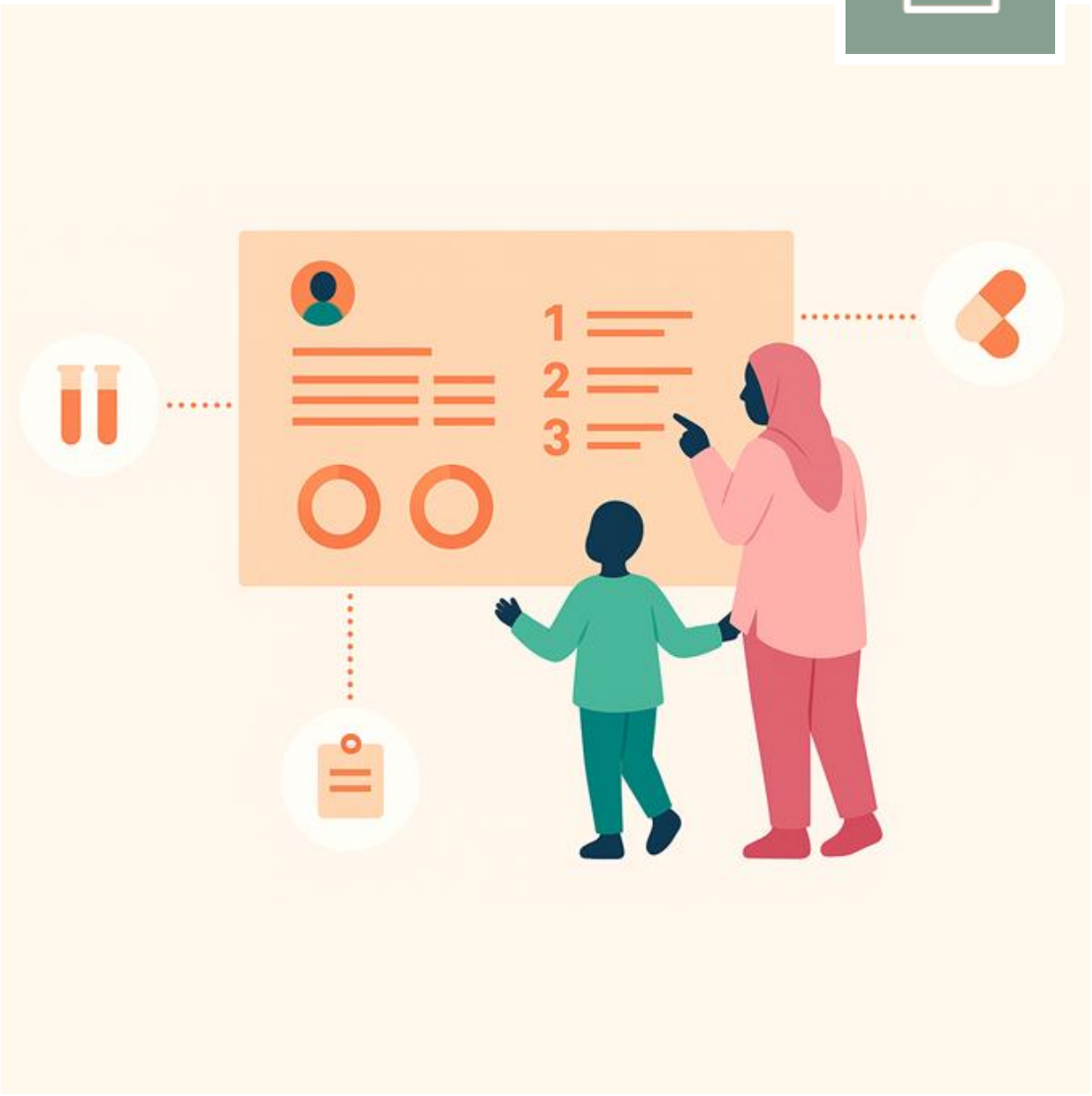

Privacy-friendly chatbots

Using artificial intelligence and language models to simplify access to health information

– Report from the sandbox project with the Norwegian Directorate of Health

Datatilsynet

Contents

Executive summary

The Norwegian Directorate of Health wishes to develop AI-based chatbots to provide advice on health and lifestyle. They will answer questions about subjects such as mental and sexual health, nutrition and physical health, physical activity and exercise, addiction and substance abuse, self-care and lifestyle, digital security and the use of public health services. The aim is to provide more accessible and tailored health information to various target groups.

This sandbox report addresses relevant data protection issues associated with these chatbots. In the report, the Norwegian Data Protection Authority assesses whether information entered into the solutions by users can be considered anonymous, and whether the Directorate of Health must establish a legal basis under data protection regulations in order to offer such solutions.

Within the sandbox project, we have reached three main conclusions:

- **Anonymity**: Data protection rules do not apply to anonymous data, but in practice, anonymising information can be challenging. We have therefore examined whether information used in the solutions meets the requirements for anonymity under the data protection rules, and assessed the risk of the re-identification of users. We have concluded that the Directorate of Health's solutions are not to be considered anonymous and consequently require a legal basis.
- **Legal basis:** It is possible to satisfy the requirements for consent for this type of chatbot, cf. Article 6(1)(a) and Article 9(2)(a). This means that the Directorate of Health can establish a legal basis for this processing. The report highlights challenges and provides recommendations on how this can be achieved. We have also concluded that, at present, no so-called supplementary legal bases exist. Therefore, when deploying these solutions, the Norwegian Directorate of Health cannot rely on the processing of personal data to comply with a legal obligation, perform a task carried out in the public interest, or exercise official authority, cf. Article 6(1)(c) or (e).
- **Technical and organisational measures:** Within the sandbox project, we have also identified several technical and organisational measures that the Directorate of Health can deploy to mitigate the risks associated with this type of processing of personal data. These technical and organisational measures were not tested during the project. The Directorate of Health must ensure that the measures outlined in the report actually function effectively before they are potentially implemented.

It is also important to emphasise that other solutions, as well as specific technical and organisational measures relating to them, must be specifically assessed on case-by-case basis. While this report guides the Directorate of Health in developing its own solution, it will simultaneously help clarify important legal questions for other organisations dealing with similar issues.

The solutions described in this report are examples of services we are likely to see more of across the public sector. Central government authorities should consider establishing a common legal basis for such use – for instance, by regulating it in the Public Administration Act or other relevant sector-specific legislation. A clear legal basis will make it easier for public sector bodies to digitise services efficiently, while making the regulatory framework more uniform and predictable. Such a foundation could also provide a basis for processing under Article 6(1)(e) of the General Data Protection

Regulation (GDPR), thereby simplifying the processing of personal data where anonymisation is not possible.

About the project

The Directorate of Health aims to ensure that more people enjoy good health, to reduce health inequalities among the population, and to ensure high quality and safe care within the health and care services. The Directorate also aims to ensure that patients and service users encounter a coordinated service.

Providing information is a core component of the Directorate's duties. This is provided through various channels, with online guidance playing an increasingly important role. The Directorate of Health recognises that the use of commercial, digital solutions such as websites, search services and chatbots is increasing. These can make it easier for users to find information while also reducing the need to handle manual enquiries.

The Directorate of Health is considering deploying artificial intelligence in the form of Large Language Models (LLMs) and methods known as Retrieval Augmented Generation (RAG). The objective is to make health-related information more tailored and accessible, improve the user experience, and adapt information to different target groups and age groups.

Digital guidance services that utilise language models to answer questions can serve as a means of providing the public with swift, tailored and accessible information. The Directorate of Health has therefore initiated the development of two potential solutions evaluated in this sandbox report:

- **HelseSvar (Health Response):** The Department for Child and Adolescent Health is exploring the possibility of using language models to provide quality-assured health information aimed at young people. The goal is to provide swift, high-quality responses based on professional expertise and reliable information sources. The service will also cover questions relating to mental and sexual health. The Directorate of Health estimates that such a service could answer up to 70,000 health-related questions per year. Target groups such as younger children and adults could also benefit from the service.
- **ETI (Easier Access to Information)** This solution aims to make it easier for parents and guardians of children and young people who require complex services – as well as those service providers – to locate relevant information. The target group consists of adult users and their relatives. Many families currently report spending about 19 hours per week locating information and coordinating services. The solution will aggregate and present relevant information from a range of public sources in a simpler manner than is currently the case. The solution is being developed in collaboration with the other agencies involved in what the Directorate of Health defines as the life event 'Seriously ill child'.

Both solutions will function as web-based chatbots. The user types questions and receives answers generated by language models, which are supported by the Directorate of Health's own information databases.

This technology, however, raises difficult questions. For example, how can the Directorate safeguard user privacy when the service is developed using free-text fields and language models? Free-text fields present the risk that users may enter, both directly and indirectly, identifiable personal data, even though identifying the user is neither intended nor necessary.

Chatbots can also provide incorrect or misleading responses. This occurs because they are designed to generate text based on a vast quantity of global and multilingual information sources, where the training data does not necessarily prioritise quality-assured knowledge about health and Norwegian

healthcare services. Consequently, the Directorate of Health sees a need to develop its own solutions and offer secure, quality-assured digital guidance services tailored to Norwegian users.

The Directorate of Health's solutions aim to make it easier for users to access health information and guidance on health services. This must be carried out in a manner that safeguards privacy and is tailored to individual needs. The ETI solution will additionally be capable of providing cross-sectoral information from actors such as the Norwegian Labour and Welfare Administration, the National Support System for Special Needs Education, and the Directorate for Education and Training.

Unlike traditional search engines, language models can adapt their responses and language to the target group. This makes the technology particularly suitable for providing children and young people with answers that are easier to understand. They can also obtain answers without having to ask an adult.

Public solutions as described above can also serve as an alternative and a counterweight to commercial services that, whether deliberately or inadvertently, disseminate health information. Commercial actors frequently employ business models based on the active identification of users through profiling, which is used for targeted advertising, among other purposes. Public services provide children, parents and healthcare professionals with quality-assured answers, and are developed with the sole purpose and 'business model' of providing safe, professionally sound information. They represent important alternatives to commercial chatbots, search services or social media for finding answers about health and the health services provided by municipalities and the healthcare system. This also increases the probability of providing higher quality information and safeguarding data protection.

During the project, we debated whether information must be stored in an identifiable manner to comply with the duty to avert under Section 31 of the Health Personnel Act. The duty to avert applies to healthcare professionals and entails a duty to notify the police and fire service if it is necessary to prevent serious injury to persons or property.

The Directorate of Health has concluded that this duty does not apply to these solutions, as it is not health personnel who process the information. The services are purely digital and automated. This is an important clarification of principle for the use of artificial intelligence in the health sector. The conclusion implies that there is no need to be able to trace user queries back to specific individuals for the purpose of the duty to avert.

What is the Norwegian Data Protection Authority's Sandbox?

In the Sandbox, participants work together with the Data Protection Authority to address privacy and data protection issues. The goal is ensuring that their service or product meets legal requirements and provides sound data protection. The Data Protection Authority provides advice to participants, but the conclusions are not official or administrative decisions or approvals. Participants are thus free to choose whether they want to follow the advice.

How the solutions utilise language models and retrieval-augmented generation

The Directorate of Health's two solutions, HelseSvar and ETI, are built on language models that generate responses to queries entered by the user.

Both solutions will provide guidance within mental health and sexual health, nutrition and physical health, physical activity and exercise, addiction and substance abuse, self-care and lifestyle, digital security and the use of public health services. ETI will provide cross-sectoral information relevant to the target group.

The user will communicate with the service via prompts. These may be questions, keywords or brief descriptions that provide the model with a basis for formulating a response.

All users will receive clear and comprehensible prompts directly within the user interface directing them not to enter personal data.

When a user submits a question, the application automatically appends a predefined context. This is a hidden prompt that is always added to the user's question, containing system-defined instructions that assist the service in locating the correct information. Additionally, this context dictates how the language model (LLM) should structure its response, with specific tailoring for health-related content in this instance.

As supplementary data the general language model (also referred to as the foundation model), relevant and quality-assured information is collected from controlled sources. This may include previously published guidelines from the Directorate of Health, public documents and publications, or historical questions and answers from existing guidance services. Historical questions and answers harvested from the Directorate of Health's other services will be particularly valuable.

The information from these controlled sources will be stored in a local database that is structured to be compatible with language models. When a user submits a question, the service first executes a *semantic search* within the local database. This means it searches for content that is relevant in meaning, rather than relying strictly on keywords. The quality-assured information retrieved is then used to enrich the query with relevant facts, thereby providing the foundation language model with the necessary context to generate more precise answers.

This method is known as Retrieval-Augmented Generation (RAG), which is essentially information-enriched text generation. This implies that the response should prioritise information built upon local, quality-assured sources rather than relying solely on 'random' and unverified information from the general language model. Nevertheless, it is critical to remain aware that language models utilising RAG can still produce erroneous or incomplete answers; it would be more accurate to state that the risk is reduced. We have previously investigated RAG in other sandbox projects. General information about the method can be found in the report for 'LawAi: final report from the sandbox project with Juridisk ABC'.

Once the solution identifies relevant information, it is combined with the user's question and transmitted collectively to the language model. The general language model then formulates a response using natural language, which is returned to the user. The use of the foundation language model is necessary to produce linguistically sound and comprehensible responses.

The purpose of retrieval-augmented generation is to increase the accuracy and relevance in responses by extracting selected, controlled and up-to-date information from external, quality-assured sources. This is in addition to the data upon which the general language model was originally trained. The user consequently receives more fact-based and up-to-date responses, particularly regarding new or otherwise inaccessible information.

The Directorate of Health's models are instructed to respond with ‘I am unable to answer this’ if they fail to locate a sufficient volume of relevant sources. The instruction aims to reduce potential model hallucinations. The response from the models is therefore more heavily dependent on local sources than on the general language model.

Data protection regulations and anonymity

A central issue in the project was whether the data processed in the solution is to be considered anonymous. If the service does not process personal data, data protection regulations will not apply. Consequently, the solutions would also become more flexible in terms of the subsequent use of the data for purposes such as documentation, ongoing development, improving the quality of responses and rectifying errors.

In assessing whether or not the data is anonymous, we have found it relevant in this project to consider:

- Directly identifiable information
- Indirectly identifiable information

What is personal data?

In order to assess anonymisation, we must first understand what personal data is. Personal data is defined in Article 4(1) of the General Data Protection Regulation as follows:

> (1) any information relating to an identified or identifiable natural person ('data subject'); an identifiable natural person is one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person;

Thus, an organisation must consider more than just unique and directly identifying information, such as names and ID numbers. Data that, when combined with other information, can be linked to a natural person, also constitutes personal data.

Such indirectly identifiable data may, for example, be information concerning gender, sexual orientation, specific diagnoses or a limited geographical area. The examples mentioned here are not coincidental; this is precisely the type of information many users will input, and which the model will require, for the Directorate of Health's solution to provide high-quality responses.

It can be challenging to determine whether a piece of information can be linked to an individual or not. However, if the conclusion is that identification is impossible based on that information – even when combined with information from other sources – it can be considered anonymous data.

Recital 26 states that the Regulation does not apply if the data has been 'rendered anonymous in such a manner that the data subject is not or no longer identifiable'.

The same recital also states that 'to determine whether a natural person is identifiable, account should be taken of all the means reasonably likely to be used, such as singling out, either by the controller or by another person to identify the natural person directly or indirectly'.

See also the EDPB's Opinion 28/2024 on certain data protection aspects related to the processing of personal data in the context of AI models, paragraph 3.2.

In the Directorate of Health's solutions, the language model is structured to process only indirectly identifiable information; nevertheless, it remains difficult to assume that this information is anonymous.

The solutions are designed to store IP addresses from user sessions separately from the questions and answers, meaning that the queries themselves cannot be linked to an IP address. However, the content of the question itself could, in theory, contain a volume of information sufficient to link it to an individual.

Although the Norwegian Directorate of Health actively encourages users to refrain from inputting personal data, and the volume of queries likely to contain personal information is assumed to be low, the Data Protection Authority concludes that the solutions can scarcely be considered anonymous. The proportion of questions that could potentially contain personal data is large enough that the use of the solutions must presuppose the existence of a valid legal basis.

Data protection regulations and legal basis

To be lawful, all organisations that process personal data must have a statutory basis in one of the six legal bases outlined in Article 6 (1) (a) to (f) of the General Data Protection Regulation.

In addition, organisations must comply with the general prohibition against processing special categories of personal data set out in Article 9(1). This applies to health data, among others, because this type of personal data is considered particularly sensitive by its very nature. However, exceptions to this prohibition exist under Article 9(2).

The Data Protection Authority envisages two possible legal bases for the Directorate of Health:

- Consent, cf. Article 6(1)(a), with an exception under Article 9(2)(a), or
- Processing is necessary for the performance of a task carried out in the public interest or in the exercise of official authority vested in the controller, cf. Article 6(1)(e) cf.Article 9(2)(g).

Exercising official authority or performing a task carried out in the public interest

Article 6(1)(e) of the GDPR states that the processing of personal data is lawful if it is necessary for the **performance of a task carried out in the public interest** or **in the exercise of official authority vested in the controller.**

Given that the Directorate of Health's solutions will necessarily involve processing special categories of personal data, and that this data cannot be anonymised (see the previous section), Article 9 of the GDPR applies. While the general rule is a comprehensive prohibition against processing this type of data, a number of exceptions to this prohibition exist. For public sector organisations, the most relevant exemption will be Article 9(2)(g) ('Substantial public interests').

Article 9(2)(g)

(g) processing is necessary for reasons of substantial public interest, on the basis of Union or Member State law which shall be proportionate to the aim pursued, respect the essence of the right to data protection and provide for suitable and specific measures to safeguard the fundamental rights and the interests of the data subject.

Utilising the legal basis in Article 6(1)(e) also requires a so-called supplementary legal basis, cf. Article 6(3) and Article 9(2)(g) of the GDPR.

A supplementary legal basis will primarily consist of an Act of Parliament or a regulation, though it also allows for administrative decisions to be accepted in certain instances.

Supplementary legal bases assessed

The project has examined potential supplementary legal bases. The Directorate of Health's activities are regulated through a number of different Acts and regulations, or via the allocation letter from the Ministry of Health and Care Services.

Section 7-3 of the Specialist Health Services Act regulates the Directorate of Health's advisory duties as they are currently performed, but in the opinion of the Data Protection Authority, it is not sufficiently precise to provide a statutory basis for the processing of personal data to the extent that may be relevant within these services.

The Data Protection Authority has concluded that the Regulations on Electronic Communication with and in the Public Administration do not provide a legal basis for the guidance bots of public sector organisations either.

Section 11 of the Public Administration Act regulates administrative agencies' general duty to provide guidance, but has been assessed as too narrow in scope to authorise this type of service.

The Data Protection Authority does not find that the Directorate of Health holds the necessary statutory authority in any supplementary legal bases for this type of processing.

Consent

Through the project, we have concluded that the Directorate of Health may use consent as a lawful basis.

In principle, consent as a legal basis should be used with caution in the relationship between citizens and public services, partly due to the imbalance of power between the two parties.

The consent of the data subject means any freely given, specific, informed and unambiguous indication of the data subject's wishes by which he or she, by a statement or by a clear affirmative action, signifies agreement to the processing of personal data relating to him or her, cf. Article 4(11) and Article 6(1)(a) of the GDPR. The solutions evaluated in this project may contain special categories of personal data, which imposes additional requirements on the consent. In such cases, the consent must be 'explicit', cf. Article 9(2)(a).

Article 7 of the GDPR lists the conditions for consent to be valid. Read more about what constitutes valid consent.

We will not delve further into the general requirements for valid consent in this report, but will instead point out relevant factors when designing consent solutions for the use of chatbots in public sector organisations.

Balance of power and public authorities

Regarding the use of consent by public sector bodies, considerations surrounding the power imbalance suggest that consent should not be utilised for this type of service.

This is addressed in Recital 43 of the GDPR:

> 'In order to ensure that consent is freely given, consent should not provide a valid legal ground for the processing of personal data in a specific case where there is a clear imbalance between the data subject and the controller, in particular where the controller is a public

> authority and it is therefore unlikely that consent was freely given in all the circumstances of that specific situation.'

However, it will be important to assess what the service offers and the extent to which it involves the exercise of official authority. It is particularly when exercising public authority that it is problematic for a public sector organisation to rely on consent. Uncertainty will quickly arise as to whether such consent is genuinely freely given, as a consequence of the power imbalance between the authority and the individual.

In our assessment, the situation is different regarding the solutions of the Norwegian Directorate of Health: these are only intended to provide guidance that is quality-assured by the Directorate and the municipalities, and which will largely be available on their websites. In other words, this constitutes a form of guidance or service provision from the public sector. No administrative decisions affecting rights or obligations will be made. Nor will there be any form of automated decision-making. The argument regarding a power imbalance will therefore carry less weight than would be the case in situations that revolve more closely around the actual exercise of official authority. In the view of the Data Protection Authority, there will be fewer concerns about using consent as a legal basis in the Directorate of Health's solutions.

Guidance from EDPS

The European Data Protection Supervisor (EDPS) published relevant guidance on 28 October 2025.

Generative AI and the EUDPR. Orientations for ensuring data protection compliance when using Generative AI systems. (Version 2). The guidance provides advice to EU institutions, bodies and agencies on how they can process personal data when deploying generative AI systems. The aim is to help them more easily meet their data protection obligations under Regulation (EU) 2018/1725. Note that while this regulation governs the processing of personal data by EU institutions, it can nevertheless be relevant to the interpretation of the GDPR.

The guidance allows for the use of consent as a legal basis for public sector chatbots. It stresses, however, that the solutions must be designed in a manner that safeguards the rights of data subjects.

> 'The use of consent as a lawful ground may apply in some limited circumstances in the context of the use of generative AI systems. Obtaining consent under the Regulation, and for that consent to be valid, it needs to meet all the legal requirements, including the need for a clear affirmative action by the individual, be freely given, specific, informed and unambiguous.'

Practical implementation of consent

For consent to be valid, the controller must make it possible to withdraw it at any time. If consent is withdrawn, the processing that has already occurred remains in compliance with the regulations and remains lawful.

The Data Protection Authority recommends that the Directorate of Health obtains confirmation of consent within the same interface and screen view as the chat interface itself.

Furthermore, we recommend that the Directorate of Health provides the following information before users consent to the processing of data:

- That the use must not share directly identifying information
- How the information will be used
- That the data will be scrubbed of any directly identifying data and the IP address
- That questions and answers will be stored for a short period for quality and development purposes will be deleted automatically
- A link to a more comprehensive privacy policy

In the Data Protection Authority's assessment, the baseline assumption is that children can also consent to using this type of service. However, this requires a specific assessment. A specific issue in this regard concerns a child's capacity to consent, i.e. whether it is the child or the child's guardian (usually the child's parents) who can give valid consent on behalf of the child. This question depends, among other things, on the age of the child.

Children's capacity to consent is regulated across a variety of general and sector-specific legislation. In principle, the age limit for consent is 18 years, but in certain areas the age of consent is lower.

The age of medical majority is 16 years, cf. section 4-3(1)(b) of the Patient and User Rights Act. This implies that children can take responsibility and make their own decisions in health matters and when requiring healthcare from the age of 16. There are some exceptions until the child reaches 18 years old. Furthermore, from the age of 12, a child may make certain decisions without the parents having a right to participation or information, cf. Sections 4-3 and 3-4 of the Patient and User Rights Act. A relevant question in this context is whether the age limit for consenting to share personal data for health purposes should follow the age of medical majority. We do not, however, take a position on this matter in this report.

In the proposal for the new Children's Act, which has now been passed, Prop. 117 L (2024–2025) the Act relating to Children and Parents (The Children Act)) Sections 6-7, provision is made for the gradual involvement of children in decisions relating to personal data.

The Data Protection Authority will not delve further into the assessment of children's capacity to consent, but recommends that a thorough and specific assessment be made before the solutions are put into service.

Concerning the rights of data subjects

In order for consent to be used as a legal basis, robust information must be provided, and the solutions must be designed in such a manner that all requirements for consent are met. The importance of providing good and comprehensible information is even greater when children are involved.

In the Directorate of Health's solutions, the bulk of the processing will occur when the user receives a response from the chatbot. Questions may arise as to how rights such as erasure and access to information are to be safeguarded when non-directly identifiable personal data is stored for a limited number of days for development and training. The Directorate of Health will have no direct means of linking the questions and answers to a specific individual. Requests for access to information and erasure could thus present practical challenges. One argument in this regard is that this processing falls under Article 11 of the GDPR, which concerns processing that does not require identification. In this report, we assume that there is little risk of re-identification, and that it practically amounts to a mere theoretical possibility. A request for access to information from a user should therefore not lead to a requirement for identification that in practice weakens data protection. A practical solution to the privacy challenges surrounding the retention of data is a short retention period and automatic erasure.

Measures to reduce the volume of personal data in chatbot solutions

Even though we assume that the Directorate of Health's solutions process personal data that requires a legal basis, we discussed measures to anonymise or reduce the volume of personal data in questions that users submit.

Directly identifiable personal data is information that can be directly linked to a specific individual. **Indirectly identifiable personal data** can, through analysis or by alignment of data, also make it possible to identify a specific person. IP addresses can in many cases be an example of the latter.

Use of free-text – risks and measures

The solutions rely on free-text fields in which users formulate their questions. The solutions will provide information and request that users do not submit information that can be linked to specific individuals. However, the risk remains that users input such information into the free-text field.

We assume that, in most cases, users will input texts that are anonymous. Nevertheless, it is important to have measures in place that handle those instances where personal data is entered. It is not practically possible to entirely prevent users from entering sensitive information into a free-text field. The Directorate of Health must therefore implement technical and organisational measures that both prevent and handle content that violates the guidelines.

Such measures will be relevant for all services that utilise free-text fields, including traditional search engines that do not utilise artificial intelligence or language models.

Methods that reduce the probability of personal data being processed

Information and instructions:

The solutions will prompt users that personally identifying information is not required to generate responses and that such data must not be entered. This will take place directly within the interface. However, it may be useful to input information about health, location and support needs in order to receive high-quality responses from the chatbots. In this regard, the Directorate of Health must strike a fine balance between the need for necessary and precise information and the desire to include as little information as possible.

Client-side filters:

The Directorate of Health has integrated filters into the user interface. These are intended to detect and remove (or instruct the user to remove) directly identifying information before the data is transmitted from the local client to the solutions. Examples of this include names, national identity numbers, ages and addresses. In this manner, the Directorate of Health reduces the probability of directly identifiable personal data being submitted and processed by the language model.

Measures that have not been assessed in the context of the Directorate of Health's solutions, but which are possible:

Server-side filtering (DLP):

Organisations that develop or provide more complex services or products can implement measures to further filter user text before it is processed by the solution. This could include measures such as DLP (Data Loss Prevention) methods, preferably implemented locally on the server side. These could potentially detect more complex patterns than simple filters.

Named Entity Recognition (NER) models:

Organisations may want to consider using NLP (Natural Language Processing) models to detect more complex structures containing personal data. An example of models adapted for Norwegian (2025) is Norwegian Named Entity Recognition, developed by the National Library and AI-Lab. Other models include spaCy, Stanza, flair, and many more. Such models can be run locally on proprietary hardware. These can be used to mask more complex combinations of personal data.

LLM models that specifically identify personal data:

There are also a number of language models that are specifically fine-tuned to identify personal data. Frequently, they can also identify different categories of unwanted or harmful data. These language models are available both as API-based cloud services or can be run locally on proprietary hardware. Examples of models (2025) that can be downloaded and run locally include: Granite Guardian (IBM), ShieldGemma (Google), Presidio (Microsoft), and a range of general models that can also be adapted to specifically identify personal data (mistral, llama, gemma, phi).

Specific solutions for text anonymisation:

There are also research projects that utilise machine learning methods to automatically anonymise text documents containing personal data, such as electronic health records, judicial decisions or conversational interactions. One example of such a research project is CLEANUP, led by the Norwegian Computing Centre. The Directorate of Health has concluded that this type of approach would not be suitable for the chatbot solutions in the current phase of the project.

How to handle indirectly identifiable information

For some users, it may be practical to enter information such as a diagnosis, municipality name or family relationships into a chatbot to obtain relevant help. This could potentially constitute indirectly identifiable data, which in turn increases the risk of individuals being identified. The dilemma that arises is that any filters that remove such information may compromise the quality of the responses. The Directorate of Health considers it neither practically feasible nor desirable to remove indirectly identifiable information for these specific services.

However, organisations may wish to consider imposing limits on the volume of text (number of characters) that users are allowed to type or paste into the free-text field (to reduce the possibility of large amounts of text being pasted that users have not read through). When a solution does not allow users to enter large amounts of information, the risk of the solution receiving directly or indirectly identifiable data is naturally reduced. This also means that services that allow users to enter large amounts of text or upload attachments will likely need to make further risk assessments regarding the possibility that the information may be identifiable.

A number of the tools mentioned above (NER/LLM/CLEANUP) will, to a certain extent, also be capable of facilitating the identification of more complex personal information structures than just directly identifying data. Organisations that use such tools must therefore adapt and test them.

The effectiveness of the measures mentioned above has not been tested or evaluated in this project. Nor are they exhaustive, but are described as fundamentally feasible measures. Organisations must assess the measures against the risks to the rights and freedoms of natural persons. The specific measures and their effectiveness must be evaluated for each individual service. We also encourage regular testing and audits to ensure the effectiveness of the measures over time.

Conclusion and the way forward

The Directorate of Health wishes to develop chatbots to provide advice on health and lifestyle. The aim is to provide more accessible and tailored health information to various target groups. The solutions can play a key role in enabling the public sector's ability to disseminate quality-assured health information and guidance on healthcare services.

In this sandbox report, the Data Protection Authority has assessed whether information entered by users into two specific solutions can be regarded as anonymous. This assessment is critical to determining whether the Directorate of Health must establish a legal basis under data protection legislation.

The Data Protection Authority's assessment is that the information in the solutions cannot to be considered anonymous. The Data Protection Authority has further concluded that the Directorate of Health may use consent as the legal basis for this processing, provided that it meets the requirements for consent set out in the General Data Protection Regulation under Article 6(1)(a) and Article 9(2)(a).

The report also outlines a number of general technical and organisational measures that the Directorate of Health and other organisations can implement to mitigate the risks associated with the processing of personal data in this type of solution.

It is also important to emphasise that other solutions and the specific technical and organisational measures relating to them must be specifically assessed on case-by-case basis. The report provides guidance to the Directorate of Health on the development of its solutions. Hopefully, it will also help to clarify important legal issues for other organisations facing similar challenges.

It is important that public sector organisations assess the extent to which a chatbot service involves the exercise of official authority. This applies particularly in cases where chatbots are used to provide guidance on citizens' rights, which could potentially limit the ability or willingness to utilise traditional case processing. A chatbot could be the first point of contact a user encounters when utilising public services. It is therefore important to distinguish between situations involving guidance and those involving a chatbot taking an active stance on, for example, a potential complaint or application.

In this context, the Data Protection Authority will also consider publishing general guidance on chatbots and data protection.

The solutions evaluated in this report are examples of services that we will likely see more of across the public sector in the time ahead. The central government should therefore consider establishing a common legal basis for such use. The legislature could, for example, regulate this in the Public Administration Act or other sector-specific legislation. **A clear legal basis will** make it easier for public sector bodies to digitise services efficiently, while making the regulatory framework more uniform and predictable. Such a basis may provide a legal basis for processing under Article 6(1)(e) of the GDPR, thereby facilitating the processing of personal data in cases where anonymisation is not possible.

Office address:
Trelastgata 3, Oslo

Postal address:
P.O. Box 458 Sentrum,
NO-0105 Oslo

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

datatilsynet.no
personvernbloggen.no