Image: Ruter AS/Nucleus AS, Magnus W. Sitter

On track with artificial intelligence

Exit report from the Ruter project

Theme: Transparency and explainability

January 2023

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.

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.

Summary

The public transport undertaking Ruter wants to use artificial intelligence (AI) to provide personalised travel suggestions to their customers in the Ruter app. The travel suggestions will be generated through the use of AI. The desired effect is to increase the use of public transport, micromobility, cycling and walking, which in turn may contribute to achieving climate and environmental goals, including zero growth in private vehicle use.

To ensure that Ruter remains an attractive provider of mobility services in an increasingly more competitive market, further development of the service will be necessary. At the same time, Ruter is reliant on the trust of the general population. Ruter therefore has the goal that any use of the users' personal data by AI must be responsible and fair.

Together with greater personalisation of the digital products, Ruter needs to provide clear, understandable and user-friendly information about their services. This sandbox project will explore how further development can take place in a manner which ensures that there is transparency and trust associated with the development and use of AI.

Conclusions

- **Transparency during the development phase:** For Ruter, it is crucial that customers are confident that their personal data will be processed in a responsible manner. Customer willingness to share personal data is a prerequisite for the development of the AI service. For consent to be valid, customers must also have an insight into and understanding of what they are agreeing to, as well as the option of being able to withdraw their consent. Ruter must therefore ensure they provide adequate information, including about how the AI service arrives at its travel suggestions, without this being overly complicated to understand. In Ruter's service, layered information is a good solution for safeguarding both considerations.
- **Purpose limitation**: Ruter wishes to further develop both the AI service in particular and other services in general by using personal data collected through customers' use of the Ruter app. Ruter needs to clearly define what constitutes the original purpose of the data's collection, and what are separate purposes which they already know when the data is collected that they will want to use the personal data for. If Ruter subsequently sees a need to use the personal data for new, unforeseen purposes, they will have to assess whether the new purposes will be compatible with the initial purpose.
- **Transparency during the usage phase:** It is also crucial when the AI service is to be rolled out that there is good, layered information in order to safeguard customer trust and ensure there is valid consent. The data subjects must be informed of each specific purpose prior to the processing of personal data taking place, and consent must be obtained for each new purpose. Transparency will be important for trust in connection with the use of the AI service.

Going forward

The discussions in the sandbox project have contributed towards more clearly defining the requirements for transparency which Ruter has to follow when developing and using AI. The assessments in this report are also relevant to other developers that wish to ensure transparency in their AI solutions.

Ruter has seen that the discussions concerning transparency, purpose and responsibility in the sandbox project are also relevant to other projects that they are working on. They therefore wish to ensure that this knowledge is transferred to other parts of their operations. They will continue to explore the use of AI when they consider that this can make a positive improvement to the customer experience or enable the service to operate more efficiently. The experiences from the sandbox project relating to transparency, purpose limitation and obligation to provide information make the company better equipped to ensure that these services are developed in accordance with the regulations, and customer rights and expectations.

Read more about the road ahead in the final chapter.

About Ruter's project

Ruter is a public enterprise that is owned by the City of Oslo and Viken County Council, which plans, coordinates, orders and markets public transport in those areas. One of Ruter's primary services is the Ruter app, which allows customers to plan trips, view departure times in real time, filter means of transport and purchase tickets.

What is the problem that Ruter wishes to solve?

Ruter is planning a new service that will provide customers with more personalised and specific travel suggestions using artificial intelligence based on their usage history. As of the start of this sandbox project, the solution is in the concept phase.

Ruter emphasises in their strategy documents that the current transport solutions are facing major changes. These changes are driven by a desire for more sustainable solutions, technological development, new business models and changing customer expectations. Ruter needs to adapt to these changes in order to continue to offer a good and stable transport service in the capital city region.

The Ruter app is one of the most important tools for customers when travelling by public transport. A good experience with the app can contribute to higher customer satisfaction. In order for the suggestions to be adequately personalised, Ruter believes that the AI model needs to learn from the customers' personal data.

The development of AI services based on customers' personal data is reliant on their willingness to share their personal data and to trust that Ruter will do a good job of protecting their privacy. Ruter therefore has a clear goal that, together with the personalisation of the digital products, they have to provide clear, understandable and user-friendly information about the services.

Ruter currently finds that they enjoy a high level of trust among customers. Transparency concerning the use of customers' personal data is vital for safeguarding this trust and for encouraging them to use new services. Transparency is also a basic requirement in the data protection regulations when businesses adopt solutions that process personal data.

In this sandbox project, we take a closer look at what requirements are set for transparency when using personal travel information to develop a service that uses AI to provide personalised travel suggestions.

Transparency and explainability

Transparency and explainability have been topics of previous sandbox projects. In the project with the Norwegian Labour and Welfare Administration (NAV), we discussed what a meaningful explanation of AI models looks like and how the explanation needs to be adapted to the intended target group.

Read more in the project's exit report here

The Norwegian Data Protection Authority has also written a general report on how to succeed with transparency in connection with the use of AI. The report is based on experiences from several previous sandbox projects.

Read the transparency report here

How will Ruter use artificial intelligence?

Ruter wants to use artificial intelligence and machine learning to learn from customers' personal data in order to offer personalised travel suggestions. The travel suggestions will be based both on the customer's own travel patterns, and where others who are located at the same place as the customer typically travel to.

The solution involves three phases:

- During the **preparation phase,** there must be a function in the Ruter app that collects personal data for the development and training of an AI model. The data will be temporarily stored on the customer's device (client). No development will take place as long as the data is stored on the customer's device. The purpose of the temporary storage is to enable Ruter to collect data points at an earlier stage of the process.

- During the **development phase**, the personal data will be sent to, stored by and used in Ruter's central systems. The purpose of this is continued development and training of an AI model. The transfer to Ruter's central systems must be based on the consent of customers who are logged-in.
- During the **usage phase**, the service with personalised travel suggestions will be launched for customers in the Ruter app if they are logged-in and consent to this. The development of the service will continue centrally at Ruter in parallel with the usage phase. Since travel patterns are constantly changing, the solution will be further developed to enable the model to be correct, relevant and up-to-date at all times.

The personal data that will be used is where and when the customer opens the app, the travel searches that are made and location. Previous searches in the app, both by the specific user and other users, will be used to train the AI model to provide better travel suggestions.

In connection with this, we would also mention that Ruter may encounter situations where they process special categories of personal data (also known as *sensitive personal data*). We will return to this below.

Basic presentation of Ruter's solution during the development and usage phases

Ruter was still in the concept phase of its solution as of the date this report was prepared. Changes may therefore have occurred when Ruter potentially launches the solution.

The development phase:

During the development phase, personal data will be collected for training the AI model. The purpose is to develop an AI model that can generate relevant, good-quality, personalised travel suggestions.

1. Ruter registers a selection of the customer's actions in the Ruter app. This includes searches performed and travel suggestions displayed.

2. Data relating to the customer's actions is sent to a "back-end" system over an encrypted line, where the customer's identity is replaced with a pseudonymised number.

3. The back-end system encrypts the customer's data and sends this to an internal sharing platform.

4. Ruter then sends the data to a server where the data is processed and analysed for it to be suitable for further use. The personal data is temporarily decrypted during this processing stage.

5. The data is sent to a machine learning platform where the AI model "learns" from the data. The data is decrypted while learning is in progress.

6. After the AI model is fully trained and validated, it generates travel suggestions based on the personal data it receives.

Visual presentation of the development phase
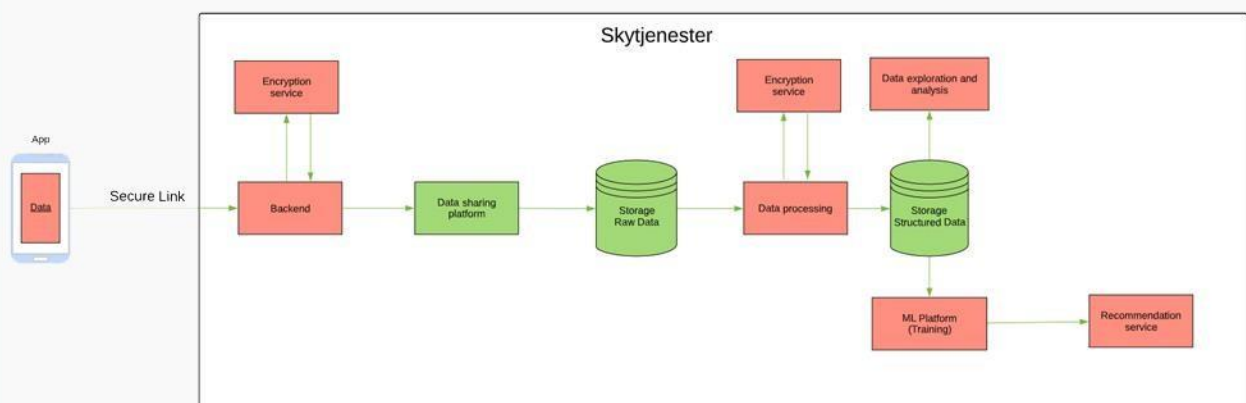
Steps two to six in the description are carried out in a cloud service. In the figure illustrating the development phase, data will be encrypted where it is green, while data must be decrypted temporarily in the red components.

As an alternative to the described solution, Ruter is considering carrying out the data flow in the development phase without using a common platform (cf. section 3 above), in addition to changing how the data is to be encrypted.

The development phase will be ongoing, because the AI model will need to continuously collect new information in order to learn and improve its travel suggestions.

Usage phase:

During the usage phase, personal data will be used to predict travel wishes and retrain the AI model. The purpose of data processing in the usage phase is to generate personalised travel suggestions in the Ruter app.

1. Customers open the Ruter app when planning a trip.

2. The app sends a request to the back-end system with the user's location. The customer's identity is replaced with a pseudonymised number before it is forwarded on.

3. The AI model sends travel suggestions back to the back-end system based on where the customer was and when the request was sent to the back-end system.

4. The back-end system sends the travel suggestion to the Ruter app, where the suggestion is presented to the customer.

5. The personal data (pseudonymised) is both stored in the back-end system and used for training the AI model.

Visual presentation of the usage phase
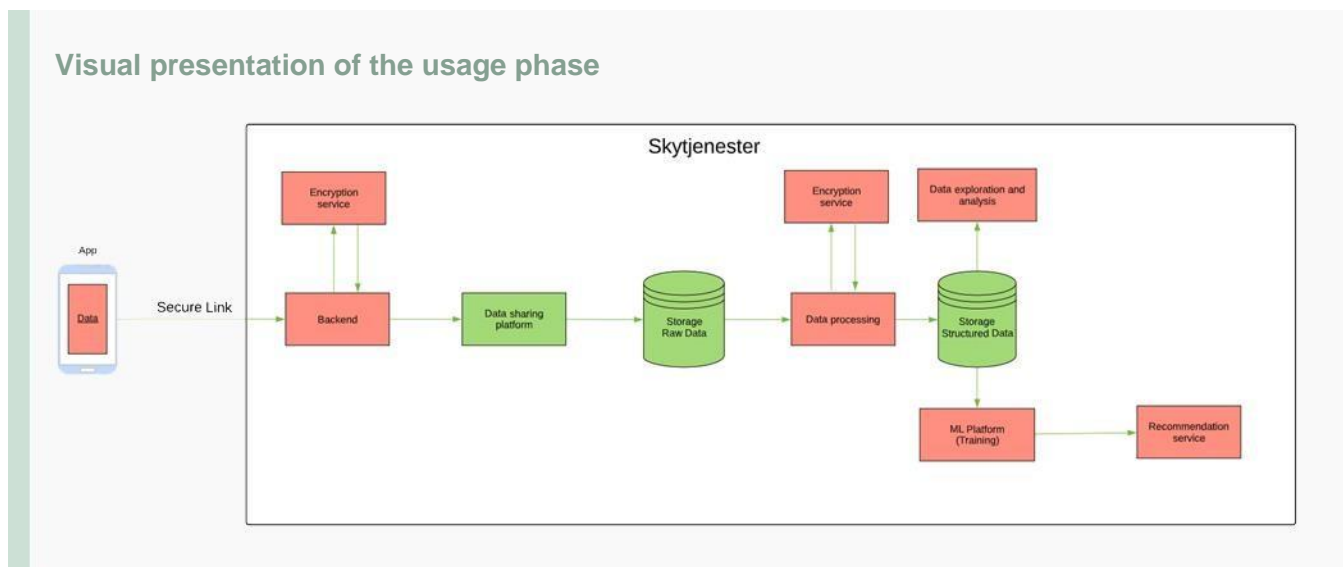
Objectives for the sandbox process

The Norwegian Data Protection Authority and Ruter have identified two primary objectives that are related to potential data protection challenges in the *development* and *use* of Ruter's AI model.

Objective 1 - Development phase: Investigate the requirements for transparency when using personal travel data to develop artificial intelligence

Delivery 1.1: Clarify what triggers the data subjects' right to information regarding how their personal data is processed in connection with the development of artificial intelligence.

It will take time from the start of local collection of personal data until the first version of the service can be presented to users. An important question is what triggers the obligation to provide information. Among other things, it must be clarified as to when data pertaining to an identified or identifiable natural person is processed in accordance with the General Data Protection Regulation (GDPR), and when Ruter's responsibility for processing enters into force.

Delivery 1.2: Clarify the data subjects' right to information in connection with the development of the AI model.

Ruter wishes to be transparent about how the customers' personal data will be used in connection with the development of the new functionality. To ensure transparency, it is important to be able to explain the purposes for the processing and how the personal data will be processed during the development of AI. Key questions are how much needs to be explained, and how to provide a good explanation before knowing what the AI model will become.

Delivery 1.3: Identify issues relating to the requirements for information in connection with the development of the artificial intelligence when consent is used as a legal basis.

Certain special information requirements apply when consent is used as a legal basis. We want to identify what information is required to ensure that consent is valid.

Objective 2 - Usage phase: Investigate the requirements for transparency when using artificial intelligence on personal travel data

Delivery 2.1: Clarify the various purposes for the processing of personal data during the usage phase, and what comes under the same purpose.

The purposes must be clearly stated even before the processing of personal data begins. We want to take a closer look at how the purposes in the usage phase can be specified. It can be difficult to define a clear purpose for the customer when Ruter themselves are still unaware of the full potential of the use of the personal data and machine learning.

Delivery 2.2: Clarify the data subjects' right to information during the usage phase.

When using AI, Ruter also wishes to be transparent about how customers' personal data will be used. A key question is how to provide a simple and concise explanation to customers while at the same time providing an adequate description of what is taking place. How much of the logic behind the AI models must and should be disclosed when they are used for personal data? Another question is whether the AI model needs to be adapted to safeguard the various information requirements.

Delivery 2.3: Identify issues relating to the requirements for information during the usage phase when consent is used as a legal basis.

The AI model will have to be further developed in parallel with customers using it in the Ruter app. We want to investigate what information a statement of consent will need to contain in order to ensure valid consent when the model is still under development.

The legality of processing personal data in the AI solution

In order for the processing of personal data to be lawful, the controller must always have a legal basis for such processing. Article 6(1)(a–f) of the GDPR contains an exhaustive list of six alternative legal bases for the lawful processing of personal data.

Assessments of the legality of the processing of personal data in Ruter's AI solution are not part of this sandbox project. The discussions in the report therefore assume that Ruter has a legal basis for the appurtenant processing activities.

However, the legal basis that Ruter selects will still influence what information they are obligated to provide.

Ruter plans to use consent (Article 6 (1)(a)) as the legal basis. This applies both to the use of personal data to train the AI model, and to the use of the AI model on personal data during the usage phase.

Consent as a legal basis

The personal data can be lawfully processed based on consent when such consent is:

- **freely given**
- **specific**
- **informed**
- **unambiguous**
- **given by a clear affirmative act**
- **able to be documented**
- **possible to withdraw as easily as it was given**

Read more about what each point entails here.

The duration of the consent will depend on what one has been asked to consent to. To avoid any doubt, the intended duration of the consent should be specified when such consent is requested. The data subjects should also be reminded at regular intervals that they have given consent and that this can be withdrawn.

One of the conditions for consent to be valid is that the consent is *informed*. It is therefore natural to further examine this requirement when we consider what information Ruter has to provide to customers who choose to consent to the processing of their personal data.

In addition to being *informed*, there are several other conditions that need to be met for consent to be valid. In this sandbox project, we have only taken a closer look at the requirements for informed consent.

Consent for special categories of personal data

Special categories of personal data are often referred to as *sensitive personal data*. This is data that requires extra protection, such as data relating to ethnic origin, religion, medical information, sexual orientation, etc. In principle, the processing of special categories of personal data is prohibited. Exceptions to the prohibition can be made through explicit consent. Read more about what makes consent explicit in section 4 of the European Data Protection Board (EDPB) guidelines relating to consent.

However, an exception cannot be made in instances in which it stipulated by law or regulation that the *data subject* cannot lift the prohibition.

Read more about the use of special categories of personal data.

Responsibility for local storage

In this sandbox project, Ruter will be the controller for the service which provides personalised travel suggestions in the Ruter app. But what responsibility does Ruter have for local storage during the preparation phase? In the project, we discussed how far this responsibility extends. This is decisive for determining when the obligation to provide information comes into effect. The question has arisen in connection with personal data that Ruter will not have access to.

As a data minimisation measure, Ruter will facilitate the local storage of data on customers' devices during the preparation phase. The purpose of local storage in this project is for Ruter to be able to use the travel data to develop AI at a later stage if the customer consents to this during the development phase. A log, which includes personal data, can then be sent to Ruter's centre for AI development at regular intervals.

While personal data is stored locally on customers' devices, Ruter will not have access to this. For customers who do not consent to central storage, Ruter will never have access to the data. The customer will also be able to delete favourite trips directly in the app. The customer will be able to delete all other data by reinstalling the app, or resetting the phone to factory settings.

Do the data protection regulations apply to local storage?

And if so, what role will Ruter play?

The starting point is that the data protection regulations will apply when processing personal data, cf. Section 2, subsection 1 of the Norwegian Personal Data Act and Article 2 (1) of the GDPR. The first question is whether personal data is processed in accordance with these provisions. The Norwegian Data Protection Authority and Ruter have differing opinions with regard to this. The Norwegian Data Protection Authority is of the opinion that personal data is already processed in connection with local storage. Ruter is of the opinion that this cannot be viewed as processing of personal data *in relation to* Ruter. Such an interpretation would mean that the data protection regulations will not apply, and that, under the regulations, Ruter will not have responsibility during the preparation phase. However, in the sandbox project we nevertheless further examined the other assessments that need to be made when the Norwegian Data Protection Authority's understanding is used as a basis.

Even if one accepts that personal data is being processed, the regulations do not always apply. There are exceptions that apply if, among other things, the processing of personal data is carried out by a natural person in the course of a purely personal or household activity, cf. Section 2, subsection 2 (a) of the Norwegian Personal Data Act and Article 2(2) (c) of the GDPR. Pursuant to recital 18 of the GDPR, such activities could include correspondence and the holding of address books, or social networking and online activity undertaken within the context of such activities. However, these rules do apply to controllers or processors that provide the means for processing personal data for such personal or household activities. The customer's own processing of personal data in the Ruter app, in the form of, for example, saving and deleting useful travel searches, will be an activity that falls outside the regulations. In instances such as this, where Ruter does not have access to the personal data, it will also be possible to make exceptions to several of the obligations under the GDPR.

An undertaking is responsible for processing when it determines the purpose of the processing and the means that will be used, see the fact box.

> ### Processing responsibility
>
> A controller is the party that determines the purpose (i.e. **why**) and means (i.e. **how**), of processing personal data, cf. Article 4 (7) of the GDPR.
>
> The controller has overall responsibility for compliance with the data protection principles and the regulations. This follows from the principle of accountability in Article 5 of the GDPR.

The purpose of storing personal data locally is that the customer will later be able to consent to personal data being sent to Ruter centrally for the development of AI. With regard to this storage, in our discussions we arrived at the conclusion that Ruter determines the purpose and means that are to be used. Therefore, for this part of the processing, Ruter will have the role of controller.

Irrespective of whether or not the regulations apply, Ruter plans to initiate measures which involve them fulfilling the responsibility that is incumbent upon the controller for this activity. Several of the exceptions to the duties of the controller are nevertheless relevant in the preparation phase, because Ruter does not have access to the personal data. For example, one does not need to fulfil all the rights of the data subjects if it can be demonstrated that the data subject cannot be identified, cf. Articles 11(2) and 12(2) of the GDPR. Read more about the requirements set for the controller here. For the sandbox project, it is the information requirements that are relevant. Ruter plans to provide information to customers even before the travel data will be stored locally. In the following chapters, we will further examine how.

The discussions in the sandbox project regarding when the responsibility for providing information becomes applicable have revealed some possible paradoxes. Ruter has chosen to use local storage as a data minimisation measure in the preparation phase. If the data protection regulations become applicable during this phase – which is the Norwegian Data Protection Authority's understanding – a potential consequence is that Ruter will have to ask the customer for access to more personal data than Ruter originally requested. If a customer asks to have a right fulfilled, for example, data portability, Ruter needs to gain access to the personal data in order to fulfil the obligation. Issues relating to responsibility for local storage may be relevant for many actors, irrespective of industry and whether they use AI. The Norwegian Data Protection Authority views this form of data minimisation as positive, and wants to contribute to actors being able to comply with the regulations in a simple and appropriate manner if they select this measure.

Responsibility in accordance with other regulations

In the sandbox project, we have only discussed the responsibility that follows from the personal data regulations. Other legislation can also stipulate obligations for Ruter, for example, the Norwegian Electronic Communications Act. This Act sets conditions for being able to store and gain access to data on the customer's equipment. However, examining the responsibilities that Ruter has under other legislation falls outside the scope of this project.

General information relating to the requirements for information and artificial intelligence

The GDPR requires that all processing of personal data shall take place in a manner that is legal, transparent and fair. When an undertaking collects and processes personal data, it pledges to provide the data subject with information regarding such processing. Use of artificial intelligence gives rise to certain particular issues concerning the information that must be provided to the data subject, because it is not always clear as to how an AI model has arrived at a result.

In this chapter, we will provide an overview of the general requirements for information in such instances. In the following chapters we will look at the requirements in the context of Ruter's project.

Feel free to read more about requirements for transparency in AI solutions in the report Artificial Intelligence and Privacy (2018).

Transparency and explainability

Transparency is a fundamental principle in the GDPR. In addition to being a prerequisite for uncovering errors, discriminatory treatment or other problematic issues, it contributes to increased confidence and places the individual in a position to be able to assert their rights and safeguard their interests. Transparency can also be of major value to the controller in order to create trust and to encourage customers to use new and complex technology.

The concept of 'explainability' is often used in connection with AI. This can be said to be a specific aspect of the transparency principle. Traditionally, transparency has been about showing how different items of personal data are used. However, the use of AI may require a different approach to explaining complex models in an understandable manner.

Explainability is an interesting topic, both because explaining complex systems can be challenging and because the way in which the requirement for transparency is be implemented in practice will vary from solution to solution. In addition, machine learning models can permit explanations that appear different to those we are used to, often based on advanced mathematical and statistical models. This opens the way for an important trade-off between a more correct, technical explanation or a less correct, but more understandable explanation.

Transparency requirement

Regardless of whether or not you use artificial intelligence, there are certain requirements for transparency when processing personal data. Briefly summarised, these requirements are:

- The data subjects must receive information on how the data is used, depending on whether the data is obtained from *the data subject* themselves or from others. We will discuss this further below. (See Articles 13 and 14 of the GDPR).

- The information must be easily accessible, for example on a website or via an application, and be written in clear and intelligible language. (See Article 12 of the GDPR).

- The data subject has the right to know whether data about them is being processed and have access to their own data if requested. (See Article 15 of the GDPR).

- It is a fundamental requirement that all processing of personal data must be done in a transparent manner. This means that an assessment must be carried out of what transparency measures are necessary for the data subject to be able to safeguard their own rights. (See Article 5 of the GDPR).

The first bullet point includes the contact details of the controller (in this case Ruter), the purpose of the processing and which categories of personal data will be processed. This is information that is typically provided in the privacy policy.

The GDPR requires that information provided to the data subjects is intelligible. It is therefore important to present the information in a manner that is clear and concise. A good means of achieving this is to provide the information in multiple layers, i.e. that one can click to obtain more information regarding specific topics. This avoids too much information being combined onto one page.

At the same time, having too many layers makes it more difficult to access the content. It is important that the information does not become overly fragmented and difficult to keep track of.

Automated decision-making

If processing can be categorised as automated decision-making or profiling according to Article 22 of the GDPR, there are additional requirements for transparency. This includes the right to know whether you are the subject of automated decision-making, including profiling. There is also a specific requirement that the individual is provided with relevant information concerning the underlying logic and the significance and the envisaged consequences of such processing.

The stronger requirement for transparency in connection with automated decision-making or profiling pursuant to Article 22 is referred to in:

- Article 13(2)(f) of the GDPR
- Article 14(2)(g) of the GDPR
- The European Data Protection Board's (EDPB) guidelines on automated individual decisions and profiling

As we will return to below, additional requirements for transparency may also apply in connection with profiling that does not come under Article 22.

Requirements for information when obtaining consent

The information that is provided to the data subjects when obtaining consent may have an impact on the validity of the consent. As mentioned, the requirement that consent is *informed* is one of several conditions that must be met for the consent to be valid.

Ruter plans to use consent as the legal basis for processing personal data when developing the AI solution. Therefore, in the report, we further examine the requirements that are set for the information in order for the consent that is granted to be informed.

The requirement for informed consent is largely linked to the right to information as described in Articles 12–14 of the GDPR. There are some additional requirements when the processing is based on consent. If sufficient information is not provided to the data subjects, the consequence will be that the consent is invalid.

What information has to be provided in connection with obtaining consent?

The data protection regulations require that, at the same time as the request for consent is made, information must also be provided relating to:

- The right to withdraw consent.
- The identity of the controller.
- The purposes of each processing activity for which the personal data will be used.

In addition, on pages 15–16 of its guidelines concerning consent, the EDPB has listed the following minimum requirements for the content of information for consent to be deemed informed:

- information about the type of personal data that will be used
- information about the use of personal data for potential automated decision-making
- information about the possible risk of transferring personal data out of the EEA without an adequacy

decision or necessary guarantees pursuant to Articles 45 and 46 of the GDPR.

The extent and type of information required for consent to be informed will vary. In some instances, it is necessary to provide additional information to what is mentioned above. The decisive factor is that the information will provide the data subjects with a genuine understanding of what they are consenting to.

What needs to be explained in connection with the development of the artificial intelligence?

During the development phase, Ruter will first collect data relating to how the data subject uses the app, what trips they search for, and where they are when the searches are carried out. This personal data is then sent to Ruter for training and development of an AI model for personalised travel suggestions, if the data subject consents to this. The customer's identity is replaced here with a pseudonymised number. The transfer of the personal data from the user's device to Ruter first takes place when Ruter has reached the stage in the development phase when the AI model is to be trained using personal data. It is uncertain as to how long it will take between the collection and storage of data on the user's local device and until the user is given the choice to share the information with Ruter.

Article 13 of the GDPR requires that information be provided to the data subject no later than at the point in time when the personal data is collected. This is solved by Ruter providing information to the data subject *before* such data collection commences. There may be nuanced differences in what information Ruter needs to provide at the time of local collection and central collection. We will return to this and to specific issues relating to informed consent.

Article 13 contains a long list of information that must be provided to the data subject. Here we will focus on Delivery 1.2, and the data subjects' right to information in connection with the development of artificial intelligence. Of greatest importance are the more complex parts of the obligation to provide information, i.e. *how* Ruter processes personal data in connection with the development of the AI model and for what purposes.

Obligation to provide information, AI and profiling

General information to the data subjects during the development phase

Ruter's project is in the preparation phase, and an important question is: How can Ruter provide the customers with the information they are entitled to when the solution has not been fully developed?

As mentioned in the introduction, the purpose of collecting and storing personal data locally is that personal data can later be sent to Ruter centrally for the development of AI. Ruter must therefore:

1. Provide information during the preparation phase that the personal data will be stored until the AI model is ready.

2. Provide new information during the development phase when the personal data is sent to Ruter centrally and the training of the AI model will commence.

During the preparation phase, it is easy to understand how Ruter will collect personal data and store this locally on the user's device. This phase will not involve AI. It is therefore a simple task to provide the data subject with information about this. The challenge is to comply with the requirement for providing information about the planned development of the AI model centrally at Ruter.

We have come to the conclusion that the list relating to the planned processing of personal data in "About Ruter's project" above is a good starting point.

The obligation to provide information has to be understood in light of the specific risk for the data subjects

We have discussed the considerations behind the transparency principle and how these influence the information that Ruter is obligated to provide to the users. Among the considerations behind the transparency principle is that data subjects must be able to safeguard their rights under the GDPR. Another important consideration is that information has an important control function vis-à-vis the controller, in that they have an obligation to explain to those concerned how their personal data is being processed.

The information that is necessary in order to safeguard these considerations will depend on how invasive the processing is in terms of the rights and freedoms of the data subjects. AI models that control what government benefits one is entitled to, or whether one should have access to government services, are examples of invasive processing of personal data. More detailed information will then need to be provided to the data subjects than for processing activities that have fewer consequences for the data subjects.

In the sandbox project, we have discussed that Ruter's model has few consequences for the data subjects, and that Ruter's use of AI does not pose any particular risk to the rights and freedoms of the data subjects. At the same time, Ruter is a publicly owned company and the sole provider of the public transport services in its area. We have therefore also discussed that Ruter is reliant on providing good information and a good user experience to ensure trust and that people will want to use their services. This particularly applies when the company intends to adopt new technology.

Special categories of personal data

One issue we have discussed in the sandbox project is whether favourite searches or travel patterns to and from the same address over time can infer that there are special categories of personal data. For example, one can envisage that a customer will regularly travel to and from a religious community, health institution or a political organisation and that this may reveal one's religious faith, medical information or political affiliation.

Ruter has been clear that they do not intend to process special categories of personal data about the customers. We nevertheless agree that Ruter may process this type of data when they collect location data and travel searches over time.

In areas with a large number of stops, such as Jernbanetorget in Oslo, a more accurate GPS position is required to provide correct and relevant travel suggestions than in areas with few stops. As a data minimisation measure, Ruter will reduce the accuracy in areas with fewer stops. This may, for example, be appropriate outside the town/city centre, where there is a longer distance between stops and where the frequency of detached houses makes it easier to identify the user based on location. If, on the other hand, the user is at Jernbanetorget, there will be a need for complete accuracy to determine which stops are closest. In areas where the GPS location is relatively accurate, there is a greater risk that the data collected will reveal journeys to and from, for example, a religious community or a political organisation. The same applies in instances in which customers themselves search for specific addresses.
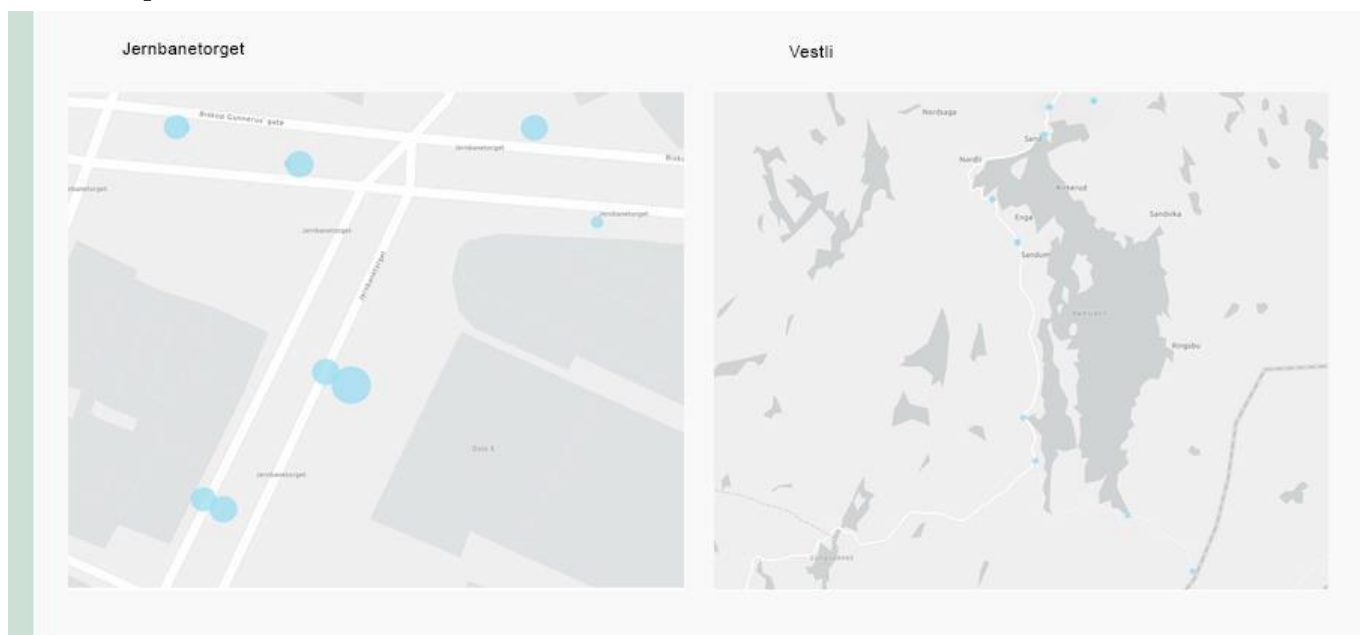
In the pictures above, the stops are marked to show distance and the various requirements for exact position.

Since there is a risk that Ruter will process special categories of data about customers, the customers must be informed of this to ensure they are aware of the risk.

Does Ruter's service involve "profiling"?

An important question for Ruter has therefore been whether personalised travel suggestions will constitute profiling pursuant to the GDPR and whether this triggers an additional obligation to provide information to data subjects.

> ### The definition of profiling in the GDPR
>
> "any form of automated processing of personal data consisting of the use of personal data to evaluate certain personal aspects relating to a natural person, in particular to analyse or predict aspects concerning that natural person's performance at work, economic situation, health, personal preferences, interests, reliability, behaviour, location or movements".
>
> (Article 4 of the GDPR)

In Ruter's service, the company will use personal data to predict travel patterns and provide personalised travel suggestions. Our assessment is that this involves automated processing that uses personal data to analyse and predict a natural person's behaviour, location and movement. Ruter's use of AI therefore constitutes profiling pursuant to Article 4 (4) of the GDPR.

In the sandbox project, we have discussed whether profiling takes place in the development phase, because the user does not receive any travel suggestions during this phase. We have arrived at the joint conclusion that Ruter will test and validate the model using personal data during the development phase. This means that profiling will therefore also take place during the development phase.

Information about profiling that is not covered by Article 22

> *"The data subject shall have the right not to be subject to a decision based solely on automated processing, including profiling, which produces legal effects concerning him or her or similarly affects him or her."*
>
> (Article 22 (1) of the GDPR)

In the sandbox project, we have concluded that Ruter's personalised travel suggestions will not have such an effect for the data subjects. Ruter's profiling is therefore not covered by Article 22.

The next question is what information Ruter is still obligated to provide regarding the profiling. Profiling is the processing of personal data, and needs to satisfy the general requirements for information in Articles 12 and 13 of the GDPR. In addition, we have discussed whether, when read in light of the recital, it can be inferred from the transparency principle that Ruter has an obligation to provide information about how the AI model functions.

Pursuant to Article 5(1)(a) of the GDPR, the controller must ensure that personal data is processed in a fair and transparent manner. Furthermore, recital 60 of the GDPR states that:

> *"The principles of fair and transparent processing require that the data subject be informed of the existence of the processing operation and its purposes. The controller should provide the data subject with any further information necessary to ensure fair and transparent processing taking into account the specific circumstances and context in which the personal data are processed. Furthermore, the data subject should be informed of the existence of profiling and the consequences of such profiling."*
>
> (Our emphasis)

Pursuant to the Article 29 Working Party's guidelines for transparency (paragraph 41), a fundamental consideration is that data subjects are able to understand in advance the scope and consequences of the processing of their personal data.

The Article 29 Working Party has also issued a statement regarding information that must be provided for profiling that falls outside of Article 22:

> *"It should be noted that, aside from the specific transparency requirements applicable to automated decision-making under Articles 13.2(f) and 14.2(g), the comments in these guidelines relating to the importance of informing data subjects as to the consequences of processing of their personal data, and the general principle that data subjects should not be taken by surprise by the processing of their personal data, equally apply to profiling generally (not just profiling which is captured by Article 22, as a type of processing."* (Paragraph 41)

A legal duty to explain the underlying logic behind the profiling discussed in this sandbox project can hardly be inferred from an expanded interpretation of Articles 13 and 14 of the GDPR. However, the principles of transparency and fairness in Article 5 (1) (a) still dictate that Ruter must provide information about the underlying logic, to the extent that is necessary for the data subjects to be able to understand how Ruter processes their personal data, and thus enabling them to exercise their rights. This is also in line with the opinions of the Article 29 Working Party that the processing must be predictable for the data subject.

In the sandbox project, we have concluded that Ruter needs to provide the data subjects with general information about the fact that they are being profiled, and relevant information about how the underlying profiling logic functions. Ruter is also focussed on providing good information about this to ensure satisfaction and trust among customers. The requirement for transparency does not necessarily mean that the source code has to be made available; however, the explanation must enable the data subject to understand why a particular decision was made.

Explanation of the underlying logic

The term "underlying logic" is used in the GDPR. This means a general explanation of how a result is arrived at. The term does not include an explanation of how the specific result that applies to you was arrived at.

When explaining the underlying logic, Ruter should strive to ensure that the information provided is meaningful, rather than using complicated explanatory models based on advanced mathematics and statistics. It is also emphasised in recital 58 of the GDPR that technological complexity makes transparency extra important.

The most important factor is that the data subjects understand *how* Ruter's service determines travel suggestions, how they are profiled and the consequences of this. As mentioned in the introduction, Ruter's planned service is not particularly invasive for the data subjects. This also influences how detailed the information Ruter has to provide to the data subject needs to be in order for them to be able to safeguard their rights and ensure transparency and predictability.

Ruter's challenge in the preparation phase, when the company will start collecting personal data, and then later in the development phase, is that it is not entirely clear what underlying logic will be used. It will depend on which models end up functioning the best. However, what is clear is that these AI models are still going to use location and travel searches with the associated points in time. Ruter must inform the data subjects that the profiling is based on these categories of personal data.

See examples of what should be included in an explanation on page 31 of the Article 29 Working Party's guidelines for automated, individual decisions and profiling.

Information relating to the transfer of data to third countries

Ruter is looking at two alternative solutions, one of which involves the transfer of data to third countries. The choice of solution will be based on Ruter's assessment of the transfer of personal data to third countries. The sandbox project did not address this. Even if the data flow should not involve a transfer, we agree that Ruter must provide information about the data flow to the data subjects and the assessments the company has made regarding transfers to third countries.

Since Ruter is using a cloud provider in this project, this could be considered a recipient of personal data pursuant to Article 4(9) of the GDPR. Ruter is obligated to provide information about recipients of personal data pursuant to Article 13(1)(e).

Information pertaining to where and how the personal data is processed is a prerequisite for the data subjects being able to exercise their rights. This is important information that the data subjects need to have in order to assess whether they wish to consent to the use of personal data.

We have discussed how it is challenging to provide information about the data flow and the assessments relating to transfers to third countries. Ruter needs to work more on this, however we agree that the simplified explanation and illustrations at the beginning of the report are a good starting point. We have also talked about how it is important that this information is presented to customers in layers to avoid the quantity of information overwhelming the reader. It is important to remember that the information must be easy for the customer to understand.

If Ruter concludes that personal data will be transferred out of the EEA, they are obligated to provide information about this and explain how Ruter does this in a lawful manner in accordance with Chapter V of the GDPR, cf. Article 13 (1)(f). The data subjects also need to be informed about where any necessary guarantees have been made available. Even if Ruter should come to the conclusion that they do not transfer personal data out of the EEA, the transparency principle dictates that Ruter should provide information about why this is their conclusion, and that they therefore do not need to satisfy the requirements in Chapter V of the GDPR relating to such transfers.

The information can, for example, be presented in layers under an overarching heading "We do not transfer your personal data out of the EEA - read more about this here".

Form – illustrations, video, layered information

Ruter plans to provide layered information, and is considering, among other things, using illustrations to explain the data flow and where personal data is transferred in the solution.

Read more about design in connection with software development.

We agree that the form has to serve a purpose, and that, for example, illustrations and video should not be used when this does not contribute to making the message easy to understand.

We have also discussed that, in some areas, it may be appropriate to have more than two layers of information. This can, for example, be done when Ruter has to explain more complicated topics, such as the underlying logic in the model or transfers to third countries. At the same time, it is important to limit the use of layers to what is appropriate, so that the information remains easily accessible and clear. For other and simpler topics, such as

storage time or contact information, it will not be appropriate to have as many layers. 19

Particular issues relating to requirements for information when obtaining consent

When collecting data for the development of the AI solution, Ruter will use consent as a legal basis. The information provided to the users must therefore be formulated in such a way that consent will be informed. If Ruter uses a different legal basis for the collection of data locally on users' devices during the preparation phase, these special requirements for information will not become applicable until consent for the development of the AI model is to be obtained.

The GDPR does not set any requirements for the form in which the information must be provided. This means that the information can be presented in various ways, including via video or audio recording. When consent is given in a written declaration, Article 7 (2) of the GDRP requires that:

> "[…] the request for consent shall be presented in a manner which is clearly distinguishable from […] other matters, in an intelligible and easily accessible form, using clear and plain language."

These requirements for form and language are closely associated with the transparency principle in Article 5 (1) (a) of the GDPR. The request for consent must be separate from other information and must be clear and concise. It cannot be inserted into general contractual terms and conditions, and it must be clear that the customer has given consent. The language must be intelligible for a normal customer. It should not be necessary for one to be familiar with difficult foreign words to be able to read the text.

Ruter is working with different proposals for how they can best inform the customer at different levels and obtain consent in the app as shown in the examples below. This can be user tested in order to obtain the best possible insight into how customers understand the message and what they are consenting to.

Examples of how Ruter's requests for consent could generally appear in the Ruter app

Ruter is planning to obtain consent through a button/tick-box that is displayed on the same page as the information provided to the data subjects in the first layer. The information page will appear in connection with an update of the app. There is also the option of notifying customers if an update is available. The first layer of information needs to satisfy both the minimum requirements for information in Articles 12 and 13 of the GDPR and the minimum requirements for informed consent.

In the sandbox project, the Norwegian Data Protection Authority and Ruter agreed that it is important to provide a simple but sufficiently detailed explanation for consent to be informed. It is crucial that the data subjects understand what they are consenting to. The development of AI can be difficult to understand. The need to provide an adequate explanation must therefore be balanced against the need for the information to be easy to understand.

Ruter has a service that is intended to reach *all* customers. The information therefore has to be communicated to a broad spectrum of people and it cannot be expected that the customers will have a complete understanding of what an AI model is. In order to understand what one is consenting to, a form of explanation is required.

In addition to information about the actual AI model, it is important that the users understand how their personal data will be processed during the storage and transfer periods. Ruter is considering preparing illustrations to provide a simple and concise description of the technical flow of data. An example of this is the illustration at the start of the report. An intelligible description of the flow of personal data is important for the data subjects to be able to assess whether this is something they wish to consent to. This will enable the data subjects to, for example, make their own assessments of whether they believe the personal data will be protected in a manner that inspires trust.

The special conditions that apply for a child's consent pursuant to Article 8 of the GDPR and Section 5 of the Norwegian Personal Data Act do not apply because Ruter has set an age limit of 15 years for being able to use the Ruter app. However, Ruter still needs to consider the fact that a 15-year-old may, for example, need information to be tailored differently to an 80-year-old.

One method of ensuring that the users actually understand the information is to conduct representative customer surveys. This is something that Ruter plans on doing.

What needs to be explained in connection with the use of the artificial intelligence?

The usage phase starts when Ruter considers the AI to have been sufficiently developed to be able to be used for predicting desired travel. The AI model will then be presented to the users of the Ruter app.

During the usage phase, Ruter plans to use personal data to predict travel patterns in order to provide travel suggestions, and for post-learning of the AI model. Ruter will also explore the option of using personal data during the usage phase for further development of the relevant service, and Ruter's services in general. The latter could, for example, include the use of statistics that can provide information for other service development, improve traffic planning, reveal which model in the Ruter app works best, and possible other forms of use.

The purpose during the usage phase

Providing information about the purposes of the intended processing of personal data is vital for compliance with the obligation to provide information. Ruter envisions that personal data collected in connection with the use of the AI solution can be useful for a multitude of purposes.

In the sandbox project, we have therefore discussed questions related to purpose limitation, including:
- What purposes does Ruter need to use personal data to fulfil?
- What falls under the same purpose?
- Which constructed examples of possible future purposes may be compatible with the initial purpose?

Personal data must only be collected for specified, explicit and legitimate purposes, cf. Article 5 (1) (b) of the GDPR. The purpose or purposes must be defined before the personal data is collected, and must be clearly communicated to the data subjects. The requirement relates to the transparency principle. The manner in which the personal data will be processed must be predictable for the data subjects. This enables them to have greater control over how their personal data is used. When Ruter uses consent as a legal basis for processing the personal data, it is also important for the validity of the consent that Ruter provides clear information about the various purposes and that separate consent is obtained for each separate purpose.

Which processing activities come under the same purpose?

At an overall level, Ruter envisions that these activities may become applicable during the usage phase:

- predict travel patterns to provide travel suggestions
- post-learning by the AI model
- further development of the relevant service
- further development of Ruter's services in general.

All these overarching activities can be divided into several, more specific processing activities. For some of the more specific processing activities, Ruter can already be certain when they collect the data that they wish to carry out these activities at some point in the usage phase. Other processing activities that may be beneficial in the long term are not possible to envisage at such an early stage. This particularly applies to the further development of the service and Ruter's services in general. We have therefore prepared some examples of new requirements that may arise in the future. In the sandbox project, we have discussed what processing activities may fall under the same purpose, and those which may be part of new purposes. We have also discussed whether or not new purposes may be compatible with initial purposes. Use of personal data for compatible purposes will be lawful, provided that Ruter has a legal basis for this.

The Article 29 Working Party, (which was the predecessor to the European Data Protection Board), wrote the following on page 16 of its opinion 03/2013 on purpose limitation:

> "For 'related' processing operations, the concept of an overall purpose, under whose umbrella a number of separate processing operations take place, can be useful. That said, controllers should avoid identifying only one broad purpose in order to justify various further processing activities which are in fact only remotely related to the actual initial purpose."

The opinion relates to the principle of purpose limitation in the previous data protection regulations. Since the principle of purpose limitation has been continued in the GDPR, the opinion can still provide guidance under the current regulations. A purpose can therefore be specific even if it encompasses several different processing activities that have a natural connection to the overall purpose.

Ruter has explained that distinguishing between providing travel suggestions and post-learning of the model serves no practical purpose. Travel patterns are constantly changing. For example, travel patterns were more static prior to the pandemic than what they are now when people have a more variable working day. Ruter wants to identify these kinds of patterns. Travel patterns are also seasonal, which means that there are major differences in movement patterns between winter and summer. The model therefore needs to be continuously adapted and improved in order to provide accurate travel suggestions. Ruter has noted that the quality of the product will deteriorate if the model does not learn along the way, and that part of the purpose of using AI would then be lost. The same applies to adaptations of the AI model that are not post-learning, for example adaptation of the extent to which the AI model should emphasise different elements, as well as the removal of unnecessary elements and errors. Ruter wants to make these needs clear to the customers.

In the sandbox project, we concluded that there could be a sufficient link between the use of personal data in the AI model to predict travel patterns and for post-learning of the model, to enable the processing activities to be considered the same purpose. An overall purpose of offering personal travel suggestions in the Ruter app can also be sufficiently specific. What is of decisive importance is that the processing activities that fall under the purpose need to have a sufficiently close connection. The connection can be close when it is not possible to achieve the purpose of a processing activity without adding an adjacent processing activity.

Page 53 of the Article 29 Working Group's Opinion 03/2013 on purpose limitation provides an example of how an overall purpose can often be broken down into several underlying purposes:

> "[...] - For example, processing an individual's claim for a social benefit could be 'broken down' into verifying his or her identity, carrying out various eligibility checks, checking other benefit agencies' records, etc.
>
> - The concept of an overall purpose, under whose umbrella a number of separate processing operations take place, can be useful. This concept can be used, for example, when providing a layered notice to the data subject. More general information can be provided in the first instance about the 'overall purpose', which can be complemented with further information. Breaking down the purposes is also necessary for the controller and those processing data on its behalf in order to apply the necessary data protection safeguards."

In purely practical terms, Ruter can provide information about the overall purpose in the first layer of information, while information about underlying purposes – such as post-learning and adaptation of the AI model – can be provided in another layer that the data subject can choose to access by clicking on a link.

Another question is whether further development of the specific service for personalised travel suggestions can be categorised under the same purpose. Further development of the service is a rather broad description. Multiple processing activities may be covered by this description. Ruter has explained that by further developing the service they wish to achieve two objectives:

1. To further develop to ensure the quality of the personalised travel suggestions, and
2. To further develop to provide added value beyond the personalised travel suggestions.

On the one hand, Ruter is continuously working to improve the service, and they claim that there is no point in presenting the product to customers without being able to further develop it. This type of further development could be called maintenance: It is about ensuring quality, not achieving something new. On the other hand, Ruter wants to further develop the service by adding new functions where they see this could provide added value. An envisaged example could be new integrations to guide the customer to make informed and efficient travel choices. Another could be new integrations to guide the customer towards making good ticket selections.
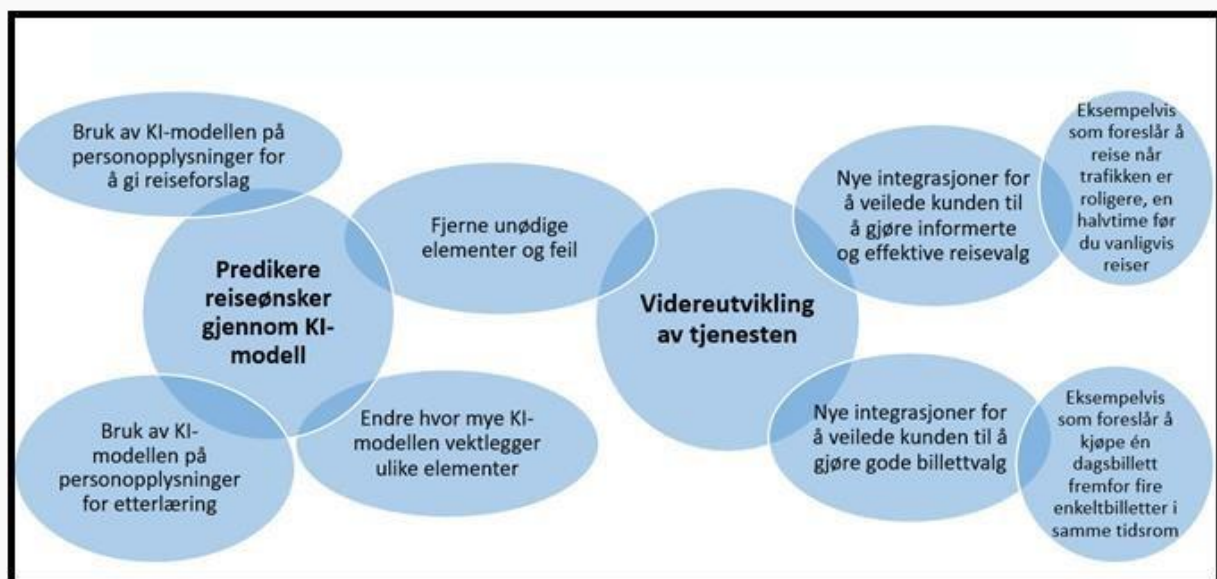
The illustration shows a proposal for the envisaged division of Ruter's overall purpose, underlying purpose and processing activities. There are smooth transitions between the various processing activities and it can be challenging to define what is covered by one specific purpose.

In order to assess what falls under the same purpose, we again need to look at the connection between the processing activities. This therefore pertains to all further development of the service that does not specifically apply to the AI model. Ruter has also explained that it is not possible to clearly state in advance which processing activities will be desirable in the long term. In order for the processing activities to be covered by the same purpose, there must be sufficient proximity to both the other processing activities and the overall purpose.

In the discussions, we arrived at the conclusion that further development covered by the maintenance category could come under the overall purpose. An example of this could be removing unnecessary elements and errors. The example of adding new integrations to guide customers towards making good travel choices is more difficult to consider as being pure maintenance. Such guidance could, for example, be a suggestion to take a more efficient route half an hour before you normally travel. We considered this to be on the borderline of what can be

characterised as the prediction of desired travel for the purpose of providing personalised travel suggestions. However, if a desire to effectuate similar integration arises, the example is of such a nature that it can be argued that it is covered by the initial purpose, or possibly that it is a new and compatible purpose. The assessment of whether a new purpose is compatible is only applicable when the purpose cannot be defined at the time the personal data is collected. If Ruter already finds, when the data is collected, that such an integration is desirable, they must assess whether it constitutes a new purpose before such collection takes place.

Another conceivable example of a new integration could be that Ruter guides the customer towards making good ticket selections, for example, by purchasing a 24-hour ticket rather than four separate tickets during the same period of time. We found that the latter probably falls outside the purpose of offering personalised travel suggestions. We also discussed whether such envisaged further development could be compatible with the initial purpose. Several elements have to be considered when making the assessment:

Compatible purposes

Pursuant to Article 6 (4) of the GDPR, when assessing whether a purpose is compatible with the purpose for which the personal data is initially collected, the following must be taken into account:

- any link between the purposes for which the personal data have been collected and the purposes of the intended further processing;
- the context in which the personal data have been collected, in particular regarding the relationship between data subjects and the controller;
- the nature of the personal data, in particular whether special categories of personal data are processed, or whether personal data related to criminal convictions and offences are processed;
- the possible consequences of the intended further processing for data subjects;
- the existence of appropriate safeguards, which may include encryption or pseudonymisation.

Compatible purposes

If the processing is predictable when the data is collected, or is a logical next step, this may indicate that the purpose is compatible. The more unpredictable that further processing will be, the more that is required to consider the purpose to be compatible.

It is important to look at how the purpose is perceived by the data subject when assessing foreseeability. On pages 24-25 of its Opinion 03/2013 on purpose limitation the Article 29 Working Party wrote that it is the content, and not the original choice of wording in the explanation of the purpose, that is decisive. The balance of power between the data subjects and the controller may also be of significance in the assessment. Technical and organisational measures may also be important. This is linked to the element relating to the potential consequences of the further processing for the data subjects, see fact box.

Using the personal data to analyse and guide ticket selection is not particularly far removed from the purpose of receiving personalised travel suggestions. However, it may come as a surprise to the data subjects who have consented to one type of analysis of their data to discover that it is also being used for another, different analysis. This argues against the purpose being considered compatible. In the discussions, we came to the conclusion that this example was borderline in terms of what can be considered compatible.

### Is further use of statistics compatible?

The Norwegian Data Protection Authority and Ruter also found that improving Ruter's other services would be unlikely to fall under the original purpose of offering personalised travel suggestions. Ruter particularly envisages that it may be applicable to generate statistics on the basis of the personal data, which in turn can:
- provide information for other service development,
- improve traffic planning, and

- reveal which model in the Ruter app works best.

Other further uses for the statistics may also be relevant for Ruter. However, it is difficult to predict exactly what use the statistics may be beneficial for in the future.

In the sandbox project we further examined whether these new purposes can be compatible with the initial purpose.

Overall purpose: Offer personalised travel suggestions
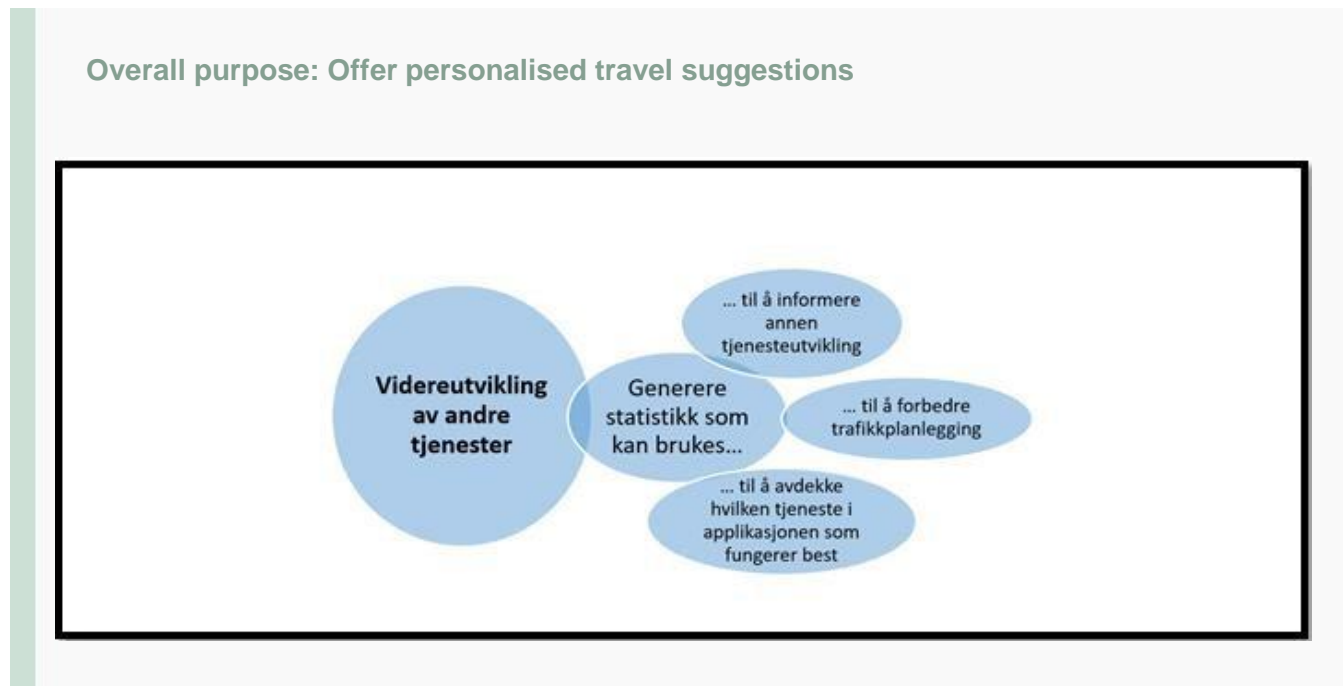
The illustration provides examples of hypothetical future purposes.

Pursuant to Article 5 (1) (b) of the GDPR, statistical purposes are not incompatible provided that the controller provides necessary guarantees for protecting the data subject's rights and freedoms, cf. Article 89 (1) of the GDPR.

In recital 162 of the GDPR, statistical purposes are described as: "any operation of collection and the processing of personal data necessary for statistical surveys or for the production of statistical results". The term encompasses a wide spectrum of processing activities, cf. page 29 of Opinion 03/2013 on purpose limitation. The use of statistics for both public and commercial purposes is covered. A commercial purpose may be the use of statistics for analysing websites or market research. Measures for protecting the personal data may, among other things, be anonymisation or pseudonymisation, including access control.
The measures must be viewed in connection with the principle of data minimisation. The data must be de-identified and protected to the extent that it is possible to still achieve the purpose.

Ruter is working on solutions for anonymising data for further use. At present, it is challenging for Ruter to achieve the desired purposes through continued internal use if they use true anonymisation. For external use, Ruter can and will anonymise the data. Read more about true anonymisation here. The statistics that Ruter wishes to use internally will still be processed in such a way that it may be difficult to derive personal data from these in a simple manner. The personal data will therefore be pseudonymised. In the sandbox project we discussed what is required to satisfy the condition of necessary guarantees.

Recital 162 of the GDPR states that: "Those statistical results may further be used for different purposes, including a scientific research purpose. The statistical purpose implies that the result of processing for statistical purposes is not personal data, but aggregate data, and that this result or the personal data are not used in support of measures or decisions regarding any particular natural person." In the discussions, we came to the conclusion that this most probably means that the statistics should not be used for purposes that require the re-identification of individuals. We also discussed whether the statistics can only be used for new purposes when personal data can

no longer be derived from the statistics. An affirmative answer to that question would appear to contradict the wording of Article 89 (1) of the GDPR. It is only when the purpose can be fulfilled by using anonymous data that the provision requires this, otherwise underlined data minimisation is required.

If personal data is reused for new purposes, information must be provided to the data subjects. Our conclusion is that Ruter should provide as much detail as possible about such reuse already at the point at which consent is given. Other information relating to new purposes must be provided no later than before further processing takes place, cf. Article 13 (3) of the GDPR. This will enable the data subjects to still have the opportunity to protect their rights. The new purposes must still be specified, explicit and legitimate, cf. Article 5 (1) (b) of the GDPR.

What information does Ruter have to provide during the usage phase?

There are a number of similarities in the development and usage phases when concerning what information Ruter has to provide and the manner in which this can be done. In the following, we further examine what specifically applies to the usage phase.

Provide information about profiling and how data is processed

As mentioned, Ruter's processing of personal data for the purpose of offering personalised travel suggestions will involve profiling. See descriptions in the previous chapter of what information must then be provided to the data subjects. Ruter does not know how the underlying logic in the fully trained AI model will function until the project is through the development phase. We therefore do not have specific underlying logic that we can identify in this report. However, we wish to provide an overall statement about what information has to be presented to the data subjects and how this can be presented by Ruter.

When the AI model is ready to be rolled out to users, it is important that Ruter has a good overview of what parameters the model emphasises in its profiling and how this takes place, such that the data subjects may receive information regarding how the model generally determines travel suggestions.

A topic that was also raised in the sandbox is whether the model needs to be changed in order to fulfil the data subjects' right to information. For example, does one have to select a model that is easier to explain in order to fulfil this requirement? We agree that no changes are required in Ruter's model to meet the requirement for information in this instance. However, we do not rule out that it may be necessary to meet other requirements in the regulations.

Information concerning the use of feedback in the AI model

Ruter wants to make it possible for customers to provide feedback on the travel suggestions through, for example, buttons with a thumbs up or thumbs down image. The company wants to use this feedback to adjust the AI model to enable it to provide more relevant travel suggestions and improve the service. In connection with this, we have discussed that Ruter has to provide information on how the feedback will be processed in a simple manner if the feedback constitutes personal data.

How can the information be provided to the data subjects

Ruter currently has an existing app with which the personalised travel suggestions will be integrated. The company therefore has the opportunity to test different methods of providing information in this solution. The information could, for example, be provided through pop-up windows.

We have discussed that a possible means of providing information about the underlying logic in an easy to understand manner could be to write a text such as the following in the privacy policy:

> *"How do we provide you with personalised travel suggestions?*
>
> *In the model, we use data about when and where you use our app, and what trips you search for at what time. The model also uses data about which trips other people in your area have searched for, at what time and where they were when they made the search. Based on this, the model calculates our*

current travel suggestions".

However, this type of wording must be adapted to the underlying logic when it becomes clear to Ruter how this will function. The wording also needs to clarify what personal data is used in the AI model.

It is also important that Ruter provides information in an intelligible manner about, for example, what profiling is and what it means for the customer, as well as the data flow and potential transfers of personal data out of the EEA.

Requirements for information when obtaining consent

Ruter plans to obtain consent for the usage phase through an information page, with a button/tick box which appears in connection with an update to the app. Like the development phase, the minimum requirements for informed consent will already need to have been satisfied in the first layer.

In the discussions, we discovered that the assessments for the development and usage phases will be relatively similar. Among other things, the information on the right to withdraw consent can be formulated in the same manner. What may be different in the usage phase will particularly relate to purpose limitation and further use for new purposes. As mentioned, it is important for the validity of the consent that the information clearly distinguishes between different purposes. Separate consent must be obtained for each new purpose.

The information provided to the data subjects has an impact on the assessments of purpose limitation. The information concerning purpose that is provided when the user's consent is obtained could, for example, influence what may be considered a compatible purpose, see the sub-chapter relating to purpose.

Going forward

Ruter has a vision of sustainable freedom of movement for everyone. In a market where customers are constantly expecting more transport options and a more individualised service, it will be important for Ruter that they own the preferred customer interface for mobility services in the region. Only then will Ruter be able to influence and ensure that the needs of individual customers also involve sustainable adaptation in terms of the environment, land use and accessibility for all.

Ruter's services are based on large network systems, which are constantly increasing in complexity as new forms of mobility and transport are added. When using artificial intelligence, one has an opportunity to adopt the use of new technology to utilise the opportunities and make use of this system more efficient. Better travel suggestions, which is the example we have worked with in the sandbox project, may result in more customers using Ruter's services. If artificial intelligence can also be used to more optimally utilise the capacity of the mobility system, it is possible to reduce costs, both financially and in terms of resources, and thus offer the general public a more sustainable transport service.

Ruter has seen that the discussions concerning transparency, purpose and responsibility in the sandbox project are also relevant to other projects that they are working on. They therefore wish to ensure that this knowledge is transferred to other parts of their operations. They will continue to explore the use of AI in their services when they consider that this can make a positive improvement to the customer experience or enable the service to operate more efficiently. The experiences from the sandbox project relating to transparency, purpose limitation and obligation to provide information make the company better equipped to ensure that these services are developed in accordance with the regulations, and customer rights and expectations.

For Ruter, trust is the foundation of their relationship with customers and the general public. It is crucial to retain this trust when developing new services that have not traditionally been part of the public transport system, or the use of new technology that many customers consider themselves unfamiliar with. Transparency is a fundamental prerequisite for trust. Good measures for safeguarding privacy, combined with a transparent and simple explanation, can lead to customers maintaining their trust in Ruter, and Ruter being preferred over similar services.

It is a goal of the Norwegian Data Protection Authority that this report contributes to providing more practical guidance, including for other enterprises that want to ensure transparency in their AI solutions. The assessments are particularly relevant for other actors that want to develop services for personalised recommendations based on the users' own data and behavioural patterns, both in the transport sector and in other areas. The discussions on how to provide information about the complicated processing of personal data in AI will also be relevant to anyone who uses complex technical solutions - both with and without AI.

Datatilsynet