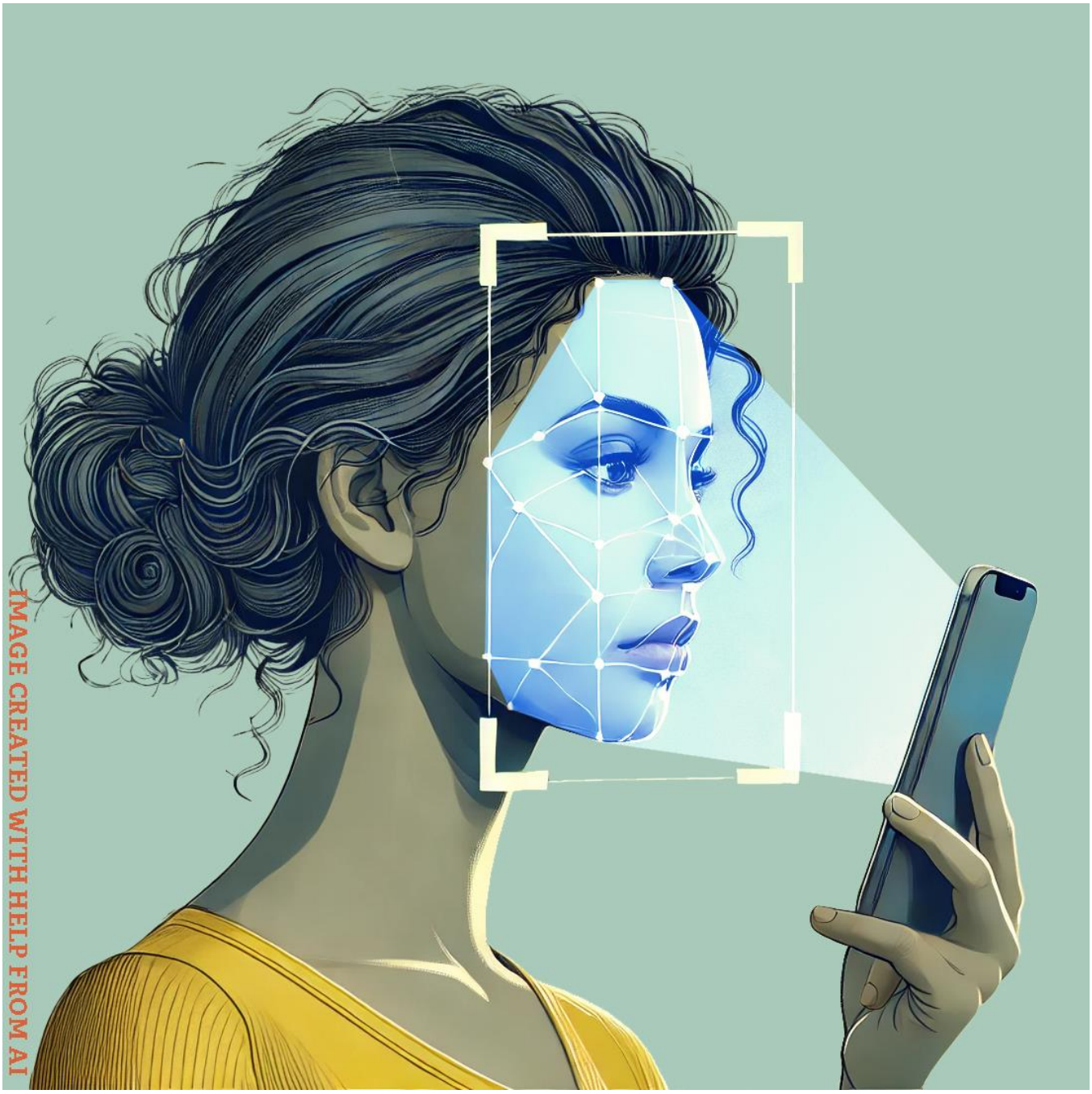

Securing Digital Identities

Biometric data and protected templates in eID solutions

Exit report from the “SALT” sandbox project with Mobai

Datatilsynet

Table of contents

January 2025

This pdf corresponds to the first version of the report, as it was published on the Data Protection Authority's website January 2025. Technology and law are constantly evolving, so there may be a need to adjust or clarify the reports over time. If this pdf differs from what is on the Data Protection Authority's website, you can assume that the website text is the current advice.

Summary

Digital services are an integral part of our modern lives, and securely and effectively verifying our identity online is necessary for using and trusting these services. There is at the same time a pressure from malicious actors to exploit weaknesses in the verification mechanisms. An increasingly important vector for exploiting weaknesses is social manipulation, where users are tricked into giving away the information needed to access online services.

Digital services typically offer several ways for a person to verify their identity, usually requiring multiple factors. This can include information about something you know (passwords or PIN codes), something you have (a specific device or a physical key) or something you are (biometrics). While passwords and PIN are widely used, we see a shift towards increased use of biometric data to increase the security of identity verification. Biometric data encompasses unique information about our physical, physiological or behavioural characteristics – most commonly referring to our fingerprints or facial images.

Mobai, with their solution "SALT", seeks to increase the robustness of digital identity verification by, amongst other things, implementing real-time check of facial images using mobile devices as users verify their identity, as an addition to other security measures.

Due to the uniqueness of biometric data and the potential consequences if it is stolen, it is subject to strict privacy regulations that limits its collection, storage and use.

To address these concerns, Mobai seeks to reduce the risks associated with processing biometric data. They aim to do this by leveraging artificial intelligence and novel machine learning techniques and, by doing so, decrease the privacy risks for the user whilst also expanding how biometric information can be used for digital verification.

In this report, the Norwegian Data Protection Authority will address key legal challenges posed by "SALT", including whether Mobai's technology can expand how biometric information can be used for verification purposes.

Our aim is to give valuable and broadly applicable legal insights, which can benefit Mobai and their partners, as well as other actors working in related fields or with similar technology.

Main findings

In this sandbox project we have addressed some key legal challenges related to how the SALT-solution works, that are also applicable to other companies that work in similar fields and with similar legal considerations.

Our legal assessments of these challenges are as follows:

- **Are facial images considered biometric data?**

  A facial image in itself does not always qualify as biometric data for GDPR purposes. However, we do, consider biometric templates as biometric data. We also consider the processing operations performed on the facial image to generate the biometric template (i.e., the biometric feature extraction), as processing of biometric data under the GDPR.

  Although facial images in themselves are not systematically considered biometric data, processing of facial images may be subject to the same level of security requirements as for biometric data.

- **Are protected biometric templates personal information subject to the GDPR?**

  We have discussed with Mobai whether protected biometric templates – that result from processing using homomorphic encryption – can be considered anonymous, or personal information subject to the GDPR. We argue that while homomorphic encryption and protected templates make the information incomprehensible to other parties without the encryption key, this does not necessarily guarantee anonymity. As long as the encryption key can be used to link the template to a real person, we contend that it falls within the definition of personal information and therefore the GDPR applies.

- **Is biometric information used for verification considered to entail the processing of special categories of personal data (article 9)?**

  There is legal uncertainty regarding whether biometric verification entails a processing of special categories of personal data as covered by Article 9(1). This is not an issue that can be solved in this sandbox project. However, based on our discussions, the Norwegian Data Protection Authority do consider that it is likely that biometric verification is covered by Article 9(1). We would therefore recommend those involved in the SALT-project to treat the biometric data used for verification purposes as a special category of personal data.

- **What is the difference between primary and secondary processing of information in the SALT-solution?**

  The primary purpose of SALT is to ensure secure verification of identities. We regard certain "pre-processing" operations – which in Mobai's case relates to the template creation and the comparison of protected templates – as intrinsic to this primary purpose. However, this is only the case when they are strictly necessary in order to achieve this purpose. We consider various processing that are not carried out for the purpose of identity verification, as processing for secondary purposes. This includes purposes related to improvements to the algorithms and the system as such.

  In line with this argument, the legal basis for the primary processing does not extend to include processing conducted for these secondary purposes.

- **It is possible to store biometric information on a central server?**

  Mobai's approach for storing and processing biometric data requires centralized storage for its enhanced security. While centralized storage offers advanced encryption, access controls, and robust safeguards, it also raises general concerns about the wide scale collection of biometric and other data that can be used for verification of digital identities.

  Mobai asserts that by employing innovative technology to generate Protected Biometric Templates (PBTs), it is possible to balance the need for security and privacy by combining the features of decentralized devices with central storage and processing.

  In addition, Mobai stores encrypted images centrally for necessary AI model training, while claiming strict data minimization practices, aiming to balance security and compliance in management of sensitive data.

  In this project, the Norwegian Data Protection Authority has assessed that Mobai's solution might enable the implementation of central biometric data storage and processing in cases where it was previously not considered secure enough to address the significant concerns associated with central storage.

About the project

Mobai is a spin-off company from the Norwegian University of Science and Technology (NTNU). They provide biometrics technology for digital services, that is used specifically in onboarding- and verification processes.

The focus in this sandbox project is their solution called “SALT”, abbreviated from: “Secure privacy preserving authentication using facial biometrics to protect your identity”. In the development of SALT – which is funded by The Research Council of Norway – Mobai collaborates with BankID BankAxept AS and Sparebank 1 Østlandet.

The Norwegian National Criminal Investigation Services’s (“Kripos”) stated in their 2023 annual report that digital fraud is prevalent in Norway and that there is an increase in identity fraud related to eID solutions. It is especially the issuing, re-issuing and use of e-IDs that are vulnerable to a variety of attacks, potentially resulting in credentials being stolen or misused.

The ambition with SALT is to pilot a solution for strong digital face verification that significantly reduces the risk of identity theft of digital accounts. It also aims to eliminate misuse and theft of biometric data.

In order to prevent this, Mobai is developing SALT to be used as a security component– that can complement existing security measures – in eID-solutions.

> **eID**
>
> eID is abbreviated from «electronic identification». An eID is a way to prove your identity on the internet. It refers specifically to those who provide identity services, such as verification and digital signing, within the European Identity Regulatory Framework.

About the SALT-solution

A key technology in the SALT-solution is an innovative encryption technique called “homomorphic encryption”. Unlike other encryption methods, homomorphic encryption allows computations on encrypted data without decrypting it.

In facial biometric systems like SALT, identities are verified by comparing two face images. This comparison produces a similarity score that indicates the degree of similarity between the faces. In the case of the SALT-solution, the comparisons are carried out entirely within an encrypted domain – using homomorphic encryption – and the result from the comparison is encrypted as well.

In order to compare face images, biometric features are extracted from images of the user’s face. Every biometric feature on the image is then given a vector-representation. This vector-representation of the image constitutes a dataset which is machine readable and fit to use for facial recognition. This dataset is called a “biometric template”. The similarity score is generated by comparing the set of biometric features in the biometric templates.

It is important to note that the biometric templates contain only plain text, and not the images themselves. However, it has been demonstrated that some degree of facial information and traits can be reconstructed from biometric templates. To prevent this, it is necessary to apply strong security measures to protect them.

The generation of “protected templates” is one such security measure. Protected templates are – in short – a security measure where the plaintext templates are transformed and encrypted. In “SALT”, the protected templates are encrypted using homomorphic encryption.

The benefits of these protected templates are that they can be used for face recognition – just like plaintext templates. However, with the current state of technology, it is assumed to be virtually impossible to re-generate the plaintext template from a protected template, and therefore not possible to restore facial images (or other characteristics) pertaining to the user.

Requirements for protected templates

According to the International Organization for Standardization (ISO/IEC 24745:2022) a protected template must meet the following requirements:

- **Revocability and Renewability:** A system must be able to revoke compromised protected templates, so they can't be further used/misused. Consequently, the system must be able to generate a fresh, uncompromised protected template and replace the original one so that the end-user can still use the solution.
- **Irreversibility:** the face image or face template shall be processed by irreversible transformation before storage.
- **Unlinkability:** the stored face templates should not be linkable across applications or databases.

How does the encryption keys work?

In this chapter we describe the role of encryption keys in the SALT-solution. The management of keys is an essential part of the solution, as the creation and use of protected templates involve the use of multiple encryption keys that each provide different security and privacy properties.

Marked with colours in the figure below, you will find an overview of keys used in the creation of protected templates. There are three encryption keys: “template transformation key”, “homomorphic encryption key” and “sealing key”. The “sealing key” is used to ensure the integrity of the system in the encrypted domain and does not add privacy properties to the protected template.

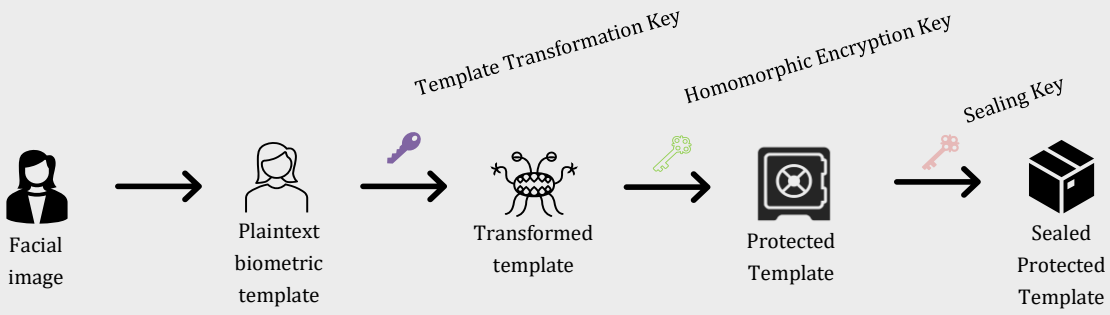

Figure 1: Overview of keys used in the creation of a protected template

The Template Transformation Key

The first key (“template transformation key”) is used to transform a plaintext biometric template to a transformed template, by combining a “seed” with the transformation process to generate a new and unique version of a plaintext template each time the process is performed. Mobai claims that this transformation provides the privacy properties of irreversibility and unlinkability, in accordance with ISO/IEC 24745:2022.

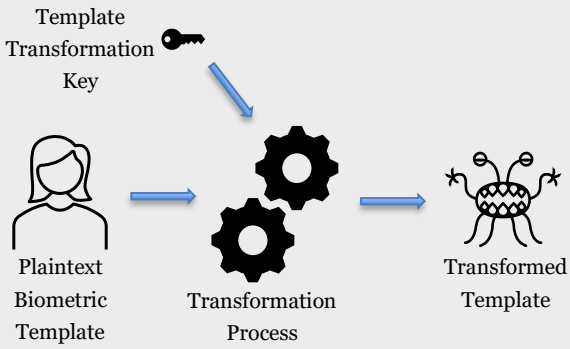

Figure 2: The template generation process takes the plaintext biometric template and combines it with the template transformation key to create the transformed template.

It is assumed to be virtually impossible to re-construct facial features from a transformed template, with the currently known technology, which can ensure the irreversibility of the data. The unlinkability (another requirement according to the ISO) between two protected templates generated from the same person is secured by using different transformation keys for each enrolment.

Furthermore, if a transformed template becomes compromised, the old template transformation key can be thrown away which will render the transformed template unusable, which can claim the revocability of the information. Furthermore, a new transformed template can be created using a new template transformation key, which ensures its renewability.

The transformed template is an intermediate step in the process of generating protected templates. The transformed template is not used for face comparison and is not stored.

The Homomorphic Encryption Key

The second stage is to encrypt the transformed template with a homomorphic encryption key. The output of this encryption is the protected template.

The encryption ensures the confidentiality of the protected template, both at storage and at use. It also ensures confidentiality of the output of the face comparison. This output, which is the similarity score of the face comparison, is the basis for the verification decision. In this system only the output is decrypted.

The homomorphic encryption scheme is based on asynchronous encryption: There are two keys – a key pair – where a data element encrypted by the first key (i.e., the protected template) can only be decrypted by the second key. An important security principle enabled by asynchronous encryption is key separation, i.e., no one entity has access to both keys. In the SALT-solution, the holder of the second key will only have access to the output. This system enables true isolation between the storage and use of the protected templates, and their decryption key. Thus, avoiding single points of compromise in the solution.

The Sealing Key

The sealing key functions as a digital signature and ensures the integrity of the protected template and the output.

This step is using an established scheme for electronic signatures. The verification process should validate this signature before each use of a protected template.

Key Management

Mobai intends to manage the keys in a way that ensures the confidentiality of these processes. The key management system will generate and manage keys in secure environments. The template transformation key is generated using a strong random number generator with sufficient length, and the homomorphic encryption key is generated using the latest homomorphic encryption scheme. A cryptographic key management policy will be implemented by following ISO 27001 standards.

How does the solution work?

A step-by-step description of the solution

Below we provide a simplified step-by-step explanation of how the SALT-solution works in practice. The explanation outlines a scenario where an end-user consents to use the SALT-solution to gain access to an online service provided by a bank.

The process consists of two phases: an onboarding phase and a verification phase.

There is also a final phase, where personal information undergoes processing for secondary purposes, including improving the fraud detection module and the machine learning algorithms. We will explore secondary purposes more closely in the next chapter.

The onboarding phase:

1. **Initiation:** The onboarding phase is initiated when the user submits a "reference image". The reference image is the face image extracted from a nationally issued ID-document, such as a national ID-card or a passport. Specifically, the image is retrieved from the Radio Frequency Identification (RFID) chip in the document.
2. **Template Creation:** A template creation module converts the reference image into a "plaintext template".
3. **Key Generation:** A key management server will generate a unique "transformation key" which is specific to each user. It will also generate a "homomorphic encryption key", which is specific to each eID provider.
4. **Secure processing:** The plaintext template is secured by a transformation process using the transformation key, followed by an encryption process using the homomorphic encryption key. This process results in the generation of a protected template. The protected template is then sealed using the sealing key. This ensures the integrity of the protected template.
5. **Storage:** The sealed protected template is stored in a storage server. This protected template is the reference template, and it is used for the face comparison in the verification phase.
6. **Verification:** To ensure that the onboarding is performed by the holder of the national ID-document, the user must verify his/her identity following the steps in the "verification phase", described below. If the verification decision is positive, the user has been onboarded for further use (i.e., ongoing verifications).

Verification phase

To gain access to a service provider's/merchant's services, the user must undergo a verification phase, detailed below:

1. **Image capture:** The user captures a live image of himself/herself (a "selfie") on their mobile device and uploads this image to the SALT-solution.
2. **Fraud Detection Module:** A fraud detection module (see info box) analyses the image to assess if it was captured live from a present person without any image manipulation or fabrication.
3. **Generation of plaintext template:** If the fraud detection module concludes that the image is authentic, the template creation module generates a plaintext template.
4. **Generation of protected template:** A protected template will be generated from this plaintext template using the same procedure as described earlier, using the same transformation key and homomorphic encryption key used for the creation of the user's reference template. The protected template is then sealed using the sealing key.
5. **Face verification and comparison:** The sealed protected template is sent to the face verification module where it will be compared with the user's protected reference template.
6. **Score comparison:** The comparison generates a similarity score and then make a verification decision based on whether or not the score meets a pre-defined threshold. The output from the process is the verification decision, which is accessible to the vendor.
7. **Authorization:** If the verification decision is positive, the user will be authorized to use the online service.

What are the components of the facial verification system?

This system involves the following modules:

- **Capture module:** This module runs in mobile applications on the users' smartphones and communicates with the camera application and the user during the selfie capturing sequence. The mobile application will typically be provided by a bank or BankID. The capture module is a software component in these mobile applications.
- **Face verification module:** This module compares a reference image (derived from an ID-document during the onboarding phase), and a live image of the face that the user takes of him/herself during the verification phase. This module generates a similarity score that indicates the probability of the images coming from the same person.
- **Fraud detection module:** This module analyses live images to assess whether they were captured live from a present person, or if there is an impersonation or manipulation. Fraud detection includes the following sub-modules
    - **Face presentation attack detection:** This module can detect if the live image is real, or if it is a digital surface or a mask. This module prevents fraudsters from presenting digital images of other users to gain remote access.
    - **Face deepfake detection:** This module detects if the live image is a deepfake, i.e., a synthetic image that is generated through machine learning or that has been swapped.
    - **Face morphing attack detection:** This module detects if the reference image is intentionally «morphed». A morphing attack consist of two or more facial identities that are digitally combined to create a single facial image that represents multiple identities. In a successful morphing attack, an ID-document can be used by multiple identities.

All these modules contain pre-trained models that uses artificial intelligence.

What personal data will be processed?

As outlined in the step-by-step description, personal information about the user will be gathered and processed. This personal information includes:

- Facial images captured with the end-user's device.
- Plaintext templates that are generated from the reference image (i.e., a nationally issued ID-document).
- Protected templates, which encompass plaintext templates that is protected by homomorphic encryption.
- Face image extracted from the Radio Frequency Identification (RFID) chip of the end-user's ID-document (for example passport or national issued ID card)
- Digital readable data from the machine-readable section of nationally issued IDs, as per the electronic machine-readable travel document standard (ICAO 9303)
- Data from the machine-readable zone ("MRZ") on the ID-document, necessary to gain access to the information on the RFID chip
- Information about the user device

As related to the step-by-step description above, Mobai will use this information for the purpose of the enrolment to a service and following verifications.

Who are the most relevant stakeholders?

In the SALT-solution, there are four main stakeholders who play distinct roles. These are:

1. **eID provider:** This actor issues the eID (such as BankID, Buypass, Commfides and MinID,) to the end-user and provides the eID services to a merchant.
2. **Mobai:** Sub-vendor to the eID provider, where Mobai provides specific components to the eID services. SALT augments these services by providing an extra security layer for services offered by the eID provider. These components include services like 1) remote onboarding to issue an eID, 2) re-issuing an eID, 3) resetting password or 4) confirming "liveliness" (i.e., that the correct person has access).
3. **Merchant (such as SpareBank 1 Østlandet):** This third-party act as a merchant and relies on the eID provider, and consequentially Mobai's components as an extra security layer, in order to ensure that they are giving access to their services to the correct end-user.
4. **The end-user:** The motivation for the end-user to use the solution is to remotely verify their identity to a service without the risk of "fraudsters" using their eID with or without their knowledge.

Personal data for secondary purposes

In this chapter we describe the use of personal information for secondary purposes.

About "secondary" and "primary" purposes

In the context of this project, "secondary purposes" refers to the use of personal information beyond the primary purpose – which is the enrolment to the e-ID-service and the following identity verifications.

The secondary purposes mentioned in this report are all known to the controller at the time when the information is first collected. We therefore assume that the personal information will initially be collected for both the primary and the secondary purposes mentioned in this report. Thus, Article 6(4) GDPR regulating processing for new purposes will not apply.

Mobai's secondary purposes include:

- Implementing improvements to the prediction accuracy for capture modules, face comparison and fraud detection algorithms, including:
    - training algorithms,
    - identifying unknown fraud types to improve algorithms and
    - improving security and make general improvements to the service
- Conducting fraud investigation after real-time sessions (e.g., in case of a dispute)
- Providing evidence for law enforcement, if and when required by law
- Bias reduction

The information that is stored for secondary purposes include:

- The live image, i.e., "selfie"
- Various device- and session-related data

Mobai consider themselves controller for the secondary purposes related to making improvements to the algorithms, and bias reduction. For the secondary purposes related to fraud investigation in case of a dispute, and providing evidence for law enforcement, Mobai initially considered themselves data processor.

In a later chapter of this report, we take a closer look on the differences between primary and secondary processing in the SALT-solution. In that regard, we look at the purpose related to training and making improvements to the algorithms and the system as such (the first of the four bullet points above). The purposes mentioned in the other three bullet points have not been part of the scope of this sandbox project. However, when it comes to bias reduction, there has been a relevant project in the regulatory sandbox of the Information Commissioner's Office (ICO) in the UK on this issue. For further reading, see the Onfido Regulatory Sandbox Final Report (ico.org.uk).

Central storage

Mobai will store the data used for secondary processing in a secure server environment with physical security measures. The data will be encrypted using conventional methods (not homomorphic encryption) and stored in a dedicated environment separated from the service offerings, where the risk of misuse and data leakage is mitigated. This is to ensure that only authorised personnel have access. Mobai may provide sharing of session data with business partners in live operations.

Storage time

In the SALT-solution, Mobai will retain personal data for 180 days from the moment the initial result of a verification transaction is passed to the merchant to allow for machine learning and improvement of the fraud detection systems. An exception is the data gathered for the training of algorithms. These will be stored for a full year, in order to acquire a high enough number of samples to ensure efficient training.

The goals of the sandbox project

Based on Mobai's use of artificial intelligence and encryption techniques, the Norwegian Data Protection Authority and Mobai have jointly identified two key areas that needs more in-depth legal consideration.

These issues are also considered relevant for other developers and enterprises wishing to use similar machine learning techniques and technologies.

1. **Assess the legal status of facial images and protected templates in an AI-based solution for verification.**
   Focus on legal status of facial images and protected templates, processing for a primary purpose of biometric verification and for secondary purposes.

2. **Assess the technical security measures for storing of protected templates to be used by the AI and facial images for training purposes**
   The design and use of technical security measures in the SALT-solution will partially depend on the results of discussions in #1.

Legal basis and legal status

Artificial intelligence often requires the processing of large amounts of data — often personal information — which is compiled and analysed on a scale that is not possible by other means.

All processing of personal data requires a legal basis to be lawful. Article 6(1) (a-f) of the General Data Protection Regulation (GDPR) contains an exhaustive list of six legal bases for the lawful processing of personal data.

It is natural to split the question of legal basis in two, based on the two main phases in an AI project; the live service phase, and the development phase. The development phase can often occur before, during and after the service is active, and the two phases often utilise personal information in different ways.

In this sandbox project, we have taken no position on whether Mobai or the eID providers have a legal basis for processing personal data in the artificial intelligence tool that Mobai is offering (the SALT-solution). This applies both to the use of the SALT-solution for the primary purpose of remote identity verification and any use of personal data for secondary purposes as mentioned in this report.

The discussions in this sandbox project presume that the data controllers, whether Mobai itself or the eID providers, find a legal basis for processing personal data when using and further developing the SALT-solution.

Why discuss legal status

For Mobai, it is important to clarify the legal status of the key data that is a part of the solution. This includes facial images, plaintext templates, and protected templates.

By “legal status”, we refer to several questions, such as: Is the data processed considered to be personal information subjugated to the GDPR? And at what point is the information considered to be “biometric”? And when is it considered to entail the special categories of personal data, subjugated to article 9? These questions are important to clarify – especially in regard to encryption techniques and machine learning – as the GDPR does not apply to data that has been completely anonymised.

Are protected templates personal information or anonymous?

The distinction between personal and anonymous data can be complex. The threshold for when data can be considered anonymous is very high.

It is easier to pseudonymise data, which replaces directly identifiable parameters with pseudonyms, which will still constitute unique identifying indicators. Pseudonymisation can make it more difficult to link a specific data set to the data subject's identity and can therefore be seen as a useful technique to strengthen privacy. However, as opposed to anonymous data, pseudonymised data is still considered to be personal information subject to the GDPR.

For more information about anonymisation and pseudonymisation, see our general guidance.

For protected templates, Mobai must assess if the associated information is anonymised, pseudonymised or simply regular personal information.

While homomorphic encryption and protected templates make the information incomprehensible to other parties without the encryption key, this does not necessarily guarantee anonymity. As long as the key can be used to connect the template to a real person, it is reasonable to say that it falls within the definition of personal information in the GDPR.

Article 4 of the GDPR supports this view. It defines personal data as information relating to an identifiable person, either directly or indirectly. Even though it is not possible to directly identify someone based on the content in the protected template, the key itself enables indirect identification.

Are facial images biometric data?

Biometric data are defined in Article 4(14) GDPR as "personal data resulting from specific technical processing relating to the physical, physiological or behavioural characteristics of a natural person, which allow or confirm the unique identification of that natural person, such as facial images or dactyloscopic [i.e., fingerprint] data".

The meaning of "which allow or confirm the unique identification of that natural person" in Article 4(14) was discussed in the sandbox project. In particular, it was discussed whether this wording entails that only personal data used for the specific purpose of uniquely identifying an individual fall within the definition. According to the Norwegian DPA this is an overly restrictive reading of the definition; data that are suitable to enable such unique identification also fall within the definition in Article 4(14). Thus, the definition of biometric data in the GDPR is essentially equivalent to the definition of biometric data in the AI Act.

A facial image in itself does not always qualify as biometric data for GDPR purposes. This is made clear in Recital 51 of the GDPR, which states that:

> "… The processing of photographs should not systematically be considered to be processing of special categories of personal data as they are covered by the definition of biometric data only when processed through a specific technical means allowing the unique identification or authentication of a natural person. …"

The term "specific technical processing" is neither defined in the GDPR nor interpreted in case law. However, in our view, the biometric templates generated by the template creation module meet the definition in Article 4(14) GDPR, and should be considered biometric data. We are also of the opinion that the processing operations performed on the facial image to generate the biometric template (i.e., the biometric feature extraction) should be considered processing of biometric data under the GDPR.

Although facial images in themselves are not systematically considered biometric data, processing of facial images may be subject to the same level of security requirements as for biometric data. For example, the creation of a database containing a large amount of high-quality facial images belonging to persons with a verified ID is likely to entail a high risk to the rights and freedoms of the data subjects whose facial images are registered in the database. For further information on this issue, see the section below on central storage of biometric information.

What are the differences between primary and secondary processing in SALT?

There are many different operations taking places when a remote biometric verification is carried out through an eID using the SALT-solution. A live image (facial image/selfie) of the user is captured through a mobile application on the user's device. Before generating a biometric template, a "pre-processing" will be performed on the image. This "pre-processing" consists of different operations e.g., image quality check, presentation attack detection, deepfake detection and morphing attack detection. These operations are performed to make sure that the presented facial image is genuine.

When the facial image has passed the "pre-processing", a protected template is created and matched with the reference template, resulting in a similarity score.

In the sandbox it has been discussed whether this set of operations could be deemed to be covered by the definition of "processing" in Article 4(2) GDPR. In this respect, it should be noted that Article 4(2) defines "processing" as:

> "Any operation or set of operations which is performed on personal data or on sets of personal data, whether or not by automated means, (…)" (our emphasis)

Common to the operations mentioned above, is that they all relate to the execution of a single verification request. When assessing whether these operations qualify as a single processing under Article 4(2) GDPR, it is natural to look at the main purpose of the processing. In this case, the main purpose is to perform a remote identity verification.

The operations described as "pre-processing" are necessary in order to ensure that the facial image to be matched with the reference image is genuine. This is not specific to remote verification. As an example, when a verification is carried out in person, for example in a bank, the bank employee will automatically check if the person standing in front of him or her is genuine, i.e., not wearing a mask etc.

We consider the "pre-processing" operations as a natural part of the verification process, and strictly necessary in order to achieve the purpose of identity verification. Against this background, we believe that it is natural to see the "pre-processing", the template creation-phase and the comparison of the protected template with the reference template, as "a set of operations" carried out for the main purpose of performing a remote identity verification.

Mobai considers that it qualifies as a processor when executing an identity verification request. Mobai will, however, also process personal data collected during both the onboarding and verification process to train and improve their machine learning models (prediction accuracy for capture module, face comparison and fraud detection algorithms and measures, including training of algorithms that contribute to improvements within mentioned areas of interest), in addition to security and general improvements of the service. For the execution of these processing operations Mobai considers that it qualifies as a controller.

Unlike the operations described as "pre-processing", the training and improvements of the machine learning models and the system as such will be performed after the completion of the identity verification. At this stage the user has already had his or her identity confirmed or rejected by the system. Thus, the processing carried out for training and improvements do not have the same link to the execution of a single verification operation as the "pre-processing" operations. These processes do however have a close connection with the main function of the service, as the training and improvements are necessary processes for the AI-models used in the SALT-solution to function as intended over time. However, the processing is not carried out for the purpose of completing an identity verification but for various purposes related to improvements on the algorithms and the system as such. We therefore consider this processing for secondary purposes.

As already mentioned, Mobai considers that it qualifies as a controller for the secondary processing purposes related to improvements on the algorithms and the system as such, while it qualifies as processors for the primary processing (the enrolment and execution of a verification request). Mobai cannot be both controller and processor for processing activities that falls under the same purpose. This is also an argument that supports the conclusion that the processing related to improvements on the algorithms and the system as such, do not fall under the primary purpose.

When is the processing of biometric data subject to Article 9?

The processing of biometric data is not always considered to be processing of special categories of personal data. This is because the processing of biometric data is only considered to be processing of special categories of personal data when processed "for the purpose of uniquely identifying a natural person", cf. Article 9(1) GDPR.

The legal meaning of the term "uniquely identifying a natural person" has been the subject of considerable debate. Some argue that the term only covers biometric identification (1:n), while others argue that biometric verification (1:1) is also covered by this term. As Mobai's intention is to process biometric data for verification purposes, the interpretation of this term is crucial to assess whether or not the envisaged processing is covered by Article 9 in the GDPR.

In this report we refer to:

"Biometric identification" as a one-to-many (1:n) comparison where the biometric template of a person with an unknown identity is compared with a database of templates in order to reveal the identity of the person in question;

"Biometric verification" or "biometric authentication» as a one-to-one (1:1) comparison where the biometric template of a person with a claimed identity is compared to a single reference template in order to verify the claimed identity

A literal reading of the wording in Article 9(1) does not seem to provide sufficient clarity to conclude on its scope of application with respect to biometric data. Some argue that "uniquely identifying" only covers biometric identification, while others point out that one has to uniquely identify someone in order to verify a claimed identity as part of a biometric verification process as well. The latter reading would find some support in the definition of biometric data in Article 4(14) GDPR, which covers personal data resulting from specific technical processing "which allow or confirm the unique identification of that natural person» (our emphasis).

A reference to biometric data was not present in the first version of Article 9 in the GDPR legislative proposal tabled by the European Commission. Such a reference was first added during the trialogue negotiations, at the request of the European Parliament. As it appears from the written debriefing of the trialogue negotiation on 24 November 2015, the legislators' intention was to regulate the processing of biometric data in line with the modernised Convention 108 of the Council of Europe, which restrictively defines biometric data that 'uniquely identify a person' to qualify as sensitive data. Nonetheless, neither the preparatory works on the GDPR nor the preparatory work on the modernised Convention 108 are absolutely clear on whether biometric verification should be deemed to qualify as processing of sensitive data.

The use of one-to-many biometric identification schemes generally has a higher impact on fundamental rights than the use of a one-to-one biometric verification system. This difference in impact between biometric identification and biometric verification is clearly evident in the proposal of the AI Act (version 13 June 2024), which imposes strict requirements on the use of AI-based biometric identification schemes. This fact could support a reading of Article 9(1) GDPR according to which only biometric identification (1:n) could be seen as processing of sensitive data.

Case law from the Court of Justice of the European Union (CJEU), do however, support a wide interpretation of Article 9, cf. case C-184/20 (see paragraph 125).

A recent opinion from the European Data Protection Board (EDPB) also suggests that processing of biometric data for verification purposes is covered by Article 9(1), see Opinion 11/2024 on the use of facial recognition to streamline airport passengers' flow (compatibility with Articles 5(1)(e) and(f), 25 and 32 GDPR. In the aforementioned document the EDPB state the following:

> "Facial recognition technology can fulfil two distinct functions – authentication and identification. While both functions are distinct, they both rely on the processing of biometric data related to an identified or identifiable natural person and, therefore, constitute processing of special categories of personal data under Article 9 GDPR." (see paragraph 21)

Previous Guidelines from the EDPB also suggests that processing of biometric data for verification purposes is covered by Article 9(1), (see Guidelines 05/2022 on the use of facial recognition technology in the area of law enforcement | European Data Protection Board (europa.eu), paragraph 12; the EDPB Guidelines 3/2019 on processing of personal data through video devices, paragraph 78).

As outlined above, there is a significant degree of legal uncertainty regarding whether biometric verification is covered by Article 9(1). This is not an issue that can be solved in this sandbox project. However, based on our discussions, we do consider that it is likely that biometric verification is covered by Article 9(1). We would therefore recommend the SALT-project to treat the biometric data used for verification purposes as a special category of personal data.

The most obvious consequence of this is that Mobai or their customers, depending on who is the controller, need to identify and demonstrate the existence of a valid exception in Article 9(2) as well as a legal basis in Article 6(1) to be able to lawfully process personal data for this purpose.

However, whether or not the processing of biometric data falls under Article 9 does not necessarily change the level of security measures required according to the GDPR. Even though biometric data was not considered a special category of personal data according to Directive 95/46/EC, the predecessor of the GDPR, processing of biometric data was subject to strict security requirements. In our view, there is nothing that indicates that the introduction of the GDPR changes this. This has also been the view of the participants from the SALT-project.

Does Article 9 apply when processing for secondary purposes?

As mentioned above, biometric data in itself is not considered a special category of personal data. It is the purpose that decides whether or not Article 9 applies to the processing.

When looking into the different processing operations relevant to this project, we have recommended that the processing operations connected to implementing improvements to the prediction accuracy for capture modules, face comparison and fraud detection algorithms, in addition to fraud investigation after real-time sessions, are considered processing for secondary purposes.

In this case the secondary purposes relate to training and improvement of AI models and the system as such. As pointed out earlier, this processing is not carried out for the purpose of completing an identity verification. Thus, we do not consider this processing "for the purpose of uniquely identifying a natural person". On this background we believe that Article 9 will not apply when processing personal data for these secondary purposes.

However, as mentioned above, this does not necessary change the level of security measures required for the processing according to the GDPR. The main consequence is therefore that Mobai for this processing solely can rely on a legal basis in Article 6(1). This is unless they are processing data falling under one or more of the other special categories of data mentioned in Article 9(1). Please note that the discussions in the sandbox project regarding Article 9 have been limited to biometric data as a special category of personal data. Other special categories of data will therefore not be addressed in this report. However, Mobai has to consider this before processing personal data for secondary purposes.

According to Article 9(4) GDPR "Member States may maintain or introduce further conditions, including limitations, with regard to the processing of (…) biometric data (…)". The Norwegian Personal Data Act section 12 sets out such additional conditions for the processing of "unique identifiers" such as biometric data:

> "National identity number and other unique identifiers can only be processed when there is a legitimate need for definite identification and the method is necessary to obtain such identification."

The impact of this national regulation on processing of biometric data for purposes related to training and improvement of AI models and the system as such, has not been further discussed in this sandbox project.

Central storage of biometric information

Processing of digital data that represent a feature of a physical person in its raw and unencrypted form (i.e., facial images, fingerprints and iris scans) represents a high risk to the individual. This is because the data is static and cannot be changed or replaced. Therefore, when processing facial images, as well as biometric templates, the data controller needs to ensure appropriate safeguards.

Centralized storage of biometric data under the GDPR is considered particularly sensitive and presents several challenges and obligations for organizations. Due to this sensitive nature, The European General Data Protection Regulation (GDPR), as well as the Norwegian Personal Data Act (section 12), emphasize the importance of strict safety measures in processing and storing biometric data.

The SALT solution has two different implementations concepts based on whether it considers end-user devices as trusted or untrusted:

- “Trusted devices” would enable local processing and local storage of biometric data
- “Untrusted devices” deems it necessary to distribute processing and storing biometric data across local devices and central systems

A primary justification for the centralized storage and processing of biometric data is that Mobai does not regard personal user devices as reliably "trusted" within their use case. This report focuses specifically on the concept of central storage and processing. However, it will also discuss the underlying rationale for centralized handling and the associated risks.

It is important to differentiate between two different types of storage of biometric data:

- Central storage, where a large amount of biometric data is aggregated in a database
- And decentralized, or local, storage, that usually happens on a users’ personal device, for example a mobile phone, PC or smartcard/hardware key.

Decentralized storage only encompasses the users’ own biometric data, captured locally and thus under physical control of the user, unless the platforms/device/service-provider also incorporate backup routines to also store the data externally. Such devices usually store local data in specially protected hardware components, like a TPM (Trusted Platform Module), or TEE (Trusted Execution Environment), which is offered in modern devices. Such solutions are already often used as an alternative for PIN codes for mobile phone screen locks.

There are many use cases for decentralized storage. One use case that many citizens already use, are different types of local device authentication or verification where a user authenticates themselves using the mobile phone’s own face verification system.

Key considerations with central storage of biometric data

Centralized storage of biometric data has several risks:

- Increased consequence of breaches: Centralizing biometric data means aggregating sensitive information in a central location or system. This creates a larger target and increased consequences from cyberattacks or unauthorized access. If a breach occurs, it could expose the biometric data of large number of individuals at once, leading to significant privacy violations.
- Illegitimate repurposing: Centralized storage makes it easier for biometric data to be used for unintended purposes. For example, if data is stored centrally without strict controls, it could be used for purposes beyond those originally intended or authorized by the data subjects.
- Loss of control by data subjects: Centralized storage can make it difficult for individuals to control their biometric data. If data is aggregated and stored centrally, individuals may find it challenging to exercise their GDPR rights, such as the right to access, rectify, or delete their data.

There are variations on storing biometric data centralized such as for instance distributing the database content across multiple databases or splitting parts of each data entry into multiple pieces for storage that can be distributed. These are mitigations that can reduce risk and consequences of central storage.

Locally stored biometric data is potentially also exposed to different types of attacks, however an important benefit of locally storing personal biometric data on user's personal devices is that it represents a lower risk for large scale confidentiality breach. It also makes it harder to systematically collect and repurpose biometric data at scale to be used for 1:n (one to many) identification purposes.

Thus, security requirements for centralized storage are comprehensive. Below, we outline the requirements associated with the various approaches:

- Encryption: Biometric data stored centrally must be encrypted both at rest (when stored), in transport and in use. This ensures that even if some types of unauthorized access occur, the data is not readable or usable. This requires proper and effective management of encryption keys. It also requires usage of efficient and secure encryption algorithms and key lengths.
- Access controls: Strict access controls must be in place to ensure that only authorized personnel can access the centralized biometric data. This includes multifactor authentication, role-based access, and regular audits of access logs.
- Pseudonymization: Where possible, biometric data should be pseudonymized before being stored centrally. Pseudonymization involves altering the data so it cannot be attributed to a specific individual without additional information.
- Data Minimization and Deletion: Organizations should collect and retain only the biometric data strictly necessary to fulfil the intended purpose. This principle mandates actively preventing the collection and storage of excessive data. Additionally, all data no longer required for the specified purpose must be promptly and properly deleted—including from backup storage—in accordance with policy guidelines grounded in the principles of necessity and proportionality.
- Regular security audits and demonstrating compliance: Organizations must regularly audit their security practices and infrastructure to ensure that the centralized storage of biometric data remains secure. This includes reviewing encryption methods, access controls, deletions procedures and any changes in the threat landscape.

Key concerns expressed with central storage of biometric data

There is considerable debate within the EU on the question of central storage of biometric data. The main questions of debate are whether it should be allowed at all, and, if yes, which specific actors should be allowed. There is no formally established consensus on these questions.

The European Data Protection Board (EDPB) has issued guidance stressing the importance of minimizing risks related to biometric data processing. While the EDPB does not outright ban the central storage of biometric data, it emphasizes strict conditions and often suggests decentralized approaches as a safer alternative. The EDPB advocates that biometric data should only be processed and stored in ways that are strictly necessary and proportionate to the purpose of the processing. Centralized storage is often seen as a last resort; only justifiable when no less risky alternatives are available.

The French national data protection authority (CNIL –Commission Nationale de l'Informatique et des Libertés) has been particularly vocal in warning against the central storage of biometric data. It advises that storing such data in a decentralized manner, or on individual devices, can significantly reduce privacy risks. Several German regional DPAs have taken a strong stance against the central storage of biometric data, particularly in the context of public administration and law enforcement. They argue that decentralized storage, combined with strong encryption and local processing (on devices), provides better protection against unauthorized access and misuse.

Certain members of the European Parliament (MEPs) and committees, particularly those focusing on civil liberties, have raised concerns about the central storage of biometric data, particularly in relation to state surveillance and mass data collection. There have been calls for stricter regulations or outright bans on central storage, especially when it comes to government databases. Civil Liberties Groups such as the European Digital Rights (EDRi) network advocate against the central storage of biometric data, arguing that it inherently creates vulnerabilities and risks that are difficult to mitigate. They push for decentralized, privacy-preserving alternatives, emphasizing the principles of data minimization and user control.

The European Court of Human Rights (ECHR) has influence on EU policies. In some cases, it has ruled against practices that involve centralized storage of sensitive personal data, including biometric data, particularly when such storage is not justified by a compelling public interest or lacks adequate safeguards.

While there is no absolute ban on the central storage of biometric data within the EU, there is a strong preference for decentralized approaches among many legal bodies, DPAs, and privacy advocates. These entities argue that centralized storage poses significant risks and that, where possible, biometric data should be stored in a way that minimizes these risks, such as on personal devices or in encrypted, decentralized systems.

Mobai's arguments for central storage of biometric information

The EU's technical specifications for eIDs prioritize multi-factor authentication, which emphasizes the importance of secure, controlled environments for biometric processing. This is outlined in standards like ETSI TS 119 461.

Mobai argues that performing biometric processing solely on devices (such as personal computers or mobile phones) presents security challenges. Arguably, although modern devices incorporate advanced security features, they are still susceptible to vulnerabilities and tampering, necessitating – in Mobai's view – trusted central environments for secure storage and processing.

Achieving a high security level thus requires combining device-specific safeguards with external processing, adding a protective layer that complicates unauthorized access for potential attackers. Increasing incidents of social manipulation, where users unknowingly disclose possession-based and knowledge-based credentials, pose substantial risks, particularly for banking accounts, eIDs, and other digital services. Effective biometric verification depends on user-specific biometric references that are stored and processed securely.

An important challenge is how to ensure the biometric reference's integrity throughout the comparison process. Many devices, including laptops and mobile phones, may be susceptible to tampering or unauthorized access. Smartphones store biometric data in secure containers protected by encryption and isolated environments, yet access to these containers is also platform-controlled (e.g., by Google, Apple, Microsoft, etc).

Mobai's solution aims to enhance protection for biometric data during both storage and processing by employing their own additional security mitigations, including homomorphic encryption. The Protected Biometric Template (PBT) provides privacy features beyond those found in standard IDs or smartphone biometrics. And by leveraging homomorphic encryption, Mobai supports properties like revocability, irreversibility, and unlinkability as per the requirements in ISO/IEC 24745:2022. Given the computational demands of homomorphic encryption, Mobai considers centralized processing the viable way to ensure trustworthy face matching results.

The PBT storage can be either centralized or decentralized. However, in Mobai's view, centralized storage offers significant security advantages, including faster updates and improved threat detection. The centralized approach also incorporates quantum-resistant encryption through SALT, whereas decentralized systems would rely on conventional encryption (meaning the need to decrypt data for pattern matching).

In summary, Mobai asserts that their implementation of centralized storage affords providers greater control of security, including key management, and of privacy using novel encryption methods. In contrast, a decentralized approach would leave service providers, like BankID, reliant on platform owners' security measures, with less direct control over PBT storage integrity, security and privacy.

The NDPA's assessment of the possibility for central storage

In this report, we have addressed key concerns with central storage of biometric data and presented Mobai's arguments for central storage. As we have pointed out, there is no absolute ban on central storage of biometric data and there is a path going forward for companies that aim to do this. What The Norwegian Data Protection Authority have pointed to in this report, is that a company must evaluate if it is necessary and proportional to store biometric data centrally and to thoroughly document the reasoning for this, including proper risk assessments of the specific protective mitigations.

In light of Mobai's solution, The Norwegian Data Protection Authority assess that they might enable the implementation of central biometric data storage and processing in cases where it was previously not considered secure enough to address the significant concerns associated with central storage.

Secondary purposes: Training and improving AI models and systems

Using data that represent a physical person, for the purposes of verifying a person's identity, requires the use of methods and algorithms that fall under the generic terms "AI and machine learning" (ML). ML models in their nature require training, quality control, tuning and adjustments to properly work over time and for improvement of quality and efficiency. Such continuous adjustments and improvements require access to training data.

To fulfil secondary purposes, Mobai stores encrypted facial images in a central database for a period of time limited by an internal policy. Mobai has indicated that they would only need a subset of images for training. The policy for what subset and percentage of data that is needed for training has not been specified, but the fraction, according to Mobai, could be as low as 10% of all images potentially available. However, the percentage could differ based on need, i.e., when there are situations or events that indicate a need to tune or train the models for specific reasons. An active and well documented practice for such data minimalization principles for model training will be important and central for the solution to be considered to have privacy by design and by default implemented.

The main difference from a technical point of view from the primary purpose is that the primary purpose includes several actors and usage of the biometric data across several organizational boundaries. That is also the reasoning for the described use of extended security measures (incl. homomorphic encryption) for the primary purpose.

For the secondary purpose, it is the understanding of the Norwegian Data Protection Authority that Mobai will not store biometric data. This is because facial images that are not specifically processed to become templates is not considered to be biometric data. However, the centrally stored data (facial images) do represent a similar risk as biometric data, as it represents a collection of data. The requirements for protecting this data are for that reason considerable. In this report, The Norwegian Data Protection Authority has not evaluated the specific security measures applied for storing data for secondary use.

The data controller should perform adequate risk assessment and implement adequate security measures for storing and processing facial images for the secondary purpose, including proper evaluation of necessity and proportionally.

The path ahead

SALT reflects an attempt to address the growing risks in digital identity verification solutions, including the increasing trend of social manipulation.

Going forward, geopolitical uncertainties, the rise of AI-driven fraud, and the potential of quantum computing (possibly weakening some widely used existing encryption algorithms) introduces new risks that call for the development of more secure and privacy-minded solutions to handle our digital identities.

Still, a more expansive use of biometric data in these solutions, also pose risks to the individual. The issue of centralized storage of biometric data, which we have addressed in this report, is subject to considerable debate. If a criminal actor would manage to break into a database that contains unprotected biometric data, it will not only be a case of theft, but they would potentially manage to acquire the entire digital persona of numerous persons.

This is why, while there is no absolute ban on the central storage of biometric data, most government authorities prefer decentralized approaches and assert the need for extraordinaire protective measures if centralized storage is used. However, there is a path going forward for companies that aim to store biometric data centrally. What the Norwegian Data Protection Authority have pointed to in this report, is that the company must evaluate if it is necessary and proportional to store biometric data centrally and to thoroughly document the reasoning for this, including proper risk assessments of the specific protective mitigations.

Another important point of discussion in this report, is the use of relatively new and novel technology like homomorphic encryption. For SALT, this is a critical component in their solution and they present an innovative use case for the technology. The Norwegian Data Protection Authority sees its potential in enhancing the security of encrypted data, as well as its potential use in other areas that can benefit from analysing and processing data while preserving confidentiality.

Datatilsynets regulatoriske sandkasse

Besøksadresse:
Trelastgata 3, Oslo

Postadresse:
Postboks 458 Sentrum
0105 Oslo

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

datatilsynet.no/sandkasse
personvernbloggen.no