest in developing and training the solution as a product. When you deploy the solution, however, you will be processing personal data for one or more specific purposes for which you have purchased the solution. You must therefore identify which processing basis is relevant for this purpose.

This may also be the case if you have developed a solution for scientific purposes on the basis of section 10 of the Data Protection Act and you subsequently want to put the solution into operation, for example, for patient treatment.

If you process so-called special categories of data or data relating to criminal offenses using the AI solution, you must be aware of a number of additional conditions and requirements for the basis for processing. You can read more about processing special categories of data below in section 3.5.

If you want to use the AI solution to make automated decisions, you should also be aware that there is a general prohibition against such decisions. Using automated decisions requires that you can identify an exception to this prohibition. You can read more about this below in section 3.6.

You must always know the basis for your processing of personal data in an AI solution *before* processing the data. You must be able to document your choice of processing basis and inform citizens about it. This also means that you cannot subsequently change processing basis and, for example, switch from processing the data based on consent to another legal basis.[14]

3.5 Special categories of personal data

The processing of certain personal data is associated with a particularly high risk for the citizen. This applies to information about:

- race or ethnic origin,
- political, religious or philosophical beliefs,
- union affiliation,
- Genetic data,
- biometric data for the purpose of uniquely identifying a person,
- health or
- sexual relationships or sexual orientation.[15]

In view of the high risk to citizens' rights associated with the processing of the special categories of personal data, the data protection rules contain a general prohibition on the processing of these categories of data. Exceptions to this prohibition can only be made in the cases set out in Article 9(2).

Data may also fall into this category even if it can only be inferred from the context in which it is held. Whether data should be considered as part of the special categories of personal data - and thus subject to the prohibition in Article 9 of the Regulation - depends on whether it is possible with a high degree of certainty to reveal special categories of data about an individual or whether the intention is to derive such special categories of data, for example, by using profiling.[16]

When using AI, pay special attention to whether you are processing special categories of personal data. In many cases, AI solutions make it possible to derive special categories of data by juxtaposing a range of information about citizens and drawing conclusions about a person's physical or mental health, political persuasion or sexual orientation, for example.

In the context of AI, this means that you may be processing special categories of personal data if you:

- can (with a high degree of certainty) infer or is intended to infer specific categories of information about citizens, or
- aims to treat citizens differently based on the inferred special categories of data.

Example 6

One municipality has developed an AI solution that analyzes GPS and other driving data from home care vehicles in order to make more efficient use of the vehicles available to the municipality.

As part of its analysis, the AI solution includes addresses that the home care service has visited. From this information, it can deduce that the citizen at the address may be receiving help from the home care service.

In this situation, the AI solution does not process health information about the citizens in question. This is because it is not possible with a sufficient degree of certainty

[14] European Data Protection Board Guidelines No. 5/2020 on consent, p. 27, point 123.

[15] Article 9(1) of the General Data Protection Regulation.

[16] See also the Danish Data Protection Agency's decision in case no. 2021-31-5478 (Radius Elnet A/S)

to deduce whether - and if so, what kind of help the citizens in question receive from home care. In addition, the solution does not aim to treat citizens differently on the basis of the analysis performed by the solution. The results of the analysis will only be used to organize the use of home care vehicles, and no changes are intended in the content of the service the citizens receive.

As a public authority, you can, among other things, process special categories of personal data when it is necessary for the exercise of your public authority.

This can happen if the processing is necessary for the establishment, exercise or defense of legal claims. This is the case, for example, when you need to assess whether the citizen is entitled to a service or whether an application should be granted. Processing can be done in order to establish your, the citizen's or a third party's legal claim.[17]

Processing of special categories of data may also take place if the processing is necessary for reasons of substantial public interest,[18] or if the processing is necessary for scientific or statistical purposes.[19] In both cases, an additional legal basis is required where the processing in question is provided for by law. Thus, it is not a specific requirement that the additional legal basis contains an explicit rule on the processing in question.

The requirements for the clarity of the additional legal basis depend on the intrusiveness of the processing in question. The processing of special categories of data is associated with a particularly high risk for the citizen and therefore the clarity of the additional legal basis is inherently more demanding.

Data relating to criminal offenses is not among the special categories of data found in Article 9 of the Regulation. However, there are special rules for when you can process this type of data. As an authority, you can process data relating to criminal offenses if it is necessary for you to carry out your tasks as an authority.[20]

3.6 Profiling and automated decisions

3.6.1 What is profiling?

The term "profiling" covers cases where collected data is used to create profiles of an individual to predict, for example, behavior or future needs. Profiling is explicitly defined in the GDPR as "any form of automated processing intended to evaluate certain personal aspects relating to a natural person, in particular to analyze or predict factors concerning that natural person's performance at work, economic situation, health, personal preferences, interests, reliability, behavior, geographical location or movements."[21]

AI solutions based on machine learning are highly effective for profiling. The solutions are often specifically trained to find certain patterns in data sets. When the solution receives input data related to a specific citizen, the model can generate the statistical probability that the citizen will, for example, exhibit a certain behavior or develop a certain disease. Such AI solutions are used, among other things, as decision support for caseworkers in public administration and for imaging diagnostics in the healthcare sector.

[17] Article 9(2)(f) of the General Data Protection Regulation.

[18] Article 9(2)(g) of the General Data Protection Regulation.

[19] Article 9(2)(j) of the General Data Protection Regulation.

[20] Section 8(1) of the Data Protection Act.

[21] Article 4(4) of the General Data Protection Regulation.

3.6.2 What are automated decisions?

Profiling can - like other forms of personal data processing - form the basis for decisions or decisions regarding the citizen. This can be in the form of decision support, where the solution generates a proposal or recommendation for a decision or action for the citizen. However, it can also be in the form of automated decisions, where the solution also makes the decision for the citizen that the solution considers to be the most appropriate in the context in question.

The data protection rules contain a general prohibition on decisions "based solely on automated processing, including profiling, which produces legal effects or similarly significantly affects the data subject".[22] Thus, the prohibition does not cover profiling, but only automated decisions (which may be based on profiling). For public authorities, the prohibition can only be waived if the law states that the authority can make such automated decisions. Consent from the citizen will generally not be able to derogate from this prohibition. See section 5.2.2 for more information on citizens' ability to give consent to authorities.

One particular issue to be aware of if you want to develop and use an AI solution for decision support is the risk of so-called *automation bias*. These are cases where, for example, caseworkers attach more importance to the system's assessment of a case than their own assessment, which leads to the system de facto deciding the case. If the AI solution is to continue to be considered decision support, a human must have independently assessed the information on which the decision is based, and they must also have the necessary authority to override the system's recommendations.[23]

It also depends on a concrete assessment when a decision "similarly significantly affects the person concerned." This would be the case, for example, with automatic rejections of bank loans or automatic screening of applicants for a job.[24] Decisions to extract citizens for control purposes, for example in the tax or social field, could in some cases also affect the individual sufficiently to be decisions covered by Article 22.[25]

As a clear starting point, a clear and precise legal basis is required if an authority wants to make automated decisions or other intrusive decisions towards citizens. The law must state that the authority can make automated decisions, but it is not a requirement that the law specifically regulates that the decision must be made without human involvement.

[22] Article 22 of the General Data Protection Regulation.

[23] Article 29 Working Party Guidelines on automated individual decision-making and profiling under Regulation 2016/679, WP 251, p. 21f.

[24] Preamble Recital 71 of the GDPR.

[25] See Naomi Lindtvedt, Kravet til klar lovhjemmel for forvaltningens innhenting av kontrolloplysninger og bruk av profilering, section 3.6.

4. The development and training phase

As mentioned, the lifecycle of an AI solution consists of a number of different phases. This is relevant, for example, when you need to identify a legal basis for your processing of personal data in the solution in question.

Processing of personal data for the purpose of developing and training the solution is, in the opinion of the Danish Data Protection Agency, considered a purpose in itself and separate from the purpose(s) subsequently pursued in operating the solution.

4.1 Relevant treatment bases

4.1.1 Task in the public interest or exercise of public authority

Public authorities' processing of personal data for the purpose of developing and training an AI solution will, as a general rule, take place for the performance of a task in the public interest or as part of the exercise of official authority.

If you want to develop an AI solution as part of your exercise of authority, you must therefore first and foremost be clear about which legal rules, executive orders or other administrative regulations oblige or authorize you to exercise the given authority. In other words, you should focus less on data protection rules in this context and examine other relevant legislation that obliges or authorizes you as an authority to perform a specific government task.

Section 222 of the Health Act about Statens Serum Institut

The provision states that the purpose of Statens Serum Institut is to prevent and combat infectious diseases, congenital disorders and biological threats. Furthermore, the provision states that Statens Serum Institut acts as a central laboratory with regard to diagnostic analyses, including reference functions. The Institute has a national role in relation to fulfilling the country's tasks pursuant to international obligations in relation to cross-border health risks. The Institute ensures the supply of vaccines for public vaccination programs and preparedness products through procurement, storage and distribution. The Institute prioritizes and organizes distribution to ensure supply and reduce the risk of wastage of vaccines and preparedness products. The Institute is part of the operational preparedness against infectious diseases and biological terrorism and the preparedness in the veterinary field. The Institute conducts scientific research and provides advice and assistance in areas related to the Institute's tasks.

The provision is an example of a general provision on the exercise of public authority, which also forms the legal basis for SSI to process the personal data necessary to perform the tasks mentioned.

Next, you need to assess whether the relevant legal basis is sufficiently clear and precise to form the basis for the development of an AI solution. The clarity of the relevant legal basis generally depends on how intrusive the processing of personal data carried out as part of the development of the AI solution will be for the citizen.

If the treatment is completely harmless, the requirements will not be particularly high. If, on the other hand, the treatment is intrusive, greater demands are placed on the clarity of the legal basis.

Section 67a of the Tax Control Act - a clear legal basis

The Customs and Tax Administration may process, including interconnect, the information held by the Customs and Tax Administration in order to develop IT systems necessary for the exercise of authority by the Customs and Tax Administration.

In addition, the customs and tax administration may collect and process all necessary information on economic and business conditions from other public authorities and publicly available sources, including comparing such information with the information held by the customs and tax administration, for the purpose of developing IT systems necessary for the exercise of authority by the customs and tax administration.

The comments to the provision state:

> "The proposed provision means that the Tax Administration will be able to combine this information for the purpose of developing IT systems that will be able to target, support and streamline the Tax Administration's exercise of authority when necessary. [...]
>
> The purpose of the proposed provision in section 67a of the Tax Control Act is to develop several different machine learning models and analysis models, etc. that will be able to recognize patterns and signs of e.g. fraud across data. However, it should be noted that the principles of proportionality and data minimization still apply when testing data. It is therefore assumed that the access to use tests will not be used to a greater extent than is necessary to ensure that the IT systems (scoring models) that could be developed under the proposal are operational and work as intended. Once the individual models have been developed, the models can be used under the provision of the Tax Control Act
> § Section 68 on record linkage."[26]

The provision is an example of legislation providing a clear framework for the intended processing of personal data. The provision thus specifies which authority will process personal data and which data. It also describes the nature of the processing (collection and combination of data) and its purpose (development of machine learning models).

In the opinion of the Danish Data Protection Agency, there are generally fewer requirements for the clarity of the legal basis for processing personal data for the development of an AI solution than for the actual operation of the solution. This is because the risks for citizens when developing the solution are lower than when the solution is subsequently used in case processing, where its data-based predictions and assessments can have real consequences for the individual citizen.

4.1.2 Research and statistics (section 10 of the Data Protection Act)

Section 10 of the Data Protection Act contains a special legal basis for the processing of special categories of personal data and information about criminal offenses for the purpose of carrying out statistical or scientific studies. Processing of personal data other than the special categories for statistical or scientific purposes will, where appropriate, be done with reference to the performance of a task in the public interest.[27]

[26] Extract from the special comments to section 3, no. 1, of Act no. 2612 of 28 December 2021 amending the Act on an Income Register, the Tax Reporting Act and the Tax Control Act (Register interconnection for the purpose of system development and exercise of authority and extended access to the eSkatData scheme etc.)

[27] Article 6(1)(e) of the GDPR on the performance of a task in the public interest.

The use of Section 10 of the Data Protection Act requires that the processing is (i) solely for the purpose of carrying out statistical or scientific studies (ii) of substantial importance to society, and (iii) that the processing is necessary for the performance of the studies.[28]

In some cases, the development of AI solutions can be of such a nature that the development of the solution can be said to be for statistical or scientific purposes. This depends on the specific purpose of the development of the solution in question.

Solely for the purpose of conducting statistical or scientific research

As an authority, you must assess whether the processing you wish to carry out in connection with the development and training of an AI solution will be for the purpose of carrying out a statistical or scientific study. In particular, your assessment may include whether the development project will:

- Provide new knowledge,
- apply the methodological standards applicable in the sector concerned,
- Comply with ethical standards,
- is done to share research results with (parts of) the outside world, e.g. for peer review and publication, and
- Contribute to the collective knowledge and well-being of society.

This is not a cumulative or exhaustive list of criteria, but merely a number of factors that can be included in the specific assessment that you must make. However, it is the opinion of the Danish Data Protection Agency that especially the first-mentioned factor - the provision of new knowledge - and the last-mentioned factor - the intention to contribute to society's collective knowledge and well-being - should be given decisive importance when assessing whether a processing activity can be considered to take place for statistical or scientific purposes.

Significant societal impact

In order to use section 10 of the Data Protection Act as a basis for processing, your development project must also have significant societal importance. Broadly speaking, the project will have "substantial social significance" when it clearly and positively benefits society. This will typically be the case when the project aims to:

- Improving the health and well-being of society,
- to improve the financial or economic situation of society as a whole,
- to contribute to knowledge in a given area, or
- To contribute to the development of more efficient products, services and processes for society.

In order for an investigation to be said to be of significant societal interest, it is not sufficient that the investigation only serves your (the authority's) own interests. This does not mean that you cannot (also) have an interest in the development of the solution in question. However, it is a condition that you can also point to something that benefits society more broadly, and the social importance of the development project must not be a subordinate or peripheral consideration in relation to your primary consideration that the project must address.

Example 7

A municipality wants to develop a decision support tool to support the municipality's assessment of notifications about children and young people not thriving. The solution must provide knowledge about which conditions can lead to or contribute to children and young people

[28] In the opinion of the Danish Data Protection Agency, the requirement of *necessity* in Section 10 of the Data Protection Act will in practice have a limited independent significance compared to the existing requirements that follow from Article 5(1)(b) and (c) of the Regulation on objectivity and proportionality.

are not thriving. The solution will support the social workers' case management when assessing notifications of poor well-being.

The solution must be developed based on historical information from previous notifications received by the municipality. This includes information about the reasons for reporting, which may include child abuse, crime, sexually abusive behavior, etc.

Developing the AI solution will partly mean that the municipality will gain new knowledge about the specific causes or conditions that lead to or have an impact on any dissatisfaction among children and young people. In addition, the development of the solution has a significant social significance, as the development of the solution must contribute to ensuring the well-being and well-being of children and young people across the municipality.

In the opinion of the Danish Data Protection Agency, the development of the solution meets the requirements for using section 10 of the Data Protection Act. This does not change the fact that the municipality is also taking care of a consideration for more efficient case processing of notifications of possible ill-being, as the primary consideration is to provide new knowledge about risk factors for ill-being and to ensure the welfare of children and young people.

However, operation of the solution still requires the municipality to identify an appropriate legal basis, which cannot be section 10 of the Data Protection Act.

Example 8

An authority is tasked with processing applications for disability-friendly refurbishment grants. Some applications are particularly time-consuming and difficult to process. This is partly because citizens often fail to send the necessary documents. The authority therefore wants to develop an AI solution that will reduce the processing time for the most time-consuming tasks. At the same time, the authority will gain more knowledge about which documents are typically missing in the cases in question and, based on this, can provide better guidance about the area on the authority's website.

The development of the AI solution must be done by training on historical time-consuming cases that contain a wide range of personal data about citizens, including health information. By taking the above-mentioned factors into account, the authority assesses that the development of the model is for scientific purposes.

However, the primary consideration that the authority wants to meet with the development of the AI solution is to contribute to faster case processing. The second consideration - obtaining increased knowledge about which documents are usually missing from citizens' applications - is peripheral to the authority's primary consideration, just as the new knowledge obtained by developing the solution is narrow and mainly relevant to the authority itself.

In the opinion of the Danish Data Protection Agency, the development of the solution cannot be said to be done for scientific or statistical purposes. This is because streamlining the authority's case processing cannot be said to provide new knowledge in the way that is characteristic of scientific or statistical research.

When you, as an authority, assess whether your development project has significant societal impact, you can start from how large a part of society the project will benefit and how much. In other words, the social impact can be considered in both breadth and depth.

If the project only affects a small group of people and only affects these people to a limited extent, you will not be able to apply section 10 of the Data Protection Act.

If the project only benefits a small group of people, but is of significant importance to those people, such as research into better diagnosis of a rare and serious disease, the project is likely to have significant societal impact. The condition may also be met where the project affects the whole of society but only modestly benefits the individuals concerned.

Tasks performed in the public sector, for example by municipalities or regions, will, by their very nature, often have significant societal importance, as the authority or institution in question plays a role that affects a large number of people and affects them significantly. At the same time, the basis and justification for the existence of public authorities etc. is that they perform a task of societal importance.

However, it is not self-evident that the development and training of AI solutions is of significant societal importance simply because the development is done with a view to use in the public sector. For example, research into intelligent handling of an authority's emails will rarely have such a significant impact on society in its breadth or depth that the processing could be considered research of significant societal importance. This does not mean that the development of such solutions cannot happen. As an authority, you simply need to identify a legal basis other than section 10 of the Data Protection Act.

Consequences of using section 10 of the Data Protection Act for development

When you have chosen to develop and train your AI solution on the basis of section 10 of the Data Protection Act, you cannot, during the development and training phase, act on the basis of the recommendations, statements, etc. that the solution generates for the citizens whose data is included in the training of the solution.[29] This applies regardless of whether you try to obtain the citizens' consent along the way, as such a switch between treatment bases is not possible.

If you subsequently put the AI solution into general operation, e.g. based on the processing basis of public authority with a clear supplementary legal basis, you can use the solution across all citizens. This means that you will also be able to make concrete decisions or measures, e.g. about social services or healthcare treatment, for those citizens whose information was included in the original training data.

The purpose limitation in section 10 of the Data Protection Act must therefore be understood to mean that you may not act on the recommendations, predictions, etc. that the solution generates while it is under development. For example, a hospital may not initiate patient treatment based on treatment suggestions that the solution generates during development. However, the hospital may initiate patient treatment based on suggestions generated by the AI solution during operation, regardless of the fact that the suggestions may concern the same people who were included in the training datasets.

Example 9

A municipality wants to gain more knowledge about its citizens' needs for rehabilitation after long-term illnesses, as it is experiencing increasing costs in this area. In order to provide this knowledge and at the same time give the individual citizen a greater benefit from their rehabilitation, the municipality develops an AI solution in collaboration with a supplier in the hope of reducing costs in this area. The solution organizes individual training courses for citizens based on a wide range of data such as their medical history. The AI solution is trained using information from previous cases.

The municipality processes data about the citizen, including health data, for the purpose of conducting a scientific study, as the reason for developing the solution is to provide new knowledge about rehabilitation and rehabilitation.

[29] This follows from section 10(2) of the Data Protection Act.

The municipality also wants to subsequently put the fully developed AI solution into operation as part of the municipality's regulatory task in the area.

The municipality can process personal data contained in the historical data on the basis of section 10 of the Data Protection Act, but cannot adjust individual training courses during the development of the solution based on the suggestions generated by the AI solution.

However, when the municipality subsequently puts the solution into operation, the municipality can customize individual training courses - also for the same citizens whose information was included in the training dataset. However, the prerequisite for the municipality to put the solution into operation is still that the municipality can identify a clear legal basis, for example in the Social Services Act, for the processing of personal data that takes place as part of the operation of the solution.

Between the development and deployment of an AI solution, the developed solution will often be tested. For example, you will apply the developed model to a limited group of people in a controlled environment to test whether the model has flaws or errors.

At this point, the processing of personal data in the AI solution will typically be for the purpose of generating predictions, recommendations or similar about a limited number of citizens included in the dataset used for testing. The solution will - in the same way as if the solution was in operation - generate an output, e.g. a prediction about the risk of a citizen needing long-term hospitalization. This output can be used to test whether the solution is error-free, for example by performing a human (in this case medical) assessment of the same citizen to verify the generated prediction. However, you cannot initiate the healthcare treatment for that citizen based on the output from the testing phase.

To the extent that the development of an AI solution can be considered to be wholly or partly for statistical or scientific purposes, the testing of such solutions may be considered part of the statistical or scientific purpose. This means that the processing of personal data in this phase can also be based on section 10 of the Danish Data Protection Act. This is because the purpose of such testing of the solution is typically to verify the accuracy of the output generated by the solution and to test the reliability of the solution. Furthermore, testing will often take place in a controlled environment and the solution cannot be considered to have been put into operation.

An inability to test AI solutions on the basis of section 10 of the Act would in many cases prevent the possibility of making a comprehensive assessment of whether the developed solution works as intended. A test of the solution may thus be essential to achieve the statistical or scientific purpose of the development and training of the AI solution.

However, it is important that you are aware that section 10 of the Data Protection Act cannot be extended to include the use of the AI solution after the testing phase as part of your daily operations.

Although it may seem obvious to develop and train an AI solution on the basis of Section 10 of the Data Protection Act, which sets out a number of more lenient conditions for the processing of personal data, it is the actual circumstances of the development project that determine whether the provision can be used. Thus, the provision cannot be used freely if the actual circumstances do not indicate that the processing of personal data is actually done solely for scientific or statistical purposes.

It should also be noted that operation of the solution - regardless of whether the development and training of the AI solution is based on section 10 of the Data Protection Act - requires a separate legal basis. Therefore, you should already in the design phase consider whether there is a relevant legal basis for your processing of personal data when the solution goes live. The Danish Data Protection Agency does not exclude that there may be cases where the operation of an AI solution is also for statistical or scientific purposes. In these cases, section 10 of the Data Protection Act can also be used for the operation of the solution in question, but it must be assumed that this will be the exception.

4.2 Duty of disclosure

When you collect data from citizens or others, you must, as a clear starting point, inform citizens that you are processing their data and why. This also applies when you collect data for the development of an AI solution.

You can read more about what information you need to give citizens etc. in section 3 on the duty of disclosure in the Danish Data Protection Agency's general guidance on data subject rights.[30]

If you are considering developing AI solutions using your own existing data, such as historical cases, be aware of your obligation to inform citizens. This is due to the requirement that citizens must be re-informed when the data is to be used for purposes other than what it was originally collected for. As described above in section 3.1, developing an AI solution is considered to be a new purpose in relation to, for example, case management in the administration.

The requirements for what information you need to provide to citizens depend on the situation. A distinction is made between the situation where you have collected the information from the citizen, for example if a citizen has submitted an application to you, and the situation where you have not collected the information from the citizen, for example if you have received information from another authority for use in your case processing.[31]

If you want to reuse data that you already have, make sure you inform citizens about the intended use before it happens.

Example 10

A municipality wants to develop an AI solution to search for previous application cases for body-worn assistive devices. The solution must be developed using old application cases. It is therefore information that the municipality has received from the citizens themselves.

The municipality has not previously informed citizens, for example in connection with the application, that the information would be used to develop an AI solution.

The municipality must therefore provide citizens with information that their data will be (re)used for this new purpose. This follows from Article 13 of the Regulation.

If you want to use data from other authorities, as a clear starting point, you must also make sure to inform citizens that you receive these datasets and what you will use them for.

Example 11

A municipality wants to develop an AI solution to support the referral of rehabilitation plans that the municipality is obligated to offer according to the Health Act. The solution will be developed using previous rehabilitation plans. The information originally comes from the regions, which have collected the information as part of patient treatment.

Neither the regions nor the municipality have previously informed citizens that the information will be used to develop an AI solution.

[30] The Danish Data Protection Agency's guidance on data subject rights, July 2018.

[31] If you have collected the data directly from citizens, the requirements can be found in Article 13 of the GDPR. If, on the other hand, you have received the data from elsewhere, the requirements can be found in Article 14.

The municipality must therefore provide citizens with information that their data will be (re)used for this purpose. This follows from Article 14 of the Regulation.[32]

However, you are not obliged to inform citizens that you are using the data for a new purpose if it is explicitly stated in the legislation that you are doing this processing of the data.

If you further process the data on the basis of an executive order issued pursuant to section 5(3) of the Data Protection Act, you do not have to inform citizens about the further processing. However, this does not apply if the new purpose is the compilation or combination of data for consolidation purposes.[33]

Financial support for recipients of elderly vouchers and lump sums for students - exemption from the obligation to disclose

According to section 6(1) of the Act on additional financial support for recipients of elderly checks and lump sums for education seekers who receive a scholarship as a supplement due to a disability or as single parents, the Danish Agency for Higher Education and Science discloses the social security numbers of education seekers to the Labor Market Supplementary Pension for use in the payment of lump sums.

The provision is an example of how an authority can fail to comply with the duty of disclosure in connection with further processing - disclosure of personal data to another authority - pursuant to Article 14(5)(c) of the General Data Protection Regulation. This is because the Danish Agency for Higher Education and Science is subject to an obligation to disclose the information in question to Arbejdsmarkedets Tillægspension, and this is clearly stated in the provision.

Disclosure of CPR data by local councils to private individuals - no exemption from the duty of disclosure

It follows from section 43(1) of the CPR Act that local councils may disclose protected names and addresses in the CPR to private individuals who have a legal interest in such information about a pre-identified person.

The provision is an example of an authority not being explicitly subject to an obligation to disclose information, but having the possibility to do so (cf. "*may disclose")*. Municipal councils will not be able to rely on this provision to avoid fulfilling their information obligation under Article 14(5)(c), as citizens cannot see from the provision that their information will be disclosed.

There are a number of other exceptions to the obligation to inform citizens about the processing of their data. This applies, for example, if the citizen is already aware of the information. Whether this exception can be used depends on what information you have given the citizens in connection with the collection of their data. If citizens were not informed at the time they provided the data about, among other things, the purposes for which the data would be used (training an AI solution for one or more purposes), you cannot use the exception because the citizens are not aware of all the necessary information.

[32] It is the opinion of the Danish Data Protection Agency that none of the exceptions in Article 14 of the Regulation or Section 22 of the Data Protection Act are relevant in this case.

[33] Section 23 of the Data Protection Act.

In addition, there may be a consideration for you as an authority or for the citizens themselves that means that the citizens should not be informed about a treatment.

In addition, if you have not collected the information from citizens, you are not obliged to inform them about the processing of their data if it is impossible or would involve a disproportionate effort. However, if this is the case, you must take other appropriate measures to protect citizens' rights, such as publishing the information on your website.

You can read more about the exceptions to the duty of disclosure in sections 3.4.3 and 3.5.3 in the Danish Data Protection Agency's guide on data subject rights.[34]

Example 12

A university researcher wants to develop an AI solution to identify possible risk factors for child and youth placement. The solution will be developed using information from an authority's register that contains information on children and young people placed in care from the last 50 years.

However, the researcher notes that some of the contact information in the register is outdated. As the researcher is unable to find the citizens' contact information in any other way, the researcher cannot inform the citizens about the processing of their data.

In the opinion of the Danish Data Protection Agency, the university can in this case refrain from fulfilling the duty of disclosure, as it is impossible or at least would require a disproportionate effort. As a compensatory measure, the researcher chooses to create a website where information about the research project can be found.[35]

You should also be aware that, as a data controller, you are not obliged to obtain more information about citizens solely in order to fulfill the duty of disclosure. In the example, this means that the university is not obliged to compare the information held by the university with information from the CPR register for the sole purpose of finding any contact information in order to fulfill the duty of disclosure, cf. Article 11 of the Data Protection Regulation.

[34] For more information, see the Danish Data Protection Agency's guidance on data subjects' rights, July 2018, as well as the Article 29 Working Party guidelines on transparency under Regulation 2016/679, WP260.

[35] The obligation to take appropriate measures in cases where it proves impossible or would involve a disproportionate effort to inform citizens about the processing of their personal data follows from Article 14(5)(b), 2nd indent of the GDPR.

5. Operation phase

When an AI solution is implemented in the authority's daily operations, it will typically be used to solve a specific task in practice. It may be an existing government task that the AI solution will help solve more efficiently. It may also be a completely new government task that the authority initially believes can best be solved using AI. AI solutions in the public sector could, for example, be used as decision support for caseworkers in the employment sector, make decisions in simple application cases, select companies and citizens for tax audits, interpret scan images in hospitals, predict patients' risk of complications following surgery and much more. Common to these purposes is that the use of such solutions in the public sector must support the exercise of authority.

5.1 Purpose

While the processing of personal data for the development of an AI solution is considered a separate purpose, the operation of an AI solution is more closely related to the performance of your official duties. Therefore, the operation of the AI solution will often not be considered a separate purpose, but simply support the existing government task. However, there may be cases where the operation of the solution is part of a new purpose. For example, this may be the case when the authority is to perform a new task that the authority has not previously performed, and which from the start must be done using an AI solution.

In any case, the use of AI solutions will typically actualize (often high) risks for the citizen, for example if the AI solution in operation aims to produce predictions, recommendations, etc. about the citizen, which the authority will act on. This will constitute an intervention in the citizen's specific circumstances, which may be greater or lesser depending on the circumstances.

Therefore, you will need to assess whether your processing basis, if any, is sufficiently clear and distinct to enable you to put the solution into operation. Furthermore, some legal bases will not be available to you for the operation of the AI solution, such as section 10 of the Data Protection Act, cf. section 4.1.2 above. In that case, you will need to identify another legal basis unless the AI solution is used for statistical or scientific purposes.

5.2 Relevant treatment bases

When the AI solution is fully developed and trained and you want to put the solution into operation, you must have a legal basis for the processing of personal data that will take place when operating the solution. If it is a dynamic model that continuously learns and develops based on the data it processes during the operational phase, a processing basis must be identified for both the purpose of development and the purpose of using the solution in operation. See more about this in section 5.3.

Personal data can be processed on the basis of one of several processing bases found in Article 6(1) of the GDPR. In practice, public authorities will typically process citizens' personal data because it is necessary to comply with a legal obligation,[36] or in order to perform a task in the public interest or in the exercise of official authority.[37] In these cases, the processing must always have a so-called supplementary legal basis in EU law or Danish law.

[36] Article 6(1)(c) of the General Data Protection Regulation.

[37] Article 6(1)(e) of the General Data Protection Regulation.

5.2.1 Legal obligation

A legal obligation does not necessarily have to be a law, but can also be rules issued pursuant to law, such as executive orders and other administrative regulations. Furthermore, a legal obligation must be sufficiently clear as to the processing of personal data that it requires. The legal obligation must therefore explicitly refer to the nature and subject matter of the processing, and you as an authority must not have undue discretion as to how to comply with the legal obligation.[38] If you as an authority base your processing of data in an AI solution on a legal obligation, it must be clear from the law that you are obliged to process the data in question and you may only process the data to the extent necessary to comply with the specific legal obligation.

Reporting earned income - legal obligation

Section 1 of the Tax Reporting Act stipulates that all employers must report their employees' income to the income register every month. Paragraphs 2-4 of the provision state the specific information in the form of income types that must be reported, including salary, bonuses, reimbursement for expenses incurred in connection with the work or used for courses and training, etc.

The provision is an example of a clear legal obligation to process, in this case process, personal data.

5.2.2 Task in the public interest or exercise of public authority

If the legal basis does not precisely state that you need to process specific personal data in order to develop or use one or more AI solution(s), it will generally be the performance of a task in the public interest or the exercise of official authority that constitutes your basis for processing. However, it must still be clear from the legislation that it is a task that you are obliged or entitled to perform. This could, for example, be the performance of tasks within the social area that you are required to perform in accordance with the provisions of the Service Act.

Support for assistive technology - mandated by the authorities

According to section 112 of the Danish Social Services Act, municipalities must in certain cases provide support for assistive devices for people with permanently reduced physical or mental functional capacity.

The provision is an example of a task that municipalities are obliged to perform.

For processing to be considered necessary for the performance of a task in the public interest, the task must be in the public interest and therefore of importance to a wider group of people. This would be the case, for example, with processing for historical, statistical or scientific purposes. It may be processing for the performance of a task in the public interest, regardless of whether a commercial purpose is being pursued at the same time, for example, and in general the concept must be understood broadly when processing is carried out by public authorities.

For personal data to be processed with reference to a legal obligation or task in the public interest or in the exercise of official authority, the processing must, as mentioned, be provided for in EU law or Danish law.

However, a specific law is not necessarily required for each processing activity. One law may be sufficient as a basis for several processing operations that are based on a legal obligation or that are necessary for the performance of a task carried out in the public interest or

[38] Ministry of Justice report no. 1565/2017, p. 117f. and p. 130.

as part of the exercise of public authority. The legal basis should also be clear and precise, and the application of the rules should be predictable for the citizens covered by the rules.[39]

The requirements for the clarity of the legal basis for your processing of personal data when operating an AI solution depend on how intrusive the processing is for citizens. In the opinion of the Danish Data Protection Agency, the legal basis must be assessed based on how direct and intrusive, for example, a decision or activity is for citizens. This applies regardless of whether the activity is burdensome or beneficial. The legal basis must be proportionate to the legitimate purpose pursued and the processing must not be more intrusive than necessary.

In the opinion of the Danish Data Protection Agency, there are different requirements for the clarity of the relevant legal basis for the development and operation of the solution. As mentioned above, the development of an AI solution does not, as a general rule, have direct consequences for citizens. On the other hand, an AI solution in operation is expected to generate predictions, recommendations, etc. that, for example, will support the authority's caseworkers in making decisions. An AI solution may also have to make automatic decisions for citizens. The consequences for citizens are therefore often greater when the AI solution is in operation, and therefore higher demands are placed on the clarity of the legal basis that forms the basis for using the solution in operation.

When assessing whether the legal basis you have identified is sufficiently clear, you should consider what data is being processed and about which individuals, including, for example, vulnerable citizens. In addition, you should consider whether the prediction, decision, etc. generated by the AI solution has an impact on the citizen's economic, educational, social, health or similar circumstances.[40] The impact can be positive or negative. Finally, you should consider whether the processing in question, including the fact that the processing is done using AI, is predictable and transparent to the citizen.

In general, the use of AI solutions cannot always be said to be intrusive for the citizen. However, the citizen-oriented use of such solutions by public authorities will typically have an impact on the citizen's life situation. Therefore, the processing of personal data by the AI solution will often be intrusive. Conversely, the use of AI solutions for more general government tasks that are not directly citizen-oriented will be considered less intrusive.

The Danish Data Protection Agency's practice (j.nr. 2022-212-3676)

In its assessment of municipalities' legal basis for using the Asta tool, the Danish Data Protection Agency stated that the requirements for the clarity of the necessary legal basis depend on how intrusive the processing in question is for the data subject. If the processing is completely harmless, the requirements will not be very high. If, on the other hand, the processing is intrusive, as was the case with the Asta tool, the requirements for the clarity of the legal basis are higher.

Asta was a tool designed to perform a machine analysis of a newly unemployed unemployment benefit recipient's risk of the person's contact with the job center becoming long-term. Among other things, the Asta tool estimated the duration of the unemployment benefit case and the contact process based on a wide range of information about the citizen.

Against this background, the Danish Data Protection Agency was of the opinion that there should be a legal basis in Danish law for the Asta tool to be used by municipalities, as it is known from
§ Section 8(2) of the Act on an Active Employment Initiative. This assessment was thus based on the description of the processing of personal data that would take place when using the Asta tool.

[39] Preamble recitals 41 and 45 of the GDPR.

[40] Preamble Recital 75 of the General Data Protection Regulation.

Requirements for the clarity of the legal basis

Stricter requirements

Direct interference in citizens' relationships

Not insignificant amount of special categories of disclosures

Includes (almost) exclusively vulnerable citizens, e.g. elderly, children, patients, etc.

Example of an example

The purpose of an AI solution is to perform a machine analysis of a newly unemployed unemployment benefit recipient's risk of the person's contact with the job center becoming long-term. In other words, the tool performs a statistically based analysis of the citizen in order to estimate the duration of the unemployment benefit case and the contact process based on a wide range of information about the citizen. This includes information from the citizen's most recent unemployment benefit cases, including information about the citizen's CV, previous contact processes, special needs, e.g. interpreter assistance, etc.

These are vulnerable citizens and the processing of their personal data is extensive. The processing has intrusive consequences for the citizens in question, as the AI solution's output is included in the caseworker's overall assessment and may have an impact on the citizen's specific financial situation.

Common requirements

No direct intervention in citizens' conditions

Few or less scope of special categories of disclosures

Includes a small number of vulnerable citizens, e.g. elderly, children, patients, etc.

Example of an example

An AI solution analyzes the municipality's waste sorting data, including the amount of different categories of waste from different districts in the municipality. The aim is to organize a more appropriate collection scheme in the municipality so that bins and containers are not emptied more often than necessary. The AI solution continuously adjusts expectations for waste production based on changes in incoming data, and the waste management company adjusts its collection routes on this basis.

Personal data about citizens will be processed to a lesser extent, as in sparsely populated areas it may be possible to link waste data from a district to individuals. The AI solution's processing of personal data may have consequences for the individual citizen in the form of less frequent or more frequent emptying of their bins.

Easier requirements

No interference in citizens' relationships

No or few special categories of personal data

Does not include vulnerable citizens, e.g. elderly, children, patients, etc.

Example of an example

An AI solution analyzes GPS and other driving data from home care vehicles to make more efficient use of available vehicles as the number of residents in need of care is increasing.

In this situation, personal data about (vulnerable) citizens will only be processed to a limited extent in the form of their residential address and the length of home care visits, and the processing has no immediate consequences for the citizens concerned. The result of the processing will thus only be used to organize the use of home care vehicles, and no changes are intended in the content of the service the citizens receive.

5.2.3 Special about consent

In certain cases, consent[41] may constitute the necessary legal basis for the processing of personal data. However, consent must meet a number of conditions to be valid. For consent to be valid, it must be *freely given, specific, informed* and *unambiguous*.[42] It is also a condition that the data subject has been informed that the consent can be *withdrawn*.[43]

The conditions for valid consent can be difficult to fulfill if you as a public authority want to process citizens' data using an AI solution.

First and foremost, this is because there will often be a clear imbalance in the relationship between the citizen and you as an authority. If the processing of information in a specific case, such as an application for a service or a permit, has an impact on the citizen's life situation - regardless of whether this is real or just perceived - the citizen's consent cannot be considered voluntary.[44]

The voluntariness condition can only be considered fulfilled in situations where there are no perceived or real negative consequences for the citizen if they fail to give their consent. This could, for example, be the case where the citizen gives consent to receive service messages by email or SMS about bulky waste collection in the local area.

The Danish Data Protection Agency's practice (j.nr. 2022-212-3676)

In 2022, the Danish Agency for Labor Market and Recruitment asked the Danish Data Protection Agency for an assessment of the issue of municipalities' legal authority to use the AI profiling tool Asta.

The purpose of the Asta tool was to perform a machine analysis of the risk of a newly unemployed unemployment benefit recipient's contact process with the job center becoming lengthy. In other words, based on a number of data about the citizen, Asta made a statistical prediction of the duration of the unemployment benefit case and the contact process for that person.

In this connection, the Danish Data Protection Agency stated, among other things, that in its opinion, consent could rarely be considered voluntary and thus constitute a valid basis for processing information about the citizen in the specific context where it was a public authority and where the public authority had control over the citizen's means of support.

This was true even if it would be possible in practice for the citizen to opt out of the treatment, i.e. avoid profiling, without it having a negative impact on the person concerned, e.g. in the form of stopping services. There would be a not insignificant risk that the citizen - regardless of this possibility - would feel pressured to consent to the treatment, e.g. to avoid appearing difficult or similar.

You should be aware that, in general, authorities must have legal authority to perform their tasks. This means that you cannot use consent from the citizen as a basis for processing if the processing falls outside the scope of the tasks you have been assigned as an authority. However, this does not preclude the authority from choosing consent as a basis for processing under the data protection rules in situations where the processing is within the scope of the authority's tasks. This may, for example, be the case if the authority wishes to

[41] Article 6(1)(a) of the General Data Protection Regulation.

[42] Article 4(11) of the General Data Protection Regulation.

[43] Article 7(3) of the General Data Protection Regulation.

[44] The Danish Data Protection Agency's opinion of July 5, 2022 in case no. 2022-212-3676 (the Asta tool). See also the European Data Protection Board's guidelines no. 05/2020 on consent, p. 8, point 16, and preamble recital no. 43 to the Data Protection Regulation.

Give citizens real freedom of choice when it comes to the processing of personal data. An example could be the authority's offer to use an app or similar.

Example 13

A municipality has developed an app that can predict where parking spaces are most likely to be available. Citizens can thus minimize the time they spend searching for parking options.

Among other things, the app shows the historical occupancy of each street at the time a driver drives by and how long it typically takes to find a parking space in that area. At the same time, the app suggests where there is the best chance of finding an available parking space.

In order for the app to function properly, the municipality requests citizens' consent to collect personal data about them, including location data.

It is the opinion of the Danish Data Protection Agency that the municipality will be able to process the citizens' personal data for this purpose based on their consent. In this situation, there are no negative consequences - neither real nor perceived - for citizens by not giving consent. The parking app is only a service offered by the municipality, and drivers are not disadvantaged in relation to the municipality in general if they do not use this service.

In addition to the voluntariness condition, the complicated processes and lack of transparency that often characterize AI solutions can stand in the way of valid consent. Consent must be specific and informed, and you must be able to handle the withdrawal of consent and stop processing the data in question. This can be a challenge in complex AI solutions where data is processed in many different ways. It is crucial for the validity of the consent that the citizen understands what their data will be used for and can opt in and out of these purposes. The more you want to use the data for, the more difficult it will be to fulfill these conditions.

As a public authority, you should therefore generally process citizens' data when using AI solutions on a legal basis other than consent. This also applies to the processing of special categories of data, which are discussed in more detail below.

In this connection, it should be noted that public authorities are in several cases subject to provisions in other legislation that require "consent" from the citizen. In these cases, you should be aware that even if the legislation in question requires consent from the citizen to process their data, this is not necessarily consent in the sense of the data protection rules. Such consent will often constitute a guarantee for citizens, but will not constitute the basis for the actual processing of personal data. In these cases, the basis for processing will often be the exercise of official authority under Article 6(1)(e) of the GDPR, where the legislation in question will constitute the relevant supplementary national legislation.

For example, consent under section 11a of the Act on Legal Certainty and Administration in the Social Field (the Legal Certainty Act) does not constitute consent under data protection law. The consent that the citizen can give under section 11a of the Legal Certainty Act aims to ensure citizens' rights and influence in the processing of cases covered by the Act, but does not constitute the basis for the processing of personal data. The basis for processing is the exercise of public authority.

5.3 Monitoring and post-learning

Once an AI solution is deployed, the model must be continuously monitored and, if necessary, retrained to ensure continued accurate output.

In a static model, the operational phase is clearly separated from the development phase and once the model is deployed, it only processes the personal data that is necessary for the operational purpose. Regular monitoring will be needed to ensure that the model continues to process and generate correct personal data, but the model itself does not change during use and you therefore have full control over its processing of personal data. If, during monitoring, there is a need to retrain the model, it is taken out of operation and retrained in a closed test environment on selected training data. The two purposes of development and operation are thus not present at the same time, and you only need to identify a processing basis for one purpose at a time.

A dynamic model, on the other hand, is continuously (re-)trained on the new data it processes while in use, and you will have less control over the processing of personal data as the model itself is constantly changing. This requires more extensive monitoring to prevent the model from evolving in an inappropriate direction, and you need to have the necessary legal authority both to process citizens' data to retrain and develop the solution and to process their data as part of the authority's operations, as the two purposes are present simultaneously.

5.4 Duty of disclosure

When considering deploying your AI solution, you should also be aware of your obligation to inform citizens about the processing of their data.

As described in section 4.2, developing an AI solution will always constitute a separate purpose that you must inform citizens about.

Most often, operating an AI solution will not constitute a separate purpose. This will usually be the case where the authority wants to solve a specific task and where the use of the AI solution is linked to the solution of this task.

In a few cases, operating an AI solution will constitute a separate purpose. This means that you will need to inform citizens that their personal data is being processed as part of your official duties *and* that you are using an AI solution for this purpose.

Especially for dynamic AI solutions, where development and training also occur while the solution is in operation, you must be aware that the development constitutes a separate purpose that you must inform citizens about. See more about this in section 4.2.

You can read more about what information you must provide to citizens etc. in section 3 on the duty of disclosure in the Danish Data Protection Agency's general guidance on the rights of data subjects.[45]

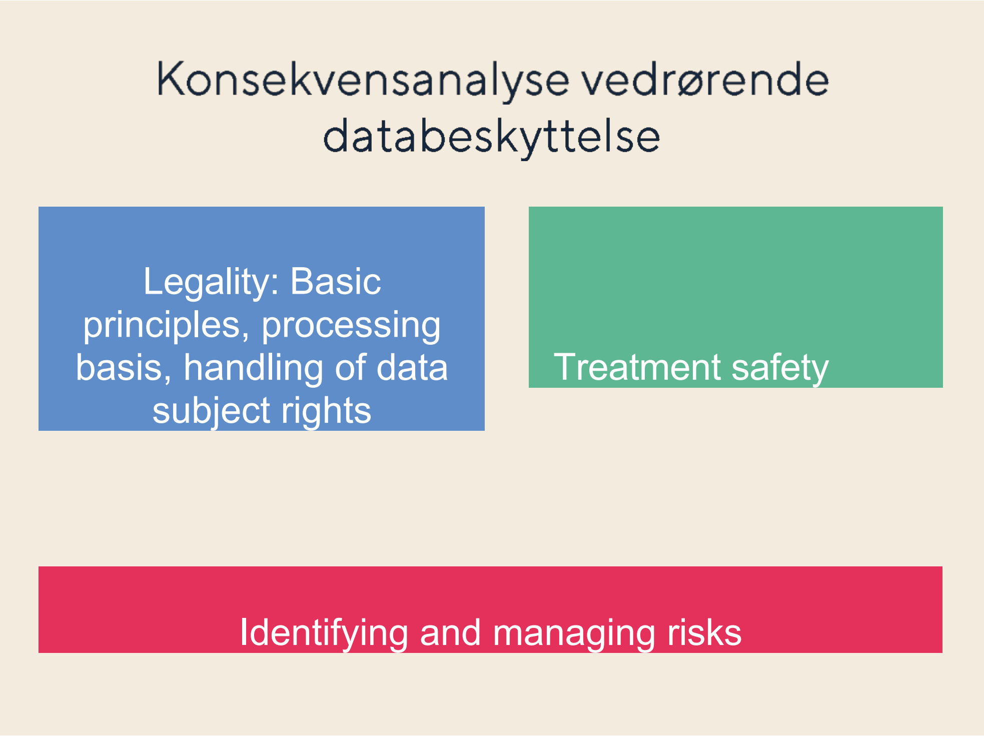
[45] The Danish Data Protection Agency's guidance on data subject rights, June 2018.

6. Impact assessment

If the processing of personal data is likely to result in a high risk to the rights and freedoms of natural persons, you must carry out a data protection impact assessment *(DPIA)* prior to processing.[46] This is particularly relevant when you want to process personal data using new technology.[47] In the opinion of the Danish Data Protection Agency, this will usually be the case when developing and using AI solutions.

A data protection impact assessment is a process that aims to:

- to describe the processing of personal data,
- assess the necessity and proportionality of the treatment; and
- to help manage the risks to the rights and freedoms of natural persons arising from the processing of personal data.



Konsekvensanalyse vedrørende databeskyttelse

Legality: Basic principles, processing basis, handling of data subject rights

Treatment safety

Identifying and managing risks

Risk assessment is also an integral part of the process of determining an appropriate level of processing security, but the impact assessment goes one step further than the risk assessment and includes, among other things, an assessment of how the intended processing activities meet the basic requirements of lawfulness.

The impact assessment also includes an assessment of the risks of deviation from the lawful and intended processing activity. In addition, the rules on impact assessments include a process for the involvement of the Data Protection Officer, as well as possible consultation with the Data Protection Authority and/or citizens. The specific purpose of an impact assessment is to ensure a systematic

[46] Article 35 of the General Data Protection Regulation.

[47] Article 29 Working Party Guidelines on Data Protection Impact Assessment (DPIA) and determining whether the processing operation is "likely to result in a high risk" under Regulation (EU) 2016/679, WP248, p. 9

investigating situations that may lead to a high risk to the rights and freedoms of natural persons.

Under certain conditions, it is a requirement to conduct an impact assessment. The conditions are set out in (i) the GDPR, (ii) the European Data Protection Board's guidelines on impact assessments[48] , and (iii) the Danish Data Protection Agency's list of processing activities that are always subject to the requirement for an impact assessment[49] .

The Danish Data Protection Agency believes that the processing of personal data as part of the development and/or operation of AI solutions will almost always trigger several of the criteria that determine whether an impact assessment must be conducted. This is because:

- AI is considered a so-called "new technology", which is one of the criteria in the DPA's list of activities that are always subject to the requirement of an impact assessment, and
- that the development and/or operation of AI often involves (i) processing of special categories of data, (ii) processing of data of vulnerable persons or (iii) processing of personal data on a large scale, which are three other criteria listed in the Article 29 Working Party guidelines on impact assessment.

The impact assessment should be carried out already in the planning and development phase of an AI solution. This way, you will be aware of and address any data protection challenges associated with the AI solution as early as possible in the process.

The impact assessment will also enable you to demonstrate compliance with data protection principles by design and by default. These rules require you to implement technical and organizational measures that ensure compliance with data protection rules - and thereby provide appropriate safeguards for citizens' rights.

The impact assessment supports this process and should be seen as an ongoing commitment, especially when, as is often the case in an AI solution, the treatment situation is dynamic and constantly changing.

The principle of accountability is a common thread throughout the GDPR. In this context, the rules on data protection by design and by default emphasize the obligation to incorporate data protection rules already from the design phase and to monitor the effectiveness of the measures chosen throughout the life of the system. The requirement of data protection by design and by default applies both during the development and design of a system and during its use.

AI systems that process personal data should therefore be designed from the outset to ensure effective implementation of data protection rules. Appropriate measures must be implemented in advance to ensure that the requirements and protection considerations of the Regulation are handled as an integral part of the entire system's processing of personal data. Among other things, it must be ensured that the training data is representative, that the system's output is reasonable, that there is no unlawful discrimination, and that data is processed with the necessary security, including, for example, by using pseudonymization.

There are no specific formal requirements or methodology for conducting an impact assessment. However, as a minimum, the analysis must include (i) a systematic description of the intended processing activities and the purposes of the processing, (ii) an assessment of the risks to the rights and freedoms of citizens, and (iii) the measures envisaged to address those risks and demonstrate compliance with the GDPR.[50] This must be a proper assessment of risks that enables you to take measures to mitigate them.

[48] Guidelines on Data Protection Impact Assessment (DPIA) and determining whether processing is "likely to result in a high risk" under Regulation (EU) 2016/679 (WP248, rev. 01).

[49] The list compiled in accordance with Article 35(4) of the GDPR can be found here.

[50] Article 35(7) GDPR and preamble recitals 84 and 90.

The use of artificial intelligence by public authorities
2023 Datatilsynet

Reprinting with attribution is permitted

Published by:
Data Protection Authority
Carl Jacobsens Vej 35
2500 Valby
T 33 19 32 00
dt@datatilsynet.dk
datatilsynet.dk

Data Protection Authority
Carl Jacobsens Vej 35
2500 Valby
T 33 19 32 00
dt@datatilsynet.dk
datatilsynet.dk