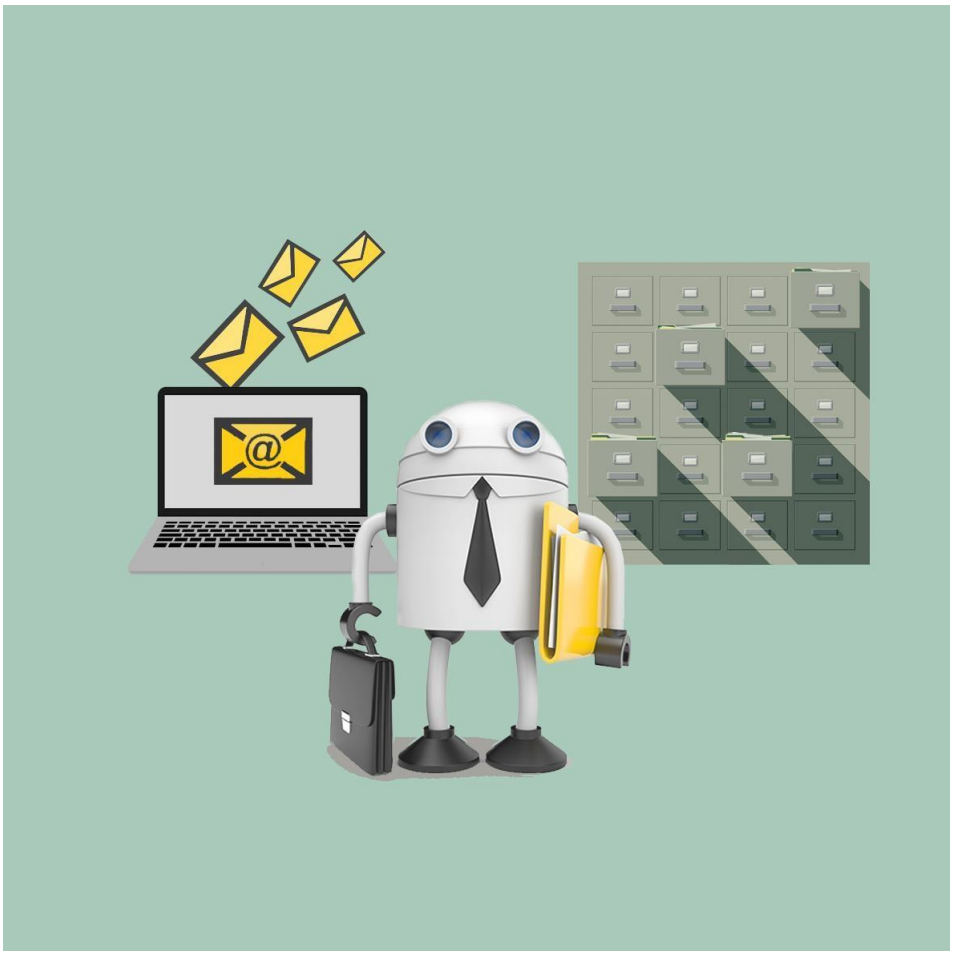


DIGITAL EMPLOYEE

Exit report from the sandbox project with Simplifai and NVE

Topics: legal basis and data protection by design

Datatilsynet

Contents

What is the sandbox?

In the sandbox, participants and the Norwegian Data Protection Authority jointly explore issues relating to the protection of personal data in order to help ensure the service or product in question complies with the regulations and effectively safeguards individuals' data privacy.

The Norwegian Data Protection Authority offers guidance in dialogue with the participants. The conclusions drawn from the projects do not constitute binding decisions or prior approval. Participants are at liberty to decide whether to follow the advice they are given.

The sandbox is a useful method for exploring issues where there are few legal precedents, and we hope the conclusions and assessments in this report can be of assistance for others addressing similar issues.

Summary

The transition from analogue to digital mail has challenged old systems for record-keeping and the archiving of mail and important documents. A report indicates that more than 25 per cent of all important information in public administration “is lost” as a result of inadequate archival procedures.

That is why the company Simplifai has developed a digital archive employee – called DAM for the Norwegian acronym – to help administrators in the public sector record and archive documentation sent via e-mail. DAM works as a form of decision support for administrators. The goal is to automate this work as much as is safely possible, in the hope that this ensures all important items are archived correctly.

In the sandbox, Simplifai and the Data Protection Authority have looked into whether data protection legislation permits public administration organizations to implement machine learning to record and archive e-mails. In collaboration with the Norwegian Water Resources and Energy Directorate (NVE), they have explored how public bodies can make informed choices when purchasing intelligent solutions, such as DAM.

Conclusions

- **Lawfulness.** Public bodies do have a legal basis for using DAM for decision support in connection with archiving and record-keeping. There is, however, some uncertainty concerning whether this legal basis extends to the use of personal data in the further development of the model (continual learning), unless the personal data is anonymized. The use of DAM also complies with national working environment regulations. The Data Protection Authority recommends technical and organizational measures, such as guidelines that prohibit or limit the use of private e-mails.
- **Data protection by design.** This project has revealed a considerable need for guidance on how the public sector can ensure data protection by design when procuring intelligent solutions. The project has resulted in general recommendations for the steps a public-sector body can take: Gather information, consider whether machine learning is appropriate for the specific need and make demands.

Going forward

Work on this project has highlighted a considerable need for guidance and tools to meet the requirements for data protection by design when procuring solutions where the data controller is obligated to ensure data protection by design.

The sandbox project recognizes the need for more actors to come together to increase knowledge of data protection by design in public procurement. It would also be useful to have practical examples of requirements the public sector can make in competitive bidding processes involving technology based on machine learning.

And – to make it easier to use machine learning to solve the challenges associated with record-keeping and the archiving of documentation in time, the sandbox project recommends new archive legislation to regulate this responsibly.

January 2023

This PDF corresponds to the first version of the report, as it was published on the Norwegian Data Protection Authority's website in January 2023. Both technology and the law are constantly evolving, so the reports *may* need to be adjusted or clarified as time goes on.

If this PDF differs from the text published on the Norwegian Data Protection Authority’s website, please assume that the website text contains the Authority's prevailing advice.

About the project

Simplifai has developed a digital archive employee – called DAM for the Norwegian acronym – to help administrators in the public sector record and archive documentation sent via e-mail. As the public sector handles key social functions, the sector has its own requirements for the handling of documents. Records must be kept of relevant correspondence, so that the public can request access to it and see what is going on. Documentation that is deemed to have archival value, and that has been used in the administrative process, must be preserved for posterity through archiving.

Simplifai was established in 2018, with the goal of making artificial intelligence simple and available to everyone. In 2022, Simplifai has approx. 140 employees and supplies various digital assistants worldwide. The main customer sectors have been the public sector, as well as banking, insurance and finance. What all of these sectors have in common is stringent regulatory requirements. That is why Simplifai is actively working to develop secure solutions based on artificial intelligence for the automation of work processes related to digital communication and documents.

The project wanted to invite a public body that could share their knowledge of and experience with archiving and data protection in practice. Resource persons from the Norwegian Water Resources and Energy Directorate (NVE) provided useful perspectives from their role as a potential buyer of DAM.

Why do we need DAM?

Back when we all sent letters through the postal service, skilled archivists in the public sector ensured that all all incoming and outgoing letters were registered under the appropriate heading. The administrators themselves did not have to consider the matter of record-keeping and archiving. Today, however, a great deal of correspondence comes directly to administrators via e-mail. In addition, there is a lot of internal communication taking place on other platforms. The amount of documentation has increased and appears rather overwhelming to many of us.

With digitalization, the responsibility for archiving and record-keeping has largely shifted, from archivists to the administrators themselves. According to a survey by Menon, the majority of administrators believe they are able to archive less than 75 per cent of all important information.

Read the report: Mangelfull arkivering koster Norge dyrt (arkivverket.no – in Norwegian only)

And this occurs at every level. In a report on archiving and transparency in public administration, the Office of the Auditor General points out that inadequate archiving and record-keeping in key cases erodes the foundation for public discourse and democratic access to information and control. Among other things, they point to the Coronavirus Commission's conclusions that key parts of the interaction between the Norwegian government and parliament in the response to the Covid-19 pandemic have not been documented. "The lack of documentation and archiving weakens access to knowledge of and insight into central assessments and decisions that were made in the response to the pandemic, deemed the biggest national crisis since World War II."

Read the report from the Office of the Auditor General here (riksrevisjonen.no – in Norwegian only)

How does DAM work?

DAM is a solution that helps the employee select which e-mails to archive and enter into the public record. It also aids in the actual recording and archiving process.

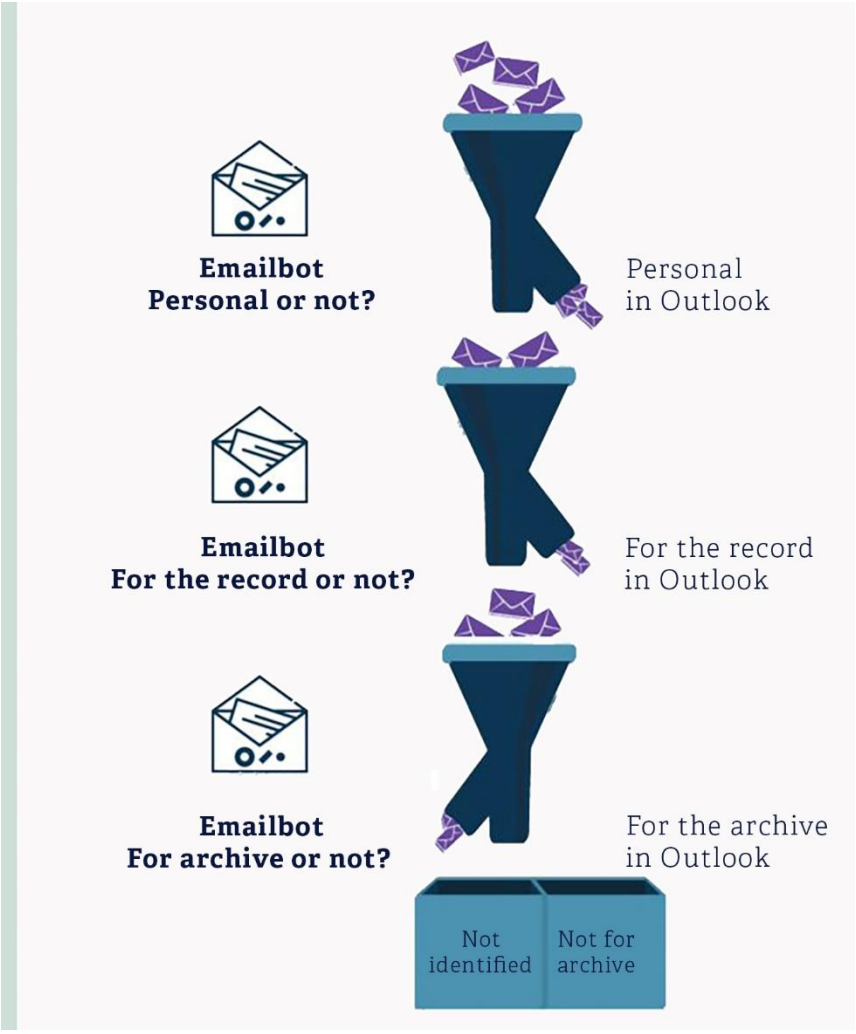
The solution has been developed using machine learning. The method used in this project is called natural language processing (NLP), specifically natural language understanding (NLU), which means that the solution recognizes text within the e-mail, including text in attachments, if relevant. A new development for the digital employee is that machine learning (NLU/NLP) has been combined with optical character recognition (OCR).

The digital employee consists of three modules – so-called bots. Before the digital employee suggests what to do with the e-mail, each bot considers one of the following questions:

1. Is the e-mail personal in nature?
2. Should the e-mail be entered into the record?
3. Should the e-mail be archived?

Based on these assessments, an e-mail will be tagged in categories of personal/not personal, whether it should be entered into the record and whether it should be archived. If a human employee agrees with the assessment, the e-mail can be tagged in the category "Archiving OK". The employee will then personally approve the archiving of the e-mail. Furthermore, the employee may personally archive the e-mail, or all e-mail tagged with "Archiving OK" can be archived by the digital archive assistant the following night, based on the employee's actions.

The employee may choose to add personal rules that exempt certain inbox content from being processed by DAM. This could include exempting certain senders or certain key words in the subject field or e-mail body.



Simplifai's goal is for the digital employee to contribute to

- more e-mails being entered into the record and archived
  - greater legal protection for all parties who are in contact with the public sector
  - better control functions for the public sector and the media
- freeing up time for employees – this time is better spent on core tasks

Objective of the sandbox process

Before Simplifai was admitted into the Data Protection Authority's sandbox, the company had participated in a StartOff project with the National Archives Services of Norway, led by the Norwegian Agency for Public and Financial Management (DFØ). The goal of this project was to explore how artificial intelligence could be used to automate or support e-mail record-keeping and archiving. DAM was developed in this project.

See the presentation "Fullautomatisert epostarkivering for Arkivverket" (youtube.com – in Norwegian only)

The main focus of the StartOff project was to strike a good balance between data protection considerations and the desire to find effective ways to meet requirements under the Freedom of Information Act and Archive Act. During the project, it became clear that the potential legal hurdles associated with the use of artificial intelligence in e-mail processing did not so much come from archive requirements, but rather from the Personal Data Act and the Working Environment Act. The project assumed that the legal complexity would be lower when DAM was applied to organizations' centralized e-mail addresses, such as post@datatilsynet.no than on personal e-mail addresses. Personal e-mail addresses, such as kari.nordmann@datatilsynet.no, were not included in the StartOff project.

Even so, it is only when DAM is applied to personal e-mail addresses that the effect of the service really kicks in, as administrative processing primarily takes place using personal e-mail addresses. At the same time, using this technology on personal e-mail addresses raises some key questions about privacy. This was the background for the application to join the Data Protection Authority's sandbox for responsible artificial intelligence.

The sandbox process has had **two primary goals:**

1. Consider whether the use of DAM is lawful for a public-sector body

Finding out whether the implementation and further development of DAM would be lawful for a public-sector body was a primary concern for Simplifai before launching this solution on the market. The primary goal of lawfulness was divided into three separate issues:

1. Explore which legal basis from the General Data Protection Regulation (GDPR) would be relevant for a body implementing the planned solution.

2. Explore whether the use of special categories of personal data is permitted under the GDPR.

3. Clarify whether the planned solution would be in conflict with the prohibition on monitoring in the E-mail Regulations.

Based on resource considerations, the project decided not to explore the transfer of personal data to countries outside the EEA through cloud computing. The project has based its considerations of lawfulness on Simplifai's assessment that the company is a data processor. This means the project has not considered whether Simplifai is a data controller or data processor.

2. Give recommendations on data protection by design to public bodies intending to procure a solution that is entirely or partially based on machine learning (artificial intelligence)

As a provider of intelligent archive solutions, Simplifai wants to develop a service that meets both regulatory requirements and other needs public bodies have in terms of data protection and privacy. Intelligent solutions are becoming more and more prevalent in the public sector, and purchasing competence is pivotal in making sure these solutions are implemented in a good way.

In this project, the Data Protection Authority, Simplifai and NVE worked together to develop recommendations for how public bodies can set requirements for solutions based on machine learning in procurement processes, to ensure that the public sector complies with the requirement for data protection by design.

Is the implementation of DAM lawful?

Lawfulness is a fundamental principle in the General Data Protection Regulation (GDPR). The fundamental principle is that anyone who wants to collect, store or otherwise process personal data must comply with the requirements of the GDPR and other relevant legislation. In the following, we will examine more closely the requirement that there must be a legal basis for processing personal data.

In the Simplifai sandbox project, we have explored what leeway exists for the use of personal data in a decision-support tool based on machine learning.

The GDPR applies with equal force in 30 countries in Europe. In addition, Norway has special rules on privacy, including privacy in working life. These special rules are laid down in Working Environment Regulations and are intended to protect workers from unnecessarily invasive monitoring or control.

How did we approach the assessment?

Initially, the sandbox project discussed whether we should base our assessment on full automation of DAM or on DAM as a decision-support tool. While automation would have the greatest effect, the risk of violating data protection legislation would also be higher, in that content from personal inboxes would be entered directly into the public record. The project decided to focus on DAM as a decision-support tool.

The project organized two workshops focused on lawfulness. As preparation for these workshops, Simplifai provided information about the technical solutions and the purpose of the different modules/bots. In addition, NVE expressed its views on, among other things, the purposes of the processing performed by the modules and the legal basis the directorate believed it had for implementing the solution. Based on these thorough preparations, the Data Protection Authority outlined a framework for use, continual learning and potential measures for making DAM more data protection-friendly.

We opted to split the legal basis issue into phases: (1) the use phase, where the algorithmic model is used in the administrative process, and (2) the continual learning phase, where the algorithmic model is improved.

Simplifai came to the sandbox with an already-developed tool. The project was limited to the question of whether Simplifai had a legal basis for developing DAM.

Legal basis for the use phase

In this phase, DAM is being used as a decision-support tool and serves as a guide for administrators when they are considering whether an e-mail is personal in nature, whether it should be entered into the record and whether it should be archived.

The processing of personal data in record-keeping and archiving currently takes place pursuant to Article 6 (1) (c): “processing is necessary for compliance with a legal obligation to which the controller is subject”.

The legal obligation to maintain an archive is laid down in Section 6 of the Archive Act and Sections 9 and 10 of the Regulations Relating to Public Archives. In addition, the obligation to keep records is laid down in Section 10 of the Freedom of Information Act.

The sandbox project recommends that public bodies base their processing on Article 6 (1) (c) even when using digital tools, such as DAM, to make the archival process more efficient and systematic.

Legal basis for the continual learning phase

The ability of artificial intelligence to learn is a major technological breakthrough. When the algorithm learns through being used, we call this continual learning. The inference (i.e. the result of the algorithm being used on a new data source) is included in the algorithm, which means the tool is dynamically adjusted to become more accurate.

Continual learning can be challenging in terms of data protection law. Even if there is a legal basis for using artificial intelligence for archival purposes, such as processing being necessary to comply with a legal obligation, this does not necessarily mean that there is a legal basis for continual learning of the algorithm.

Article 6 (1) of the GDPR lists six alternative conditions that may provide a legal basis for the processing of personal data. The alternatives most relevant in this case are the conditions found in (c) and (e):

- Processing is necessary for compliance with a legal obligation (c).
- Processing is necessary for the performance of a task carried out in the public interest (e).

In both cases, the GDPR requires statutory authority for the processing in other legislation, see Article 6 (3).

The current Archive Act includes no wording to indicate that e-mails or other material with archival value can be used in the further development of machine-learning systems or other digital tools. DAM would be on firm ground if the Archive Act explicitly provided, that material processed in connection with archival obligations may also be used to improve or further develop digital tools. Even so, continual learning may be considered as a consequence of using the tool, thus establishing a legal basis pursuant to Article 6 (1) (c), as in the use phase.

For society, there is a clear need to improve the efficiency of the archival process and increase associated legal protections. Deciding how access to archive material used in the development of machine learning should be handled, is a matter for legislators. The Data Protection Authority has seen examples where some organizations are authorized to use collected data in the development of IT systems, see Section 45 b of the Act Relating to the Norwegian Public Service Pension Fund.

One way to approach this problem is to design the algorithm in such a way that the inference does not contain any personal data. In this case, the GDPR would not apply, and the algorithm can be trained without restrictions. Another option, if the inference does contain personal data, is to anonymize the data before the algorithm is set to engage in continual learning. Then it would no longer be personal data, and the model can be trained without needing a legal basis.

Another option could be to enable each DAM-user to turn the algorithm's continual learning capability on or off, either in general or with the option to exempt individual e-mails from inference. This could also be positive in terms of transparency concerning how the tool processes personal data, and could help ensure personal e-mails, etc., are not included in the training. One side effect of this approach, however, may be that the underlying data could reinforce an incorrect archival practice.

Is the processing of special categories of personal data lawful?

In the inbox of an administrator in NVE, DAM could find information from the local trade union as well as an e-mail to the boss about illness. Some employees may also use their work e-mail for private correspondence. In these e-mails, DAM may find information about the administrator's religion and sexual orientation, for example. All of these categories are considered special categories of personal data, see Article 9 of the GDPR. The processing of these categories of personal data is only permitted if the conditions set out in Article 9 (2) have been met.

The sandbox presumes that the condition of "processing [being] necessary for reasons of substantial public interest", see Article (2) (g), would apply to NVE's use of DAM. This is reflective of the preparatory works to the NAV Act, where the ministry assumes that efficient processing by the Labour and Welfare Administration and the Norwegian Public Service Pension Fund would be covered by a "substantial public interest", as laid down in Article 9 (2) (g).

Another key part of this condition is that the processing must be "proportionate" to the aim pursued. In its assessment of proportionality, the sandbox has emphasized that the suggestion will be limited to the administrator only, and that the impact on privacy will therefore be limited. In addition, we recommend establishing guidelines that prohibit or limit the use of the organization's e-mail address for private purposes, in order to limit the extent of personal e-mails being processed by the solution.

Prohibition against monitoring, see the Email Monitoring Regulations

So far, we have focused on the GDPR. It is, however, also relevant to consider DAM in relation to the Email Monitoring Regulations and their prohibition on "monitoring the employee's use of electronic equipment, including use of the Internet" (Section 2 (2)).

So it may be lawful for an employer to implement DAM, provided its use does not entail monitoring the employees' activities. Could the implementation of DAM be considered monitoring of the employees' use of electronic equipment?

What counts as "monitoring" is not defined in more detail in the regulations. The preparatory works of similar regulations in the previous act highlight that the measure must have a certain duration or take place repeatedly. Monitoring differs from isolated incidents of access, which is permitted in several circumstances. The preparatory works also emphasise that it is not solely a question of whether the purpose is to monitor. The employer must also attach weight to the question of whether the employee may perceive the situation as monitoring.

Past decisions by the Data Protection Authority are not conclusive with regard to whether the employer is required to actually see the personal data for it to count as monitoring. Monitoring is a broad term, and a natural linguistic understanding of the term may entail that collection and systematization are also affected by the prohibition. The fact that the provision is aimed at the *employer's* monitoring indicates that the employer must, at the very least, be able to access the data concerning the employees in order to be subject to the prohibition. This was the sandbox's position in the project with Secure Practice as well.

After discussions in the sandbox project, there was a consensus that the prohibition on monitoring would not extend to the use of DAM for decision support. We have emphasized that the solution merely makes a suggestion in the administrator's inbox, and that the information about the categories does not go any further.

In addition, the sandbox project recommends that technical and legal measures be implemented to prevent the employer from having access to the information collected by the solution about the individual employees. One such measure could be to implement a rule that prohibits or limits the use of the organization's e-mail addresses for private purposes, as mentioned in the section on special categories of personal data.

Data protection by design when procuring intelligent solutions

The principles of data protection by design have been around for a long time, but they were not legally established until the GDPR entered into force in 2018. This requirement is based on the idea that a system must have data protection measures built-in from the beginning in order for the user of the system to be able to comply with data protection rules. Tacking on measures to "fix" data protection problems after the fact would, in the long run, result in solutions that are less data protection-friendly than those that have these features built in by design.

Article 25: Data protection by design and by default

1. Taking into account the state of the art, the cost of implementation and the nature, scope, context and purposes of processing, as well as the risks of varying likelihood and severity for rights and freedoms of natural persons posed by the processing, the controller shall, both at the time of the determination of the means for processing and at the time of the processing itself, implement appropriate technical and organizational measures, such as pseudonymization, which are designed to implement data-protection principles, such as data minimization, in an effective manner and to integrate the necessary safeguards into the processing in order to meet the requirements of this Regulation and protect the rights of data subjects.
2. The controller shall implement appropriate technical and organizational measures for ensuring that, by default, only personal data which are necessary for each specific purpose of the processing are processed. That obligation applies to the amount of personal data collected, the extent of their processing, the period of their storage and their accessibility. In particular, such measures shall ensure that, by default, personal data are not made accessible without the individual's intervention to an indefinite number of natural persons.
3. An approved certification mechanism pursuant to Article 42 may be used as an element to demonstrate compliance with the requirements set out in paragraphs 1 and 2 of this Article.

Who is responsible for ensuring data protection by design?

Compliance with Article 25 requires that programmers and developers implement suitable technical solutions. Only those who are developing the solution are able to include good data protection measures from the beginning. Even so, the responsibility for compliance with this requirement does not rest with the software provider or the data processor, but with the data controller.

See the Data Protection Authority's guide to data protection by design.

This is how the Regulation works. The data controller is responsible for the processing of personal data, see Article 5 (2). Responsible data controllers will demand good data protection measures when they purchase products. In this way, providers who can guarantee good data protection by design will have a competitive advantage. This requires that the data controller purchasing the machine-learning solution is knowledgeable about

- data protection by design
- machine learning
- contracts
- procurement rules (for data controllers who are subject to the Public Procurement Act)

What should be included by design?

The system should include the following by design:

- basic data protection principles
- the data subject's rights and freedoms

This includes everything, from abstract requirements, such as fairness, to more practical requirements, such as procedures for deletion. So, when has a provider been able to include data protection by design? The law does not explicitly define what are considered appropriate technical and organizational measures. The Data Protection Authority's guide to data protection by design emphasizes that the measures must be capable of effectively ensuring data protection and provides examples of key elements for each data protection principle. The data controller must decide which measures and which levels of measures are required, by considering the following factors:

- the status of technological development
- implementation costs
- the nature, scope, purpose and context of the processing
- the risk to the rights and freedoms of data subjects – likelihood and severity

In other words, Article 25 of the GDPR provides that the context of the processing shall be a determining factor in how and to what degree data protection is to be included in the solution by design. All of the GDPR's provisions shall nevertheless apply, but how a specific procedure for deletion is built into the solution must be determined on the basis of what is the most suitable approach.

Data protection by design and machine learning

Machine learning may complicate the assessment of which measures would be best suited for ensuring data protection by design. The reason for this is that machine learning challenges data protection and privacy in the following ways:

- Most buyers acquire algorithms that are already developed (off the shelf) and repurpose them for various purposes
- Training machine learning solutions requires vast quantities of data
- We still do not know very much about how machine learning may affect individual rights and freedoms and have unexpected consequences, such as
    - discrimination
    - lack of transparency

Machine learning algorithms may be extremely opaque. It can be difficult to explain the logic on which the results of a machine learning solution are based – this is often referred to as a *black box problem*.

If the risk to the data subject's rights and freedoms is related to discrimination, for example, the measure must reduce the risk of discrimination. This means the algorithms must be designed so as to prevent discrimination.

Machine learning solutions are often bought as so-called "off-the-shelf" products. This means that the solution has been developed by someone other than the data controller for the particular use of the tool. The buyer will be responsible for acquiring a solution that has data protection by design. The buyer (data controller) may appoint a data processor to guarantee compliance, but in order to select appropriate measures at the appropriate level, the data controller must also understand the impact the solution will have on the privacy of the data subjects. Because machine learning solutions are often complex and

difficult to explain, this will pose a greater challenge for data controllers than technology that does not include machine learning.

In the sandbox project, we wanted to highlight how data protection by design could be ensured in connection with the procurement of intelligent solutions. In order to gain an understanding of what public bodies need help with, we chose to conduct a case study to map the landscape and prepare some basic recommendations.

How did we prepare our recommendations?

In order to clarify which information public-sector bodies need to make good decisions in their procurement of machine learning solutions, the project group decided to conduct an interview with a public body, NVE. The purpose of this interview was to establish an overview of the needs public-sector bodies have for information about data protection by design. Based on this interview, we chose to prepare recommendations with points public-sector bodies can follow up on. NVE was also invited to provide feedback on an early draft of these recommendations, to see if they met their needs. In addition to NVE, representatives from the police service and Simplifai also provided feedback on the draft recommendations.

NVE was chosen as a case study because they had already been in contact with Simplifai about their archive system. As a public-sector body, NVE is an interesting choice for a case study because their activities do not entail broad processing of personal data, compared to, say, a municipal authority. Even so, NVE does need to process some personal data about its employees, and DAM would be part of that. NVE is also a public body with experience of developing its own proprietary solutions, which means they would have an interesting perspective in a procurement situation such as this one.

Feedback from NVE after the interview indicates that there is a considerable need for guidance on data protection in connection with the procurement of intelligent solutions. The other workshops in the sandbox about data protection by design also confirm the need for guidance.

Information about data protection by design is also provided by the Data Protection Authority in its activities outside the sandbox – in supervisory activities and in monitoring compliance with the GDPR. We see two factors, in particular, driving the considerable need for information.

Why is the need for information so great? Our findings

One interesting finding is that data protection by design seems to be exclusively associated with information security. Information security is a field that has been explored much more than data protection, so that may be part of the reason. We see a demand for advice on how organizations can request information and make demands beyond information security when they check a solution's data protection by design. These experiences are shared by the Privacy Commission.

Read more about this in the Privacy Commission's Official Norwegian Report 2022: 11 (regjeringen.no – in Norwegian only)

Machine learning and artificial intelligence are so complex that most organizations without specific technical expertise lose sight of which measures are considered "appropriate". The difference in technological expertise between the developer of the tool and the data controller is even greater in machine learning than it is in other contexts. This gives providers a key role in providing information about the product. Our perception is that customers so far have only to a very limited degree requested data protection by design for off-the-shelf products based on machine learning.

Background for choices

Because both data protection by design and machine learning are complex and, to a certain extent, new concepts, our recommendations are somewhat generalized for the time being. Examples of the types of requirements that can be demanded of providers of machine learning tools have nevertheless been included, because they are especially relevant for machine learning solutions.

Recommendations for data protection by design in the procurement of solutions based on machine learning

Build competence

Employee training is a key measure for ensuring data protection by design in practice. Build employee competence with regard to data protection by design, machine learning, contracts and procurement.

Both the data security officer and the data protection officer can be good resources in procurement processes involving machine learning solutions, and should be involved as early as possible.

Here is a list of useful sources:

Data protection by design

- Data protection by design and by default | Norwegian Data Protection Authority

This guide covers all of the data protection principles and provides simple recommendations for how these can be included by design. The guide provides a good description of the requirement for data protection by design.

- Software development with data protection by design and by default | Norwegian Data Protection Authority

This guide goes through the different stages of the development process and describes how data protection can be included by design. It is written for people with technical knowledge, but everyone can benefit from the practical recommendations for how to include data protection by design.

Machine learning

- Artificial intelligence and privacy | Norwegian Data Protection Authority

Easily understandable report describing artificial intelligence and how it relates to privacy.

- Free online introductory course on artificial intelligence, available to all (elementsofai.com)

A comprehensive and well-designed online course on artificial intelligence for those wishing to learn more about the mechanisms behind it.

- Norwegian Board of Technology report on artificial intelligence and machine learning (pdf – in Norwegian only)

Comprehensive report on artificial intelligence from a technical perspective. This report also covers potential challenges posed by machine learning.

How to define requirements and perform evaluations in a procurement process

- Anskaffelser.no (Public Procurement – Information in English)

General information about procurement, especially how a public procurer can go about collecting information about products and defining requirements.

Consider: Is machine learning the most appropriate approach?

Which need is the machine learning tool intended to meet? Our recommendation is to consider the context in which you will be processing personal data and to consider whether a rule-based tool can achieve a more privacy-friendly solution than a machine learning tool. If, for example, you are using the tool to support decision-making processes that affect citizens, you must define different requirements for the tool than if you are using it to make internal decision-making processes more efficient.

> ### Rule-based tools
>
> Rule-based tools are tools where the algorithm is static and based on fixed rules, as opposed to algorithms that are dynamic and make predictions based on patterns in the source data.

As an aid in this comparison, we refer to some challenges posed by machine learning tools, as highlighted in a report by the Privacy Commission of 26 September 2022.

First, machine learning algorithms demand large quantities of data to develop an accurate model. (Read more about that on page 17 in NDPAs AI report.) Consistent principles in the GDPR include the principle of data minimisation, which entails limiting the collection of personal data to what is necessary to accomplish the purpose of the processing, and the principle of purpose limitation, which entails limiting the use of personal to the purpose for which it was originally collected. Furthermore, the provision concerning data protection by design specifies that the data controller shall ensure that, "by default", only personal data necessary for the purpose of the processing is processed, see Article 25 (2).

Second, the Commission emphasizes that the source data collected and used to train a machine learning algorithm may contain errors and defects. If such defects are present in the source material, the result, i.e. the predictions of the machine learning algorithm, will be affected by these defects. This means the organization must be very conscious of which data is used to train a machine learning algorithm. Depending on the context of the processing, the data material may also be outdated, and therefore misleading. In order to maintain good data protection throughout the lifespan of the solution, it will therefore be necessary to make regular adjustments to maintain the accuracy of the predictions.

Furthermore, machine learning algorithms will rarely allow for sufficient transparency and predictability. Machine learning solutions are generally not transparent, according to the Privacy Commission's report. In many cases, machine learning algorithms will also be dynamic. This means that their logic may change, even after the algorithm has been implemented.

Finally, the Commission points to the risk of predictions from machine learning solutions being used without critical reflection. This issue was also emphasized in the exit report for the NAV sandbox project. In practice, it could mean that what was intended as a decision-support system in reality becomes an automated decision-making system.

The Data Protection Authority would also like to emphasize the importance of the organization taking into account which of the data subjects' rights and freedoms the data controller is responsible for protecting, and whether the data controller is still able to uphold this obligation if a machine learning solution from an external developer is implemented.

If you conclude that a machine learning tool is the best solutions for your needs, what steps can you take to define requirements for the product?

Ask questions, dig deeper and make demands!

We have some recommendations for how you, as data controller, can request the documentation you need to assess data protection by design in the various solutions offered.

Get the solution explained in a way you understand

Machine learning solutions can be very complex, and it is important that everyone who will be using the solution understands how it works. In order to meet the requirement for data protection by design in the procurement phase, it is important that those involved in assessing the bids submitted understand how data protection can be safeguarded when it is used.

We recommend that those who are in the process of procuring a machine learning solution request an easy-to-understand, detailed description of what the solution actually does.

Request to see data flows and processing records

A provider acting as data processor must be able to account for which processing activities they perform on behalf of the data controller, see Article 30 (2). As data controllers, you can ask to see this processing record before you enter into a contract for procurement of a machine learning solution. Even providers who are not data processors should be able to account for which types of personal data will be processed by the solution. The processing record is essential for gaining an overview of how the processing actually takes place. It would also be useful to view the processing record in light of the accessible description mentioned above. It could also be relevant to get a description of how the collected data moves in the solution, i.e. the data flow.

Ask how the transparency requirement is handled

An even bigger issue related to the use of machine learning solutions is how to ensure one is able to explain decisions made with decision support from a machine learning solution.

As described above, the lack of transparency is a recurring issue we encounter in connection with machine learning solutions. Even so, the organization has an obligation to inform the data subject. This obligation to provide information extends to the underlying logic in certain types of automated decisions.

It is therefore important to ensure, before a contract is signed, that there are ways to present how the algorithm weighs variables and how accurate the algorithm is. The latter can, for example, be handled by the solution indicating how likely it is that the prediction is accurate.

Ask what the mechanisms are in place to identify and mitigate algorithmic bias

Machine learning creates some new problems related to system ethics. Potential algorithmic bias may challenge the principle of fairness in Article 5 (1) (a) of the GDPR.

See also the Data Protection Authority's AI report, p. 15 (pdf).

The Privacy Commission points out that this type of bias may occur when there is a lack of transparency in the solution. Furthermore, these biases will be exacerbated if the solution is used without critical reflection or is fed incorrect data.

In connection with procurement of a machine learning solution, it may be a good idea to find out whether the solution has mechanisms in place to identify potential bias, how often the algorithm should be adjusted and how. If it is possible to identify the situations in which the algorithm may be less accurate, it will be easier to implement appropriate measures to reduce the consequences of this bias. Another alternative would be to retrain the algorithm as soon as its accuracy falls below a pre-defined tolerance.

Going forward

In this project, we learned that the biggest hurdles to establishing a legal basis for the processing of personal data in connection with machine learning arise in the continual learning phase. If an organization wants to use machine learning to solve the problem of not all relevant documentation being archived and entered into the record in a timely fashion, the sandbox project recommends that new archive legislation be implemented to responsibly regulate this issue.

The participants and the Data Protection Authority have had an opportunity to discuss practical challenges and potential measures to promote data protection by design in connection with procurement of solutions based on machine learning.

By taking a deep dive into the rules for data protection by design, Simplifai has gained a better understanding of the kinds of requirements they may be faced with from potential buyers of DAM. NVE, on the other hand, has gained a better understanding of the responsibilities data controllers have and how it can define requirements for solutions in a procurement process. The Data Protection Authority has learned more about the need for guidance actors have and how it can continue working to provide specific recommendations and examples.

The sandbox project recognizes the need for more actors to come together to build competence in the public sector and prepare practical examples of requirements public-sector bodies can include in competitive tenders for technical solutions. We would like to see templates setting out requirements for data protection by design in digital tools similar to those available for provider qualifications at anskaffelser.no. Such examples in template form could also provide increased predictability for providers.

If the proposal for a European Artificial Intelligence Act[2] is adopted, there will be less need to use contracts to regulate product requirements. This proposal includes requirements for selected products that use machine-learning technology, not just for product users.

[1] EUR-Lex - 52021PC0206 - EN - EUR-Lex (europa.eu)

The Norwegian Data Protection Authority's regulatory sandbox for responsible artificial intelligence

Office address:
Trelastgata 3, Oslo

Postal address:
PO Box 458 Sentrum
0105 OSLO

sandkasse@datatilsynet.no
Phone: +47 22 39 69 00

datatilsynet.no/sandkasse
personvernbloggen.no
twitter.com/datatilsynet