



Project no. ICT-2007-216339

TURBINE

TrUsted Revocable Biometric IdeNtitiEs

Grant agreement for: Large-scale integrating project (IP)

Theme 3: ICT - Information and Communication Technologies Secure, dependable and trusted infrastructures

D1.2.1

Services and schemes for multiple trusted identity

Due date of deliverable: M9

Actual submission date: M9

Start date of project: 1 February 2008

Duration: 36 months

Organisation name of lead contractor for this deliverable: KUL

Project co-funded by the European Commission within the Seventh Framework Programme (FP7/2007-2013)		
Dissemination Level		
PU	Public	X
PP	Restricted to other programme participants (including the Commission Services)	
RE	Restricted to a group specified by the consortium (including the Commission Services)	
CO	Confidential, only for members of the consortium (including the Commission Services)	

Table of Contents

Glossary	3
1. Executive Summary.....	6
2. Introduction	7
2.1 Project scope.....	7
2.2 Background and objectives of the report.....	7
2.3 Organization	8
3. Preliminaries	9
3.1 Building blocks.....	9
3.1.1 Logic parties.....	9
3.1.2 The authentication process.....	9
3.1.3 System implementation components.....	11
3.2 Identity management settings and system implementation models.....	11
3.2.1 Identity management settings.....	12
3.2.2 System implementation models.....	13
3.3 The trust model architecture.....	15
3.3.1 Public-key infrastructures	16
3.3.2 Certificates and protected templates	17
4. Enrolment	18
4.1 Identification	18
4.2 Privacy requirements.....	18
4.3 General scheme	19
5. Verification	22
5.1 Privacy requirements.....	22
5.2 General scheme	22
5.3 Identification scheme.....	24
6. Revocation.....	25
6.1 Privacy requirements.....	25
6.2 General scheme	25
7. Multiple identity management	27
7.1 Management of multiple identities on the same token	27
7.1.1 Identity and Service providers	27
7.1.2 Identity corruption	27
7.2 Multiple identity management scenario	28
8. Token characteristics	30
8.1 Token as trusted personal device	30
8.2 GlobalPlatform Security Domains for secure multiple pseudo identities using different applications.....	31
8.3 Draft GlobalPlatform based architecture for multiple TURBINE service providers	33
9. Duplicate enrolment check scenarios	35
9.1 A solution for DEC using a “weak link scenario”	35
10. Biometry trust management	38
10.1 Performance level Structure	38
10.1.1 Criteria on performance.....	38
10.1.2 Technical criteria.....	39
10.2 Services description	39

10.2.1	<i>Prescription control</i>	39
10.2.2	<i>Access control</i>	41
11.	Conclusion	43
12.	Bibliography	44
Appendix A – BIOSIG 2008 Paper		46
1.	Introduction	46
2.	Challenges	46
3.	Requirements for biometric templates	48
4.	Reference architecture	49
4.1	<i>Pseudo identities</i>	50
4.2	<i>Pseudo identity creation</i>	50
4.3	<i>Pseudo identity verification: PI recoder (PIR) approach</i>	52
4.4	<i>Pseudo identity verification: PI verification (PIV) approach</i>	52
4.5	<i>Pseudo Identity expiration</i>	52
4.6	<i>Pseudo Identity revocation</i>	53
5.	Architecture overview.....	53
6.	Conclusions	53
7.	Acknowledgments.....	54
	References.....	54

Glossary

<u>Term (Acronym)</u>	<u>Description</u>
Anonymity	Anonymity is being not identifiable within a set of subjects, the anonymity set (see [PRI-D14.1c]).
Anonymity set	The set of all possible subjects in a given data collection context (see [PRI-D14.1c]).
Auxiliary Data (AD)	Subject-dependent data that are required to recreate Pseudo Identities during verification or for verification in general. Auxiliary Data are part of a Protected Template but are not necessarily stored in the same place as Pseudo Identities. Auxiliary Data may contain data elements for diversification (i.e., Diversification Data).
Biometric characteristic	Biological and behavioural characteristic of an individual that can be detected and from which distinguishing, repeatable biometric features can be extracted for the purpose of automated recognition of individuals (see [SC37SD2]).
Biometric identity	In the scope of this document, a biometric identity is the combination of a pseudo-identity that is derived from a subject's biometric data, thereby implicitly referring to a protected template, and another non-biometric, e.g., civil, identity that is linked with it, if any.
Biometric reference	One or more stored biometric samples, biometric templates or biometric models attributed to a biometric data subject and used for comparison (see [SC37SD2]). EXAMPLE: Face image on a passport; Fingerprint minutiae template on a National ID card; Pseudo-Identity in a database.
Captured biometric sample	Analogue or digital representation of a biometric characteristic that is output of a biometric capture subsystem (see [SC37SD2]) or any data derived thereof, which uniquely characterize an individual within a set of all possible subjects in a given data collection context.
Certification Authority (CA)	
Certificate Revocation List (CRL)	
Civil Identity (civil ID)	Identity attributed to an individual by a state (e.g. name, date of birth, social security number) (see [PRI-D14.1c]).
Comparison (compare)	Estimation, calculation or measurement of similarity or dissimilarity between recognition biometric sample(s) and biometric reference(s) (see [SC37SD2]). NOTE: Match (v) is deprecated as a synonym to compare (v).
Controller	The natural or legal person, public authority, agency or any other body which alone or jointly with others determines the purposes and means of the processing of personal data (see art 2.d of [Dir9546EC]).
Data Protection Authority (/ies) (DPA)	
Diversification Data (DD)	Diversification Data represent those data elements of the Auxiliary Data that may be used for diversification to allow renewal of Pseudo Identities and to prevent database cross-matching.
Duplicate Enrolment Check (DEC)	
Entity authentication	Entity authentication (or identity verification) is the process whereby one party, the verifier, is assured, through acquisition of corroborative

	evidence, of the identity of a second party, the claimant, involved in a protocol, and that the second has actually participated, i.e., is active at, or immediately prior to, the time the evidence is acquired (see [HAC1997]).
Feature Extractor (FE)	
Global Platform (GP)	
Hybrid verification	Verification that requires the combination of a claimed identity, a biometric sample and (part of) a protected template stored locally on an asset. E.g., the pseudo-identity is stored in a central database and the auxiliary data are stored on a token held by the subject.
Identifiability	Identifiability is the possibility of being individualized within a set of subjects, the identifiability set (see [PRI-D14.1c]).
Identity	Any data that uniquely characterize a data subject within a given anonymity set (see [PRI-D14.1c]).
Identity Management (IDM)	
Identity Provider (IdP)	<p>An entity that provides identity services, by issuing credentials or assertions thereby making claims about another entity, or subject, and its attributes and thus allowing the other entity to manifest its identity (digitally), i.e., to convince a third entity, called the relying party, of its identity.</p> <p>In the context of the TURBINE project, identity checks are needed to make a link between the biometric pseudo-identity and another identity, e.g., the civil identity or some billable identity. Then, by asserting, explicitly or implicitly, that a protected template was properly created, an identity provider provides trust to a service provider. It is possible, but not required, that during the authentication of the subject, the identity provider and the service provider interact to validate the status of the identity, e.g., through some online or offline revocation status services offered by the identity provider.</p> <p>The distinction between subject, identity provider and service provider is a logical distinction.</p>
Identity Reference (IR)	The identity against which a claimed identity is matched during identity verification. This can be a civil identity, a pseudonym or a pseudo identity in case no link with another identity is required.
Identity verification	See entity authentication.
Integrated Secure Platform for Interactive tRusted pErsonal Devices (Inspired)	
Issuer Security Domain (ISD)	
Operating System (OS)	
Ownership verification	Verification by means of a protected biometric template that is stored on a certain personal asset, for example a smartcard, a paper ticket or magnetic stripe medium. The task is to verify the ownership of the asset. This process does not strictly require a claimed identity.
Privacy	The privacy is the protection of the personal data of the user. The access to those data must be controlled, restricted and protected. In systems based on biometric data for access control, the privacy also concerns those data.
Protected (biometric) Template (PT)	Combination of a Pseudo Identity and Auxiliary Data required for biometric verification that satisfies privacy and security requirements.

Pseudo Identity (PI)	Pseudo Identities represent an individual or data subject within a certain context by means of a protected binary identity derived from the Captured biometric sample. A Pseudo ID does not contain any information that allows retrieval of the original biometric measurement data, biometric template or true identity of its owner. The pseudo identity does not mean anything outside the service context.
Pseudo-Identity Comparator (PI)	
Pseudo-Identity Encoder (PIE)	
Personal Identification Number (PIN)	
Pseudo-Identity Recoder (PIR)	
Pseudo-Identity Verification/Verifier (PIV)	
Pseudonym	A pseudonym is an identifier of a subject other than the subject's civil identity (see [PRI-D14.1c]).
Public-Key Infrastructure (PKI)	
Recipient	A natural or legal person, public authority, agency or any other body to whom data are disclosed (see art 2.g of [Dir9546EC]).
Secret	The secret is a data that is used in a cryptographic algorithm. It may be, for instance, an encryption key.
Secure Channel Protocol (SCP)	
Service Provider (SP)	An entity that offers services to other entities, called service consumers, and that requires guarantees about the identity of the consumer with whom it is interacting to be able to personalize or to charge for the services. When a service provider relies on an identity provider to have this certainty of identity it is called a relying party. In the context of biometric authentication, strong guarantees imply assurance of the link between a physical and a logical identity.
(Data) Subject	An individual who can be identified, directly or indirectly, in particular by reference to an identification number or to one or more factors specific to his physical, physiological, mental, economic, cultural or social identity (see art 2.a of [Dir9546EC]). In the context of biometric authentication, the subject is a natural person using his biometrics to convince other entities of his identity.
Supplementary Data (SD)	Data intended for security amplification by means of possession, knowledge or application-based secrets that are both required during enrolment and verification and are not stored with the biometric template(s).
Supplementary Security Domain (SSD)	
Trusted Third Party (TTP)	
Universal Integrated Circuit Card (UICC)	

1. Executive Summary

Deliverable D1.2.1 of the TURBINE project presents the work carried out in work package 1.2, of which the objectives are to provide a generic definition of the different identity management services that are needed to manage multiple identities that are derived from fingerprint biometrics. The different services include schemes for enrolment or creation of biometric identities, for revocation of these identities and for verification, i.e., the use of these identities.

Before describing the general procedures, the different building blocks in identity management are introduced along with an overview of different system implementation models. Each scheme is then presented as a flow of transactions between a data subject, a service provider and an identity provider. The flows are preceded by privacy requirements that are relevant for the schemes.

The basic schemes are further studied in the context of managing multiple identities on the same token. As an example the scheme is used in a scenario of ePrescriptions. However, having multiple identities is not always desired and procedures for duplicate enrolment checks are elaborated. Particular attention is paid to the characteristics of a token carrying multiple identities and a draft architecture for multiple TURBINE service providers is proposed based on the GlobalPlatform specifications.

Finally, the notion of scalable trust and performance is addressed. A performance level structure is developed and then applied on two services that can benefit from such a structure: prescription control and access control. Although the scope is entirely different, similar performance levels can be specified.

As a reference, the paper "Reference Architecture for Biometric Template Protection based on Pseudo Identities" that was presented by the consortium at the BIOSIG workshop in Darmstadt, Germany, September 2008 is included in appendix. This paper summarizes deliverable D1.1.1 and reflects the work that was carried out in WP1.1 of the TURBINE project. This document, along with deliverable D1.1.1 and results from sub projects 2 and 3, will serve as input for work package 1.3 where the goal is to describe the functionality of the generic and the airport security demonstrator. Based on this description the architecture of the two demonstrators will be defined.

2. Introduction

2.1 Project scope

The **TURBINE** project “TrUsted Revocable Biometric IdeNtitiEs” proposes a multi-disciplinary privacy enhancing authentication technology. Based on innovative developments in cryptography and fingerprint biometrics, it aims to resolve the current privacy concerns regarding the use of fingerprint biometrics for ID management.

To achieve this it will develop and evaluate the foundation and application of revocable protected biometric templates and pseudo-identity bit-strings using fingerprint data. It will provide:

- Cryptographic techniques applied to fingerprint biometrics to obtain a non-invertible and protected;
- Pseudo-identity bit-string for enrolment and subsequent verification;
- Multiple re-generation of independent unique bit-strings based on the same fingerprint;
- Revocable and multiple pseudo-identity management scheme based on these unique bit-strings;
- Highly reliable biometric fingerprint 1:1 secure verifications using these unique bit-strings;
- Detailed verification performance analysis, evaluated against very large public and private fingerprint databases;
- Comprehensive risk analysis and system security;
- Contribution to developing international standards for biometric template protection.

Its primary objective is to develop and then demonstrate that the technology and its performance in practice are sufficiently mature for deployment¹ as a solution to large scale eID requirements. Expert groups will advise the consortium on i) data protection, privacy issues and ii) requirements of key application sectors for eID management solutions. Furthermore, a comprehensive verification test, demonstrator environment will evaluate how single fingerprint data of an individual may be used to generate several secure unique pseudo-identity bit-strings with different levels of trust. It will include revocation and issuance of an equivalent re-generated biometric identity based on the same specific fingerprint data without weakening the overall security.

2.2 Background and objectives of the report

This document presents the result of the work carried out in work package 1.2 of the **TURBINE** project that focuses on a **general scheme to generate and to revoke trusted identity**. The objective of this work package is to define a generic scheme that specifies identity management services based on the use of fingerprint biometrics. Several application profiles are analysed in a generic way targeting schemes that are easily transposable to other biometric modalities. The nature of the work is service oriented; for different scenarios the interaction between the involved parties and components are described and their roles are analyzed.

Work package 1.2 consists of three tasks.

- The **Creation and revocation** task addresses the basic procedures in biometric identity management, namely enrolment, verification and revocation, thereby covering the entire lifecycle of a biometric identity;
- The task on **Multiple identity management** studies how to manage multiple identities that are generated from the same characteristic and stored on a single token, in an attempt to prevent the privacy issue of cross-matching. It also addresses the inverse problem of duplicate enrolment checks;

¹ This objective is in line with the objectives of other privacy-oriented projects, such as PRIMELife. Similarly, TURBINE aims at delivering new privacy-enhancing mechanisms that allow to establish trust infrastructures for identity management in real life [PRIMELife].

- The **Biometry trust management** task specifies services that obtain stronger accuracy by using multiple fingers or other biometric data in the same authentication session.

This deliverable integrates the results of these tasks.

2.3 Organization

The next section (section 3) defines the model for biometric authentication and introduces some of the building blocks (subject, identity provider, service provider...) that recur in the presented schemes. The authentication process is formally defined and placed in the context of biometric verification. Different identity management settings are described and the trust model that is adopted in **TURBINE** is positioned against the PKI model.

Sections 4, 5 and 6 present general schemes for the basic identity management procedures: enrolment or creation of the biometric identity, verification or use of the identity, and revocation. Each scheme indicates the role of the involved parties, the data that is exchanged and the flow of transactions. Attention is given in particular to the protected biometric template and the characterisation of the token on which it is stored.

Section 7 tackles one of the main challenges in the **TURBINE** project: the management of multiple identities derived from the same characteristic of a single subject. At each service provider, a subject will use a different pseudo-identity as a countermeasure to the risk of cross-matching data from different applications. One of the options in **TURBINE** is to store the protected templates on a single token. The required characteristics of such a token are described in section 8. Despite the privacy-enhancing measure of using different identities with different service providers, multiple identities may allow identity fraud. For services like visa applications duplicate enrolment checks are needed, which in turn may require central storage of biometric data. These issues are addressed in section 9.

The final objective of scalable trust and reliability is discussed in section 10, where a performance structure using multiple fingers is specified for two practical applications, namely prescription control and access control.

Appendix A of this document contains the text of a paper that was accepted and presented at the BIOSIG workshop in Darmstadt, Germany, September 2008 (see [Biosig08]). This paper presents challenges and requirements on biometric template protection methods along with a technology-neutral description of a reference architecture. It summarizes deliverable D1.1.1 ([D1.1.1]) and reflects the work that was carried out in WP1.1 of the **TURBINE** project.

3. Preliminaries

In the schemes presented after this section, a number of building blocks will recur. We briefly introduce them and elaborate on some aspects that are important for the remainder of this document.

3.1 Building blocks

3.1.1 Logic parties

In identity management a logical distinction is made between three entities:

1. Subjects or service consumers, e.g., travellers, online customers,
2. Service providers or relying parties, e.g., airline companies, online shops, and,
3. Identity providers, e.g., governments issuing passports, credit card companies.

Subject

In the context of biometric authentication or identification a subject is a natural person that is interested in using some services and that is willing to use his biometrics to get them. The subject can also be obliged to use certain services in order to fulfil his duties, e.g., tax declaration.

Service provider (SP)

A service provider offers services or benefits to subjects (or users). To personalize the services or to be able to charge a subject, a service provider will require strong guarantees about the identity of the entity, i.e., the subject, with whom it is interacting. To have this certainty of identity it may rely on a third party that provides identity services; hence, a service provider is often also called the relying party.

Identity provider (IdP)

An identity provider provides identity services: it issues (digital) identities and makes it possible for a subject to manifest its identity (digitally) when interacting with service providers. It does this by issuing credentials or assertions to a subject thereby making claims about the subject and its attributes. With these credentials or assertions a subject can convince a service provider of its identity and claimed attributes. The identity provider also takes care of the registration of the subject and the identity checks that are required to issue a digital identity.

In the context of the **TURBINE** project, identity checks are needed to make a link between the biometric pseudo identity and another identity, e.g., the civil identity or some billable identity. Then, by asserting, explicitly or implicitly, that a protected template was properly created, an identity provider provides trust to a service provider. It is possible, but not required, that during the authentication of the subject, the identity provider and the service provider interact to validate the status of the identity, e.g., through some online revocation services offered by the identity provider.

3.1.2 The authentication process

A subject will have to authenticate itself towards a service provider when it wants to invoke some service. To be able to position the role of the different components we adopt the definition for entity authentication as given in [HAC1997]:

Entity authentication (or identity verification) is the process whereby one party, the verifier, is assured (through acquisition of corroborative evidence) of the identity of a second party, the claimant, involved in a protocol, and that the second has actually participated (i.e., is active at, or immediately prior to, the time the evidence is acquired).

Authentication protocols have evolved from static shared-secret schemes, e.g., based on passwords, to challenge-response protocols, where the claimant proves knowledge of a secret without revealing it. The security of those protocols is based on factors such as something you know (passwords) or something you possess. Biometrics take this one step further by basing the security of the protocol on something inherent to the claimant (something you are). On the one hand, they provide stronger guarantees about the link with the physical identity of the subject. On the other hand, biometrics are the only factor that allows efficient identification of a subject among a set of subjects. This will lead to the possibility of duplicate enrolment checks.

Reference data for verification

All authentication schemes have in common that the verifier needs to be able to corroborate the evidence provided by the claimant. Validation of this evidence can only occur if there are some reference (verification) data, e.g., the hash of a password, a shared secret key, a public key to verify digital signatures, or in the case of biometrics a reference template.

The role of the protected template is crucial. The verifier, or service provider, requires that the authenticity of the protected template is guaranteed when it is used as reference for verification. This can happen in two ways; either the service provider stores the protected template securely, or it is given by the claimant but then the service provider will use some mechanism such as a digital signature on the template to validate it. The service provider will use the claimed identity as a reference to look up the protected template and this reference is called the identity reference (IR). It can be a civil identity, a pseudonym or a pseudo identity, which is a special pseudonym generated from the biometric data.

Enrolment: creation and registration

Also common in all authentication protocols, with the exception of identity based techniques, is the execution of an initial one-time set-up or enrolment. During the set-up some parameters will be selected for the system and the subject, and the reference data will be created. Depending on the application, these actions may be performed by different parties.

Enrolment of a subject is the entire process of actions that are required before a subject and a service provider can execute an (automated) authentication protocol. This includes at least two sub processes, namely the creation of reference data and the registration thereof.

Registration

Within the scope of the **TURBINE** project the reference data is a protected template that includes the so-called pseudo-identity. The registration may include the following steps:

- A check of some other identity and the linking thereof with the pseudo-identity;
- Transfer of the protected template to the service provider;
- Authentication of the protected template by a trusted party (the IdP) to prevent forgery.

The link with another identity is required for identification purposes. Identification refers to the ability of distinguishing subjects and their related data. Because the pseudo-identity only identifies the protected template, it may be required for billing purposes or liability reasons, that the pseudo-identity is linked with another identity, e.g., the civil identity of the subject or a financial identity (credit card number). This other identity can be the identity reference, i.e., the identity against which the claimed identity is matched, but it is not required. The service provider can work with a pseudonym or the pseudo identity until abuse is suspected. Then the service provider will request the identity provider to link back to the other, e.g., civil, identity.

The party that registers the pseudo-identity is the authentic source of information regarding the status of the pseudo-identity (and the corresponding template).

Revocation

The counterpart of enrolment in the pseudo-identity lifecycle is revocation and renewal of the pseudo-identity. There can be no revocation without registration and the actual revocation, i.e., the blacklisting of the pseudo-identity or the replacement of the protected template, can only be performed at the party where the registration occurred.

3.1.3 System implementation components

We distinguish between the data that is stored in the template protection system and the components that deal with them. A protected template includes a pseudo identity (PI) and auxiliary data (AD). As described in section 3.1.2, the pseudo identity is linked to an identity reference. The definition of identity reference as proposed in [SC27BTP] can include three types:

- Civil identity, i.e., date of birth, name, address etc. as registered in a civil register (civil law countries) or by a unique number issued by the State, e.g., a social security number. A civil identity is valid by law and thus most privacy sensitive in biometric applications;
- Pseudonym, i.e., nickname, client or card number, etc. A pseudonym can be limited to one application and is considered to be more privacy-friendly than a civil identity;
- Pseudo identity (PI), i.e., an identity derived from biometric data and part of a protected biometric template. A pseudo identity can be seen as a special pseudonym with high entropy for discrimination and security. If the identity reference is a pseudo-identity then the protected template is simply not linked to another identity. In the analysis of the system implementation models (section 3.2.2), we will not consider this case.

The different functional components that generate or process these data are described in the reference architecture of [D1.1.1] (see Appendix A). We briefly review their roles in the TURBINE setting.

During enrolment a biometric reference is created as follows. A biometric sample is captured from the subject, e.g., a fingerprint image, and then processed by a feature extractor (FE), e.g., to produce a minutiae template. Then a pseudo identity and, possibly, additional auxiliary data are generated by a pseudo-identity encoder (PIE).

During verification, features are again extracted from a captured sample and then passed on either to a pseudo-identity recoder (PIR), or to a pseudo-identity verifier (PIV). A pseudo-identity recoder recreates a pseudo identity from the extracted features that is then compared by the pseudo-identity comparator (PIC) with the pseudo-identity that is stored in the protected reference template. The pseudo-identity verifier will not regenerate the pseudo-identity, but will verify the claimed pseudo-identity directly using the presented biometric data and the protected template.

Concerning system implementation, the location where the protected template and identity reference, civil identity or pseudonym, are stored and where data processing and verification happen can be described as follows [SC27BTP]:

- **Server**, a centrally located computer supplying services. This computer is remotely connected with the client via a network, and is sometimes referred to as a “biometric authentication server”;
- **Client**, a remotely (and/or distributed) connected computer with the server to request services. Generally this represents a PC or its equivalent executing a general purpose operating system which may exist in the form of a KIOSK. A biometric sensor unit is usually connected to the client. PDAs and certain smart mobile phones can be considered clients.
- **Token**, a physical device supporting for instance biometric storage and comparison. Generally this is an embedded device for storing data such as USB memory sticks or RFID tags in the e-passport and possibly for running the comparison programs such as Match-On-Card applications.

3.2 Identity management settings and system implementation models

To illustrate the different possible identity management settings and the role of the various components in the system implementation models, some examples are presented.

3.2.1 Identity management settings²

Open or third party-managed setting

In the third party setting, the identity provider is a trusted third party that is not the same as the service provider. Multiple service providers can rely on the same identity provider. This setting is the closest to the public-key infrastructure (PKI) model and is applicable to all scenarios where a central administration, such as the government, issues digital identities that are linked to some other, e.g., civil, identity.

A good example that works in this setting is prescription control. Consider a subject that is registered to an eHealth service that proposes an electronic prescription service. To get his/her medicines the subject needs to authenticate at the pharmacy. As will be described further on, the authentication process follows a Performance Level Structure. The required confidence in the subject's identity depends on the medicine he/she is getting.

The identity provider is a national health care administration and the pharmacists are the service providers. Clearly, there will not be a different biometric identity for each pharmacist and it is unlikely that all protected templates will be stored in a central database. The subject will carry a token with him that contains his biometric identity related to the prescription service and the protected template stored in the token will be authenticated by the identity provider, who will link it with the subject's identity by which it is known in the health administration. It is the task of this administration to check the subject's civil or health care identity and to ensure a proper link with the biometric identity.

In this setting, the subject often cannot choose any of the parameters or techniques with which his biometrics are protected. It is important to note that for applications where the civil identity is involved, the link between the civil identity and the pseudo identity must be known to the IdP or the SP for auditing. The subject may choose some parameters as long as this does not change the PI that was issued by the central administration, unless the chosen parameters are sent to and known by the IdP. In applications for duplicate enrolment checks, the subject cannot make any choices.

The setting of having one pseudo-identity for different service providers does not necessarily conflict with the privacy requirements as laid down in [D1.1.1]. For example, efficient eGovernment requires the ability to link data about a subject for a particular purpose. However, for different purposes, and thus in different contexts such as finance or ePassports, other pseudo-identities will be used.

Closed or organization-managed setting

In the organizational setting, the service provider and the identity provider are the same. The service provider usually has already checked the subject's identity, e.g., because it is an employee of the organization. When creating a protected template for a subject's biometric identity, the company has complete control over the template creation. The pseudo-identity is usually linked with a company-specific pseudonym, such as the employee number. An example of this setting is found in access control, e.g., in airports, or in banking applications. Because the template was created by the service provider himself, there is no need for a PKI-based trust infrastructure if the templates are stored securely and the service provider relies on the security of the storage facilities. If the templates are partially stored on the client side, it is still needed to check the validity of the templates.

User-centric or anonymous settings

This final setting is the one that gives maximum control to the subject and preserves its privacy the most. Simple examples include logical access control to computers or physical access control, i.e.,

² The settings overlap to some extent with the types of identity management systems as defined in the FIDIS project [FID-D2.3]. A first type is defined as an identity management system for account management in organizational systems, e.g., identities that are assigned by the state. Another type is an identity management system for user-controlled context-dependent role and pseudonym management, i.e., the subject can choose his identity. Both types are possible in a closed environment, e.g., within a company, or an open environment, e.g., free web applications.

biometric locks that work in identification mode. Applicability is limited to some applications, particularly when complete anonymity is required. An example of a commercial application that does not require a link to a credit card number or financial identity is coin-based vending machines that keep track of your credit. Upon enrolment a balance is associated with a subject's pseudo-identity. The machine then offers two functions; increase balance and sell product. The first will increase the balance after successful authentication and having inserted coins. The second will provide a product and decrease the subject's balance after successful authentication.

3.2.2 System implementation models

This section analyzes some facets that need to be considered before an actual protected template biometric system is implemented. The analysis is based on the system implementation components described in Section 3.1.3, and focuses in particular on the location of the different data and functional components, which should be carefully considered before deployment.

According to the storage and data processing configuration, the different biometric system components in a biometric template protection system can be configured into different models that are suitable for different identity management settings. Adopting the 6 application models proposed in [SC27BTP], we present different classes of implementation models for a biometric template protection system in Table 1. The list of implementation models presented here is not exhaustive. However, the chosen models cover most practical applications and those relevant within the scope of TURBINE. A more fine-grained distribution of the components, e.g., PIC on server and PIR on client, or more enhanced cases, e.g., where data are split "securely" among different locations, are possible, but the actual distribution depends on the target application.

These models include the data and the components that were defined in section 3.1.3, where the identity reference is redefined³ here to be either a civil identity or a pseudonym (we do not consider the case where the identity reference is the pseudo identity). In addition we discuss the implications of these models on the required security measures, on the privacy of the subject and on the resources that are needed.

In Table 1, Server, Client and Token are as defined in section 3.1.3. In most cases, use and storage of the IR depends on the application. Storage includes various situations in which PI, AD and IR (excluding PI) are distributed over Server, Client or Token; and functionality includes various situations in which FE, PIR, and PIC are distributed over Server, Client or Token.

Table 1: Implementation models of biometric template protection systems

Model⁴	Storage	Functionality
Class A: CM	PI, AD and IR on server	FE on client PIR and PIC on server
Class B: CMDS	PI and IR on token or server AD on token	FE on client PIR and PIC on server
Class C: CMLF	PI, AD and IR on server	FE, PIR and PIC on client
Class D: LM	PI, AD and IR on client	FE, PIR and PIC on client
Class E: LMDS	PI and IR on token or client AD on token	FE, PIR and PIC on client
Class F: LMDSF	PI and IR on token or client AD on token	FE and PIR on client PIC on token

³ Here the redefinition of IR is valid only for illustration of the system implementation models. In cases where a PI acts as an IR, the IR can be omitted in all classes in Table 1.

⁴ CM: central model; CMDS: central model with distributed storage; CMLF: central model with local functionality; LM: local model; LMDS: local model with distributed storage; LMDSF: local model with distributed storage and functionality.

Class A (CM): Central Model

Security

Database and network security required

Privacy

Privacy concerned for: central storage / processing; subject having no control of biometric data

Resource

Storage and processing resources required from the Server but not from the Client, except for the feature extraction.

Application Scenarios

This class of models is suitable for the open or third party-managed identity setting and the closed or organization-managed setting, in which there are identification and verification applications without token and without storage / processing resources in Client.

Class B (CMDS): Central Model with Distributed Storage

Security

Network security required

Privacy

There is no privacy issue for the case of storage in Token, except for the processing and the matching, which still occur on the server.

For other cases of storage in both Token and Server, privacy is concerned for central storage

Resource

Processing resource required from Server; storage resource required from Token and Server.

Application Scenarios

Models where the PI (and IR) is stored along with the PT on a token are suitable for the user-centric or anonymous settings; Models where the PI (and IR) is stored on the server are suitable for the open or third party-managed identity setting and the closed or organization-managed setting. In both cases, there are identification and verification applications with token but without storage / processing resources in Client.

Class C (CMLF): Central Model with Local Functionality

Security

Database and network security required

Privacy

Privacy concerned for: central storage / processing; subject having no control of biometric data

Resource

Storage resource required from Server; and processing resource required from Client

Application Scenarios

This class of models is suitable for the open or third party-managed identity setting and the closed or organization-managed setting, in which there are identification and verification applications without token; without storage resources but with processing resources in Client.

Class D (LM): Local Model

Security

Database security required

Privacy

Privacy concerned for: central storage / processing; subject having no control of biometric data, but better than other models using a centralized database

Resource

Storage and processing resources required from Client

Application Scenarios

This class of models is suitable for the open or third party-managed identity setting, the closed or organization-managed setting and also the user-centric or anonymous settings, in which there are identification and verification applications without token; with storage resources and processing resources in Client.

Class E (LMDS): Local Model with Distributed Storage

Security

Burden of both database and network security can be minimized.

Privacy

For the case of storage in Token, privacy is improved.

For other cases of storage in both Token and Client, privacy is concerned for: central storage

Resource

Processing resource required from Client; storage resource required from Token, or Token and Server.

Application Scenarios

Models where the PI (and IR) are stored along with the PT on a token are suitable for the user-centric or anonymous settings, models where the PI (and IR) is stored on the server are suitable for the open or third party-managed identity setting, the closed or organization-managed setting and also the user-centric or anonymous settings, in which there are identification and verification applications with token and with storage / processing resources in Client instead of Server.

Model F (LMDSF): Local Model with Distributed Storage and Functionality

Security

Burden of both database and network security can be minimized.

Privacy

For the case of storage and matching in Token, privacy is well guaranteed.

For other cases of storage in both Token and Client, privacy is concerned for: central storage

Resource

Processing resources are required from Client and Token; storage resource required from Token, or Token and Server.

Application Scenarios

Models where the PI (and IR) are stored along with the PT on a token are suitable for the user-centric or anonymous settings, models where the PI (and IR) is stored on the server are suitable for the open or third party-managed identity setting, the closed or organization-managed setting and also the user-centric or anonymous settings, in which there are identification and verification applications with token and with storage / processing resources in Client instead of Server. Applications requiring better privacy can be achieved by deploying matching in Token.

3.3 The trust model architecture

We now reconsider the involved parties and define a formal trust model. The trust model used within **TURBINE** is a set of defined relationships between a Subject, an Identity Provider and a Service Provider. Figure 1 illustrates the trust model architecture. The arrows reflect the different relations and do not necessarily imply real data flows.

The subject wants to use a service but wants to be absolutely sure that his data are protected and that the system is secure (Figure 1, arrow 1).

The Identity Provider (IdP) is a system entity that creates, maintains and manages identity information on behalf of subjects (Figure 1, arrow 3) and provides assertions of subjects authenticity to other providers (service providers for instance). It authenticates subjects and releases selected information about them. It can be an automated process or set of processes, a

subsystem, a person or group of persons. It verifies and validates the authenticity and the integrity of the subject's data and identity. The IdP provides identity credentials that will be eventually presented to the Service Providers.

The Service Provider (SP) manages and distributes services and solutions to subjects. The SP doesn't deal with identities but relies on an Identity Provider and will require the subject to register with the IdP chosen by the SP (Figure 1, arrow 2). It deals with identity credentials provided by the IdP, e.g., assertions, or the subject (with his/her token) (Figure 1, arrow 4). It wants to be sure that the subject is each time who he claims to be without knowing who he actually is.

When an IdP and a SP start a relationship based on trust, they first define a set of legal and technical requirements (Figure 1, arrow 5) that will define these relationships. Among the technical requirements, there are the prerequisites the IdP requires to securely communicate with the SP and to assert the subject's identity, while still preserving the subject's privacy. This means that the IdP requires data specific to the SP to create identity credentials that can be read and validated by the SP.

Among the services and the subject there may be two different kinds of relationships:

- An authentication relationship which is a relationship between an authentication infrastructure provider and an entity registering in order to be authenticable through SPs.
- A business relationship that gathers all business transactions (information exchanges, transactions such as voting, prescription management, access to a restricted area).

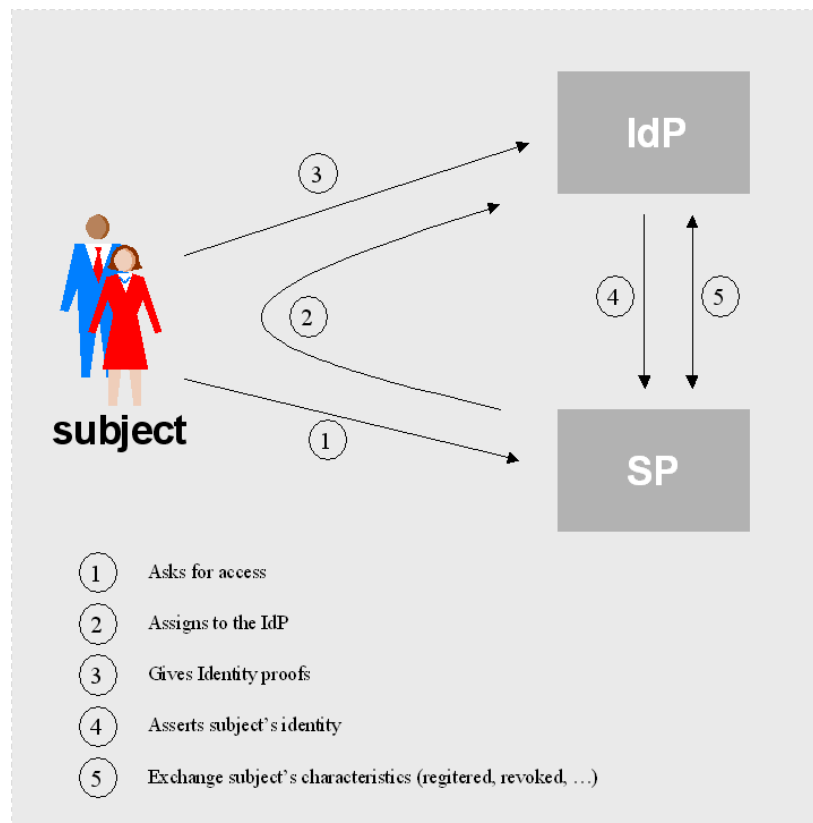


Figure 1: Trust model architecture.

The trust model architecture is used to define the general scenario flows.

3.3.1 Public-key infrastructures

As indicated already, the open or third party-managed setting easily maps on the PKI model. Public-key infrastructures rose with the development of asymmetric cryptographic techniques.

Public-key cryptography suggested that the problem of having to exchange shared keys securely was solved. A secure channel was no longer needed because symmetric keys were replaced with key pairs that consisted of a private (signing or decryption) key and a public (verification or encryption) key. Unfortunately, a party A that wishes to communicate with a party B still needs assurance that a public key really corresponds to the private key held by B and that a genuine copy was retrieved.

The most common solutions nowadays are based on public-key certificates. A public-key certificate is a data structure that stores the public key of an entity and some information that uniquely identifies the entity. These data are then signed by a trusted third party (TTP). Additional information may be stored in the certificate such as a validity period or expiry date. A party that receives a public key, e.g., along with a digital signature, only needs to store a copy of the public key of the trusted third party to be able to verify the integrity and authenticity of the received public-key certificate and contained public-key. The trusted third party is the identity provider and is called a certification authority (CA). It is the single point of trust for the service providers. The standard format for public-key certificates is the X.509 standard [RFC2459].

In case a private key is lost or stolen, the corresponding public-key certificate needs to be revoked. The X.509 standard also defines a format for certificate revocation lists (CRLs). This list serves as a blacklist in which serial numbers of certificates are contained. The list is updated on a regular basis and published by the certification authority. It is also possible that a certification authority provides an online service to check the validity of a certificate. A protocol for this kind of service has been defined and is called the Online Certificate Status Protocol (OCSP) [RFC2560].

3.3.2 Certificates and protected templates

Digital signatures are often used in authentication schemes and compared to biometric authentication, the public key can be considered as the equivalent of the biometric reference data. Without the reference data identity verification is not possible. The certificate model and its structure can be applied to biometric template protection and derived pseudo-identities. In the presented application of the prescription control, the eHealth administration will set up a certification authority that, instead of binding an identity with a public key, certifies protected templates and asserts their link with another identity. Particularly for the purposes of binding an entity's identity and some of his attributes another type of certificates has been defined, namely attribute certificates [RFC3281]. The advantage of this approach is that the revocation services (CRL or OCSP) can be adopted without further modification.

For privacy reasons it might be undesired that the identity linked to the protected template or pseudo-identity is referred to in the certificate. This can be solved by including a pseudonym in the certificate that only makes sense for one particular service provider. The complexity, however, increases at the side of the certification authority (identity provider), which now has to offer a conversion service, with possibly extended access control. E.g., for liability reasons legislation might require the possibility to link the pseudonym with the civil identity.

4. Enrolment

As indicated in the previous section, registration of a pseudo-identity depends on a subject's ability of being identified in a system. In other words, a service provider must be able to distinguish a subject from others and if needed, link to another identity.

4.1 Identification

There are various ways of identifying subjects in an identity management system:

1. **Identification** understood as uniquely or globally identifying a data subject:
 - a. Based on the subject's civil identity;
 - b. Based on any other factor, including client or card number, location data, number code, physical characteristics, etc.
2. **Pseudonymity** is understood as *a mechanism* whereby pseudonyms are used as identifiers instead of the use of an identifier that refers directly to the civil identity or the data subject. Pseudonyms can prevent linking data or transactions across different contexts or applications. They provide varying degrees of accountability depending on the other identity they are derived from and who is able to link back to it. A pseudonym can be created:
 - a. From a civil identity by a controller or by the data subject itself (type 1);
 - b. From any other factor, including client or card number, (type 2); or
 - c. From an anonymous credential by a controller or by the data subject itself, i.e., a claim of the data subject that can be verified in various ways (based on ID, e.g., being over 18 years old or by review of a register, e.g., the list of employees or the list of doctors) without referring to a specific data subject (type 3).
3. **Anonymity** is understood as that one cannot sufficiently identify the data subject within a set of subjects while the data subject remains accountable. This can be achieved:
 - a. Based on anonymous credentials (e.g., see [IDEMIX], [PRlide08]).

4.2 Privacy requirements

1. The use of pseudonyms as a privacy enhancing mechanism in **TURBINE** means that **TURBINE** could reach accountability but also anonymity, i.e., the pseudonym cannot be linked back to a single subject (type 3), depending on how the pseudonym is created (from a civil ID, any other identifying factor or a credential) and used.
2. As a privacy-enhancing IDM solution, **TURBINE** should specify and use all three types of pseudonyms, as identity reference, for the specification of the requirements for the enrolment and the creation in particular applications.

For specific service profiles, a specific pseudonym type will be required depending on the application. For example, for eGovernment and eVoting whereby a check of the voter against the voters list (based on the registration in the civil registers) will be necessary, a pseudonym type 1 will be required. For other eGovernment applications, a pseudonym type 3 may be required (e.g. for participating in polls). For eHealth and banking applications, a pseudonym type 1 to identify the patient/client may also be required.

3. For the use of pseudonym type 1, a proper *authentication procedure* for identification and verification of the civil identity, possibly in combination with identity certificates, will have to be defined. In addition, a *legal basis* for the identification and the verification of the civil identity is in principle required.
4. Because of the risks that the pseudonym as identifier could be used to link all the information available for that identifier, especially if this would be available to a

government, **TURBINE** may want to provide specifications for *sector-specific or context-specific identifiers* for one particular service profile (e.g., eGovernment). This means that different pseudonyms, possibly derived from the same data or identity, should be used for different applications.

5. In order to enhance security, not only the use of encrypted data, but also the use of encoded data is often recommended or required, i.e. the personal data (e.g., the biometric data) with a *reference code* instead of the name of the individual. In case of research projects, the use of *encoded data*, if no anonymous data can be used, is sometimes even required as a principle by law. The reference code will somewhere (offline or online, in the same or preferably another database) be linked or linkable to an individual, and for the IDM system, this would be the pseudo-identity.
6. The collection, the processing and the storage (including retention thereof) of any personal data shall in an application, including at registration, at all times be 'adequate, relevant and not excessive'. In practice, this means that personal data shall be *limited* to a minimum and only insofar as required for the purpose and finality of the application (Art. 6.1.c of [Dir9546EC]). For the IDM system, this implies that the needs for specific personal data related with the IDM in general, such as for the profile, the smart card registration and use, and with the setup and deployment of the pseudo-identities, shall be limited to a strict minimum.
7. The *place of storage* of the pseudonym (local or central) is crucial in a privacy enhancing IDM and either choice shall be carefully defined and motivated.
8. *Consent* of the data subject shall in most cases be obtained. The subject must have the possibility to withdraw his consent, i.e., consent is not given for life.1.

4.3 General scheme

The basic flow for enrolment is depicted in Figure 2 and goes as follows.

1. The subject wants to register and presents identifying attributes (for instance, elements of his civil identity such as name, date of birth) to the system.
 - The identifying attributes will determine the possibilities for the identity reference to which the pseudo identity will be linked. Ideally the identity reference is a pseudonym, even when full accountability or linkability to the civil identity is desired. The required attributes depend on the application profile (purpose and finality) and must be limited to a minimum.
 - Preceding this step, the subject's consent to use his biometrics in the application for which he is enrolling, must be obtained. This use is not necessarily confined to one application or one service provider.
2. The IdP validates the subject's identity by means of a well-defined procedure.
 - A proper authentication procedure should be in place in case the identity reference is the civil identity or a pseudonym derived from it. Access to this procedure should be available in the enrolment system.
3. The IdP notifies the subject of a successful validation and requests biometric data.
 - The amount of biometric data that is requested, e.g., one or two fingerprints, should also be limited to a minimum.
4. The subject presents a biometric sample.
 - Any biometric system should deploy mechanisms, e.g., encryption, that protect the fresh biometric data that is transferred from the sensor to the place where the PT will eventually be created or stored.
5. The IdP creates a protected template (PT) that contains a pseudo identity (PI) and, possibly, auxiliary data (AD). Both are derived from the biometric sample and are specific to the SP. The captured biometric data are deleted.
6. The IdP sends the user his PI and sends a certificate to the SP or the user as a proof of the Protected Template authenticity and of the link between the PT and the identity reference.

Depending on the procedure, the AD are sent to the subject and/or to the SP. The PI may also be sent to the SP.

- The IdP certifies the link between the PT and the identity reference.
- As indicated in section 3.2.2, there are different possible places to store the PT, including the PI, and the IR. The location depends on the actual deployment and should be carefully considered, in particular when the IR is a pseudonym with a high linking potential.

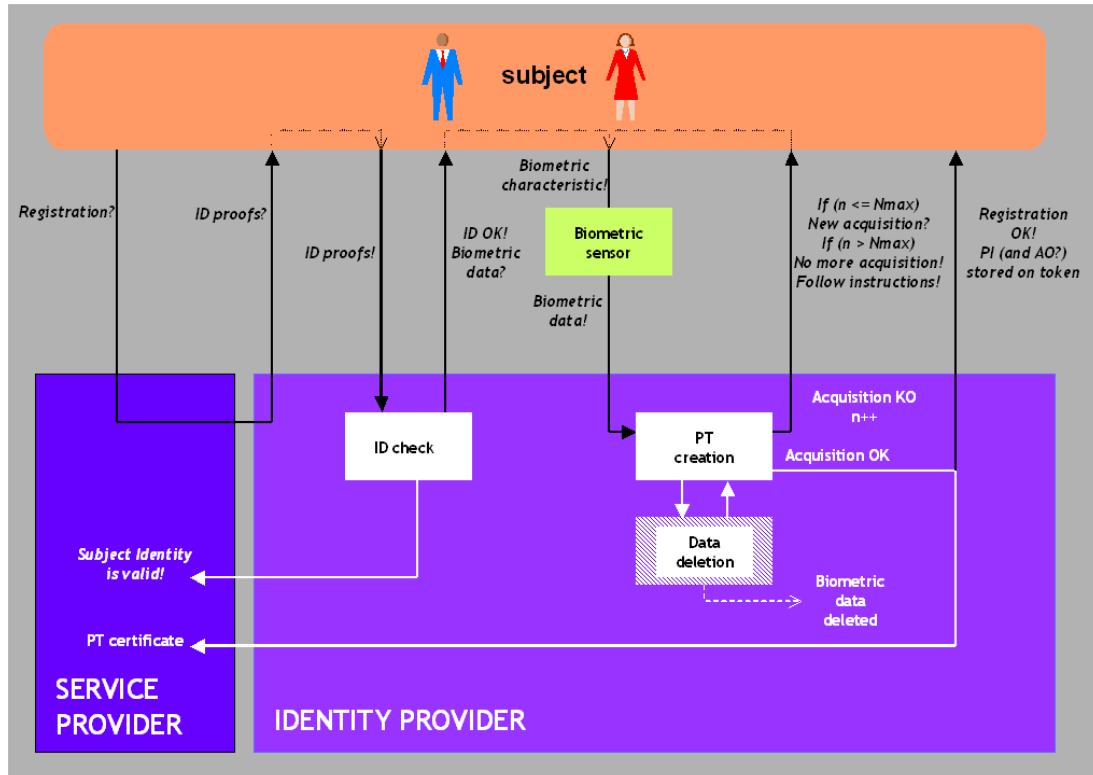


Figure 2: Creation flow.

Comments

- If the subject wants to create another PT based on the same biometric characteristic, for another service functionality or identity status, the validation of the subject's identity can be replaced by an identity verification procedure.
- If the identity or the proofs are already used in a previous registration, it can be needed that the service informs the subject and gives instructions for further investigations on a possible ID corruption.
- If Identity validation fails (invalid identifying attributes, identity duplication, registration not authorized) the IdP rejects the subject, signals an error and gives instructions for an alternative procedure, outside the service scope, via the service.
- If the biometric data check fails more than the maximum number of tries authorized, specific instructions are given in case none of the tries succeeded. At each new try, instructions for a better acquisition are given (ex: "Press the finger more", "shift the finger to the bottom", etc.).
- In case of duplicate enrolment the biometric characteristic is already used for another identity. If the system supports detection of duplicate enrolments, the IdP rejects the subject, signals an error and gives instructions for an alternative procedure, outside the service scope.

- In the case of a self-controlled environment the subject can decide where to store the created PT.
- If the IdP and the SP are the same entity, no certificate needs to be created. If not, then the IdP should provide mechanisms to verify the certified link between the PT and the IR.
- An SP can rely on multiple identity providers or can accept subjects that are registered with an IdP other than the one suggested by the SP. It is then required that the SP and the other IdP agree upon some policy, which implies the presence of an infrastructure supporting identity federation⁵.

⁵ The STORK project (Secure idenTity acrOss boRders linKed) is a Large Scale Pilot (LSP) that tries to establish identity federation at the European level by ensuring cross-border recognition of national electronic identity (eID) systems and enabling easy access to public services in 13 Member States without replacing the existing national schemes [STORK].

5. Verification

5.1 Privacy requirements

1. The privacy requirements for the enrolment phase (Section 4.2) also hold for the verification phase. More in particular, the use of means to protect the fresh personal data is required, as for instance encryption.
2. Directive 95/46/EC⁶ on the protection of personal data (see [Dir9546EC]) imposes a *detailed information* obligation upon the controllers towards the data subject (Art. 10 & 11). The latter shall be informed *inter alia* about the identity of the controller and the purpose(s) of the processing. The Directive also requires in addition that data subjects have access to additional information "without constraint *and at reasonable intervals*", in particular (i) confirmation as to whether or not data relating to him are being processed, the purpose of the processing, the categories of data concerned, and the recipients or categories of recipients to whom the data are disclosed, (ii) communication to him in an intelligible form of the data *undergoing processing* and of any available information as to the source of the data, and (iii) knowledge of the *logic involved* in any automatic processing of data concerning him at least in the case of the automated decisions (Art.12). This obligation requires for the IDM system that if the data subject exercises the right of access (see also below), the appropriate information about the functioning of the IDM system shall be given. For the use of the various pseudonyms on e.g., a card, an *interface* for the card usage between the application and the card which provides sufficient information to the data subject about which pseudonym is used, for what application and how, will be important.
3. The Directive only allows the *transfer* to a third country if such country ensures an 'adequate level of protection' (Art. 25 1). Derogations from this principle apply, for example with the *consent* of the data subject or if necessary for the conclusion or performance of a contract (Art. 26). For the IDM system, this principle could imply that the data subject is informed (e.g., by the interface) of such transfer outside the EU and could authorize such transfer.

5.2 General scheme

The basic flow for verification is depicted in Figure 3 and proceeds as follows.

1. The subject presents non-biometric data as a claim of his identity.
 - o Depending on the identity reference in the application, the subject will either present his identity reference, or the pseudo identity. The latter is preferred. Only when needed, the SP should be able to link the PT with the IR, e.g., the civil identity. Otherwise the SP should rely on the IdP.
 - o Prior to the first verification, the subject's consent should be given to use his biometric data in this particular application. It could have been given during the enrolment phase, however, if the SP was not specified, the subject should give his consent to the SP.
2. The subject provides a proof of identity based on fresh biometric data, taken from his biometric characteristic and sent to the SP.
 - o As in the enrolment phase, the collected data should be protected throughout the entire verification procedure.
3. The SP verifies if the PT is valid with the biometric characteristics sent by the subject.
 - o The locations from which the PT, PI or IR, are retrieved play an important role, however, they do not influence the different steps and their order in this flow.

⁶ Hereafter referred to as "the directive".

- The subject should know how his data are processed, by who and where, particularly when the PT or a part of it is stored by the SP.
4. The SP notifies the subject of the successful verification and gives him/her access to his/her profile and data. The captured biometric data are deleted.

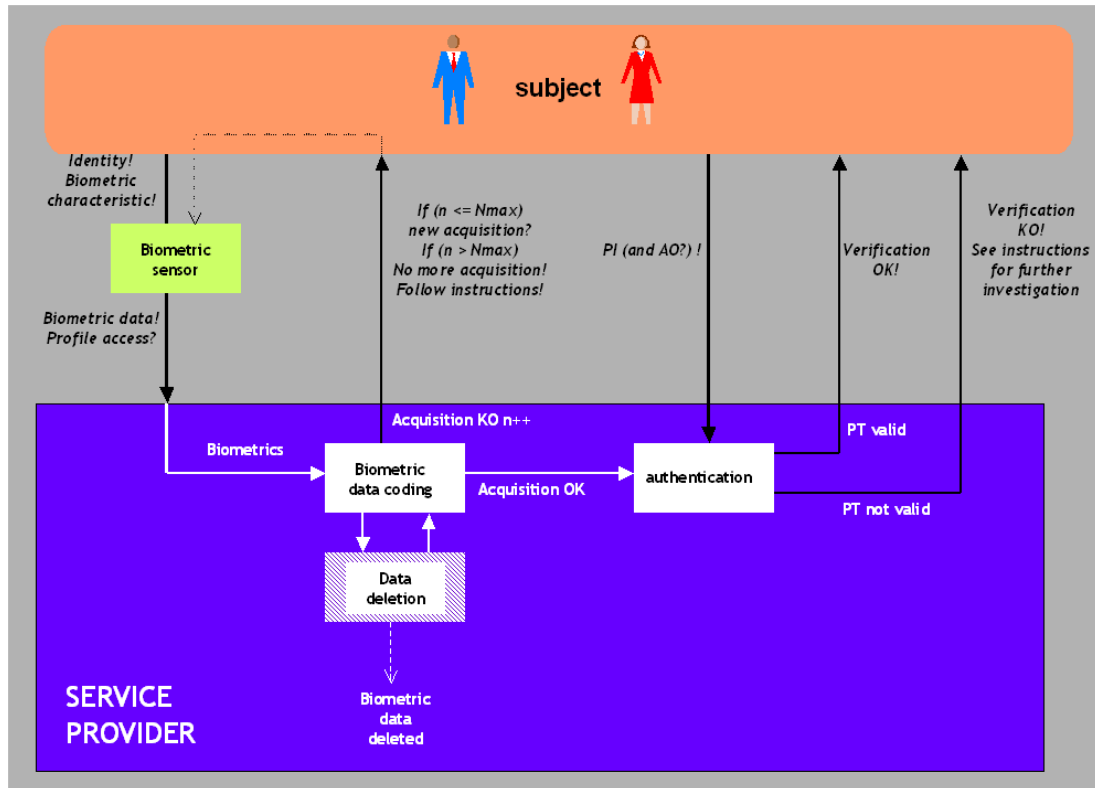


Figure 3: Verification flow.

Comments

- If the IdP and the SP are not the same, the SP should verify the PT validity without any biometric characteristic, using the certificate provided by the IdP at the PT creation.
- For step 2, if the quality of the captured biometric sample is bad, the subject is asked to capture again his biometric characteristic until a certain number of times (cf. registration section). If after the maximum number of acquisitions the quality is still bad, the service informs the subject that the limit is reached and turns the subject toward detailed instructions to follow in case of acquisition failure.
- The subject provides claimed identity and proof in one transaction. If the pseudo IDs do not match or if the claimed identity (if any) is not valid, the service returns that the couple claimed identity/pseudo ID is not valid. Then the service gives instructions for further investigation for ID corruption or technical failure. This may be useful when the service does not want to inform the subject about the identity status (registered or not). This may prevent reverse engineering techniques (“claimed identity ok but biometric not ok”).
- If the subject is not registered, the SP signals a failure of the verification process and requests the subject to try again.
- If the SP needs additional information (a password for instance), it requests extra information.
- If the Verification fails, the SP signals a failure of the verification process and requests the subject to try again until the maximum number of allowed attempts is reached. If the maximum is reached, the subject’s identity is locked out.

- A secure connection, i.e., confidential and integrity-protected, should be established between the subject and the SP to avoid tracking of the subject, by monitoring network activities or observing the biometric data. Subsequent verification attempts (in case of a verification failure) must occur in the same session.

5.3 Identification scheme

A verification scheme can also be extended to an identification scheme (also called a one-to-many verification scheme) when no identity is provided by the subject. In that case, the goal is to recover the subject given his/her fresh biometric data among a set of registered users.

For identification, i.e., verification of a subject's identity among several identities, a central database is often required. Identification is needed, for example, for negative identification functionality or duplicate enrolment check (DEC). The use of the central database is then required to be proportional, and, if applicable, appropriate template protection methods should be used which at least exclude (to the extent that it is possible):

- That templates reveal 'additional information',
- Cross-linking,
- Function creep (use for other purposes than initially intended).

Access control is another classical scenario where identification is useful. A subject is given access to some facilities, without the use of any token or inputs from the subject, after the capture of a biometric sample and its retrieval in the list of authorized users.

AFIS (Automated Fingerprint Identification System) is a good example of an efficient identification system where the use of template protection techniques could bring benefits to a subject's privacy.

6. Revocation

6.1 Privacy requirements

1. The privacy requirements of the enrolment and the verification phases (Sections 4.2 and 5.1) also hold for the revocation procedure.
2. The Directive [Dir9546EC] requires that a data subject *has access* to the data processed about him and *can rectify, erase or block* the processing of data which are incomplete or inaccurate (Art. 12). This requires that an appropriate procedure is defined for the IDM system and put in operation in case the IDM system is compromised. There should in principle be no additional costs involved at the charge of the data subject to exercise its right to rectify, erase or block.
3. The Directive explicitly states that personal data must be kept in a form which permits data subjects to be *identified for no longer than is necessary* for the purposes for which the data was collected and processed (Art. 6.1.e). In the **TURBINE** IDM system, this requirement should be translated in a functionality which contributes to the effective deletion of all personal data (a) once the transaction is completed and no legal obligation to keep such references exist, (b) upon revocation of a biometric identifier and (c) upon ending the use of the application by the data subject.

6.2 General scheme

The basic flow for revocation is depicted in Figure 4 and proceeds as follows.

1. The subject asks for PT revocation to the SP, the SP reassigns the subject to the IdP.
2. The subject presents identifying attributes to the IdP. This identifying attributes are the same as those used for registration.
 - o Again the proportionality principle holds and only the minimum number of attributes that are needed to perform the revocation should be presented.
3. The IdP validates the subject identifying attributes to the SP which validates the subject's request.
4. The IdP revokes the PT. The PI and AD are deleted from where they are stored. The certificate sent by the IdP to the SP to authenticate the PT is stored in a revocation list managed by the IdP (and maybe the SP).
 - o It is the responsibility of the IdP to know where the PT and other data are stored.
 - o The subject should be well informed about the impact of this procedure.
 - o To the extent that it is possible, the SP should remove all application data that are related to the subject and that can be linked with data from other applications through the identity reference. Special care should be taken if the IR is the civil identity or derived from it.
5. The subject requests re-registration or de-registration (if the contract is terminated).
6. Execute re-registration or de-registration scenario.

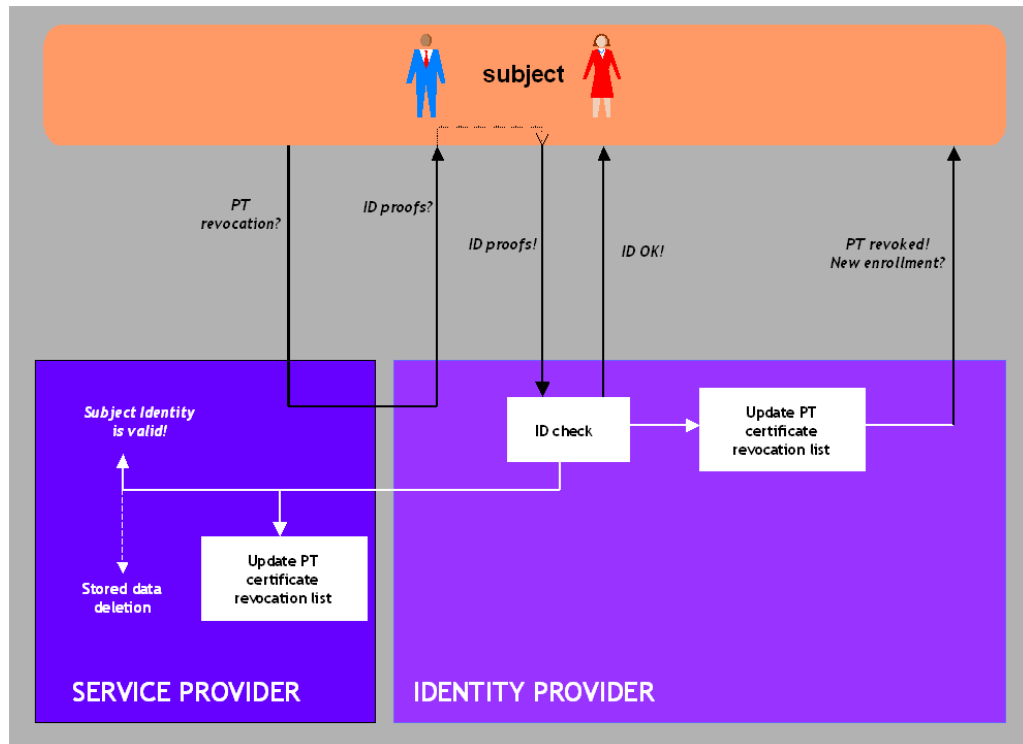


Figure 4: Revocation flow.

Comments

- If possible, the PT revocation should be carried out after a PT verification procedure, either in addition to or in replacement of steps 2 and 3. Then the service reassigns the subject to the IdP for the creation of a new PT. If the PT is revoked because verification no longer works, e.g., due to aging, then this is not possible.
- The revocation may be engaged by the IdP or the SP. In that case the PT certificate is blacklisted and the PI and or AD stored by the SP are deleted.
- If the revocation is related to an Identity corruption, the revocation is asked via the IdP only. The identity verification procedure is heavier than for PT revocation. The revocation procedure for the PT is the same as described in the previous bullet. Then all the SP that refers to the same identity are informed of Identity corruption and all the PT certificates are blacklisted.
- For revocation, the IdP requires additional information.
- If the revocation of the requested identity is not allowed or the identity is not registered, the SP notifies the subject of failure and specific instructions are given.
- If the re-registration is not allowed, the IdP rejects requests for re-registration and specific instructions are given.
- Many of the alternative steps depend on the policy of the target application.
- In the case of ownership verification, the only way to revoke an identity is by means of a watch list, containing either entire PIs or asset references. It may not be feasible to use supplementary data that can be controlled by the identity or service provider, since it may be hard to provide this data upon verification.

7. Multiple identity management

7.1 Management of multiple identities on the same token

In this section, we will study service definition and constraints in the management of multiple identities on the same token. Without loss of generality, we will assume that each service provider will manage only one identity, ignoring and being unable to access additional identities used by other service providers.

7.1.1 Identity and Service providers

Let's consider two identities for one subject, which means that the subject registered to two different and independent service providers. There may be two different situations:

1. The service providers are related to the (same) IdP that issued the two identities.
2. The service providers are related to different and independent Identity Providers⁷.

The Pseudo Identity creation, use and revocation for one SP are absolutely transparent to any other SP. This issue is developed at the token level in section 8.1.

One IdP for all the SPs

If there is one IdP that communicates with all, or at least a set of the service providers, there should be some requirements on the IdP:

- There is no information of any kind that can be transmitted between services (exception may be made with specific conditions in case of Identity corruption, see section 7.1.2),
- An SP cannot know if a registered subject is also registered in any other SP,
- An SP cannot modify information stored by the IdP.

One IdP for each SP

If there is one IdP that communicates for each service providers there should be some requirements on the IdPs:

- There is no information of any kind that can be transmitted between IdPs (exception may be made with specific conditions in case of Identity corruption, see section 7.1.2),
- An IdP cannot modify information stored by another IdP,
- An SP cannot modify information stored by the IdP.
- An SP cannot know if a registered subject is also registered in any other SP.

7.1.2 Identity corruption

The IdP contains a set of data about the subject (civil identity, social security number, etc.) that may be used to validate the subject's identity and to allow his/her registration.

The SPs that refers to the same IdP to validate the subject identity may use:

- The same set of attributes,
- Different sets of attributes with common attributes, or
- Different sets of attributes without any common attributes.

There is one specific situation where there may be, under particular conditions, communication between SPs via the IdP: if an identity is corrupted.

⁷ In practice, a service provider may rely on multiple identity providers, e.g., an international trader working together with several national identity providers.

An identity is corrupted when:

1. Two different subjects use the same identity,
2. One or several identity attributes are used by at least two different subjects, or
3. A subject uses one or several identity attributes he doesn't own.

In the first case, the identity must be revoked and deleted on the token. The data stored by the identity provider are not corrupted. There is no need for the IdP to inform the other SPs.

In the second and third case, the data stored by the identity provider are not corrupted. Those data may be used by other SPs. The IdP should update his data in order to inform indirectly the other SPs that some of the identity attributes are corrupted. Then the other SPs may decide whether or not to revoke their identity.

7.2 Multiple identity management scenario

Prerequisites

In this scenario we use ePrescriptions⁸ as an example of a multiple identity management application. The scenario described below, closely resembles the delivery of ePrescriptions in Sweden (see [Apoteket] for more details). Over 40% of all prescriptions in Sweden are transferred from the doctor to the pharmacy electronically via a health extranet or by using web based prescribing. The prescription is stored in a national mailbox that can be accessed by a pharmacist chosen by the patient at the time he collects his medication. Our scenario presents a TURBINE privacy-enhanced version of this application.

The subject is registered with different health institutions, e.g. hospitals, pharmacies, and medical test laboratories. To prevent the different institutions from using a pseudo identity as a unique identifier and link information between databases, the subject uses different pseudo identities in contact with each of these institutions. All of these pseudo identities may be stored on the same token, e.g. the subject's health card.

In this example the subject has at least two pseudo-identities stored on the token:

1. One of them is used when accessing the personal health database through any kind of terminal (e.g. public terminals available on hospitals and pharmacies or a personal computer equipped with card reader and fingerprint scanner). In this application, where security and privacy requirements are very high, on-line verification and central watch-list functionality are feasible, thus hybrid verification is used.
2. The other one is used when the subject wants to get treatment at the pharmacy. Depending on level of prescription control needed (see section 10.2.1), even off-line verification can be used, which implies an ownership verification model.

Scenario

1. The subject visits a doctor which writes a prescription. The prescription is sent to the subject's prescription inbox in the health database.
2. Back at home, the subject uses the personal computer to gain access to the health database. Through a secure connection, the central health database authenticates itself to the token, which by this authentication knows what "box" to open. In this box the additional data, required to perform hybrid verification, is stored. Fresh biometric data from the subject are used together with the additional data to create a temporary pseudo-identity (PI_1^*). This is sent to the central application which compares it to a previously (at registration) stored pseudo-identity (PI_1). Since they match, the subject gains access.

⁸ See also [PR1ePr08] for aspects of patient privacy and the use of multiple identities for the same subject in drug supply chains.

3. Using the system, the subject finds the prescription in the inbox and grants the pharmacy access to it. This is done by establishing a link between the prescription and the pseudo-identity the subject uses at the pharmacy.
4. At the pharmacy, the subject puts the personal token into the pharmacy's terminal. The terminal authenticates itself and is gained access to its box on the token where additional data needed for the verification process is stored. Using fresh biometric data from the subject together with the additional data, a temporary pseudo-identity (PI_2^*) is created and compared to the original (stored on the token at registration) one (PI_2). The pharmacy's system compares the two pseudo-identities and since they match the system can now use the pseudo-identity to search for linked prescriptions.
5. The prescription linked to the subject's pharmacy-pseudo-identity is found and the subject may get his/her medicine. The pharmacist updates the prescription to note what medicine has been given.

8. Token characteristics

8.1 Token as trusted personal device

Among the different studies, central, local or hybrid storage of subject data, one of the TURBINE identity management solutions provides means for storage of protected templates and other data related to identities and pseudo-identities on a token. The token is carried by the subject as his or her trusted personal device. The trusted personal device could be a smart card, but also a device of another form factor, e.g. a USB dongle equipped with smart card technology. Trusted personal devices of different form factors have been the topic of the FP6 project InspireD (INtegrated Secure Platform for Interactive tRusted pErsonal Devices), a European initiative dedicated to the future of the secure device industry. The vision of InspireD was to base next-generation secure devices on a new common platform approach for Trusted Personal Devices. A comprehensive set of requirements for these tokens has been compiled by the InspireD project (see [Ins-D1], [Ins-D8]). In the following, those requirements relevant for the use in TURBINE identity management systems are cited, complemented by requirements specific for the application of tokens in TURBINE.

When protected templates are stored on a token carried by the subject, the token should be unique, it should not be duplicated or recycled (used for another subject), and therefore it should contain a certificate and a key pair that allows it to prove its authenticity. The key pair should be generated on the token and the private key can never leave the token. The certificate is created at the token activation and could not be modified. It may contain additional information on the token characteristics such as its lifetime and is shared by the identity provider and/or the service provider to allow token revocation and the use of token revocation lists.

This token contains a storage unit and carries biometric data in a protected form and may contain identity information (depending on the trust model policies). The token should follow a specific protocol to secure its communications with the IdP or the SP. An example of such a token is the ePassport as defined by the ICAO standard [ICAO06], which states that an extended access control mechanism must be used to access sensitive data, such as fingerprints or iris templates, stored on the chip of an ePassport.

Depending on the IdP/SP policies, the token may contain:

- A biometric sensor, the biometric data would then be sent in a protected form to the IdP for verification;
- An authentication process, the verification is then carried out within the token. The token then sends the verification decision to the SP and/or the IdP. The decision transfer should be secured to avoid any attack such as: information gathering or change in the decision. To avoid replay attacks, fresh randomness is required from both the token and the reader analogue to challenge response authentications.

The token shall fulfil the following security requirements:

- The token shall be designed as a tamper resistant device protected against side channel attacks and physical attacks.
- The token should be a trust enabler to access to services proposed by third parties (on-demand application/library loading, co-operation with static security equipment, etc.)
- It should ensure third parties (i.e. neither the issuer of a token nor the user) that the token is trustworthy, while protecting user privacy by providing: strong user authentication and identification services (e.g. PIN, biometrics...), pseudonymity or anonymity of the user, trusted attestation.
- The token should offer an execution environment suitable for applications that employ secrets, without leaking information.
- It should allow secure remote or local configuration or enhancement (including applications, OS services, and data) and application download control in a way similar with the GlobalPlatform model [GP] or its evolutions. Access to resources shall be only granted to authorised processes clearly identified and controlled by the OS.

In the case of multiple identity management, the subject may use multiple identities on the same token. This means that one token contains and manages several Pseudo Identities (PIs) for several SPs. Let's consider 2 service providers, SP1 and SP2. Protected Template PT1 is created for SP1 and Protected Template PT2 for SP2. Both PTs are stored in one token, entirely (PI and auxiliary data AD) or partially (only PI).

It is important to make sure that each SP cannot download all the PTs contained in the token. This would allow it to track down for which service they are created for and maybe access the data. In order to avoid this, the token should be divided into "boxes" that do not communicate. The service has only a "key" to access its own "box" and this "key" must not reveal anything about its owner. That is, it should not be possible to link a token to a specific service provider through the key or any identifier of the box. A service provider may use the same key for all tokens, but where security constraints so require, the key should be unique per token.

The SP can write (for PT creation/revocation) and read (for PT verification) only in its own box. SP1 can only read or modify PT1; it has no access to PT2 or the rest of the token.

The token is lost

If the token is lost, the user must inform all the services that stored pseudo identities on the token about the loss and follow their revocation procedure.

An identity must be revoked

When an identity is revoked, the SP optionally revokes the identity directly on the token at the next connection following its own revocation procedure.

8.2 GlobalPlatform Security Domains for secure multiple pseudo identities using different applications

An important requirement is the separation of the applications representing the different service providers. This can be realized by concepts described in the GlobalPlatform Card Specification [GPCS06]. The GlobalPlatform Card Specification is a secure, dynamic card and application management specification that defines card components, command sets, transaction sequences and interfaces that are hardware-neutral, operating system neutral, vendor-neutral and application independent. The GlobalPlatform Card Security Requirements Specification provides guidance for selecting card configurations most appropriate to the security policies set up by the Card Issuer and Application Providers. Card vendors are also provided with guidance to implement security functions in a consistent manner.

Even for smart cards that are not compliant with the Global Platform specification - as they may be used for Banking, Health, and ID applications - the concepts should nonetheless be considered for implementation of securely separated application domains.

In the context of TURBINE, mainly the concept of security domains is relevant. The GlobalPlatform Card specification proposes this concept for cards with more than one application, e.g. an identity card with an additional payment application or electronic purse. The security-relevant data (keys, PINs etc.) of each of the different applications can be located in a secure area of the card called a Security Domain. Security Domains act as the on-card representatives of off-card authorities and enforce the security policies defined by the owner.

The card has at least one Security Domain, the Issuer Security Domain, which is the on-card representative of the card issuer, see Figure 5. To access the card securely, the card issuer of a card uses a secure connection called a Secure Channel. Once a connection is established, the Secure Channel provides an end-to-end secure communication path between an on-card security domain and an off-card entity. This concept is illustrated in Figure 5.

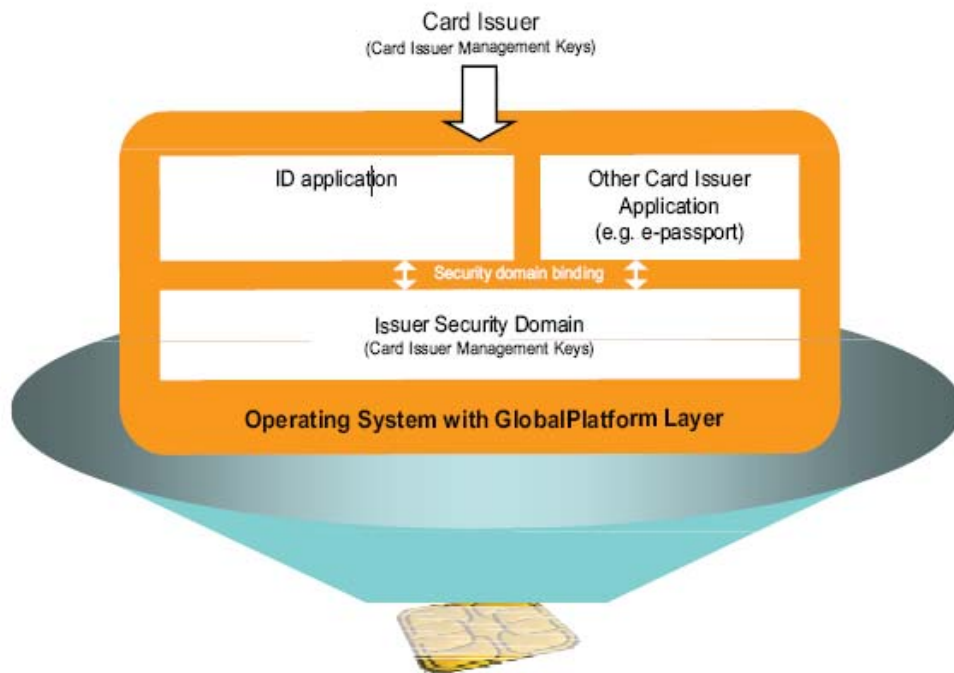


Figure 5: Global Platform Card Issuer Security Domain [GPIDM07].

The issuer of a card can decide to add an application from a business partner called an application provider. In the TURBINE terminology an application provider is a service provider. In this case the separation of responsibility (who owns and controls the application or who owns and controls the platform, for example) is necessary. The role of card issuer and application provider are quite distinct, and if the card platform provides facilities to keep these responsibilities separate, then application providers may put their applications on the card platform without inflicting vulnerabilities to the card issuer or other application issuers. This is illustrated in Figure 6. GlobalPlatform offers the issuer a means to give application providers their own Security Domains on the chip, while assuring that applications, application data and application management remain totally separated.

The GlobalPlatform Specifications are well suited for this type of multiple actor deployment. A card platform that abides by the GlobalPlatform Card Specifications allows an issuer to authorize an application provider to exclusively and independently occupy a Security Domain within the card, while enabling the issuer to retain control over the card and its own applications in a secure and standardized manner. With a Secure Channel to their own Security Domain, application providers can now load, personalize and update their applications, or even operate the application services they control on the card, while meeting the necessary standards and regulation for privacy and security. GlobalPlatform's Card Specifications support symmetric and asymmetric keys and either can be used by the issuer to access a secure area and open a Secure Channel for card management. A Secure Channel can use a synchronous or asynchronous mechanism to allow the card and application providers to authenticate and transfer secure information to the Security Domain. There is often a requirement to load new applications onto existing cards which have already been deployed, as this saves reissuing cards. GlobalPlatform supports post-issuance download in a very natural way; the application loading and activation mechanisms defined by GlobalPlatform do not make any distinction between pre- and post-issuance loading as the mechanisms and Secure Channels are identical.

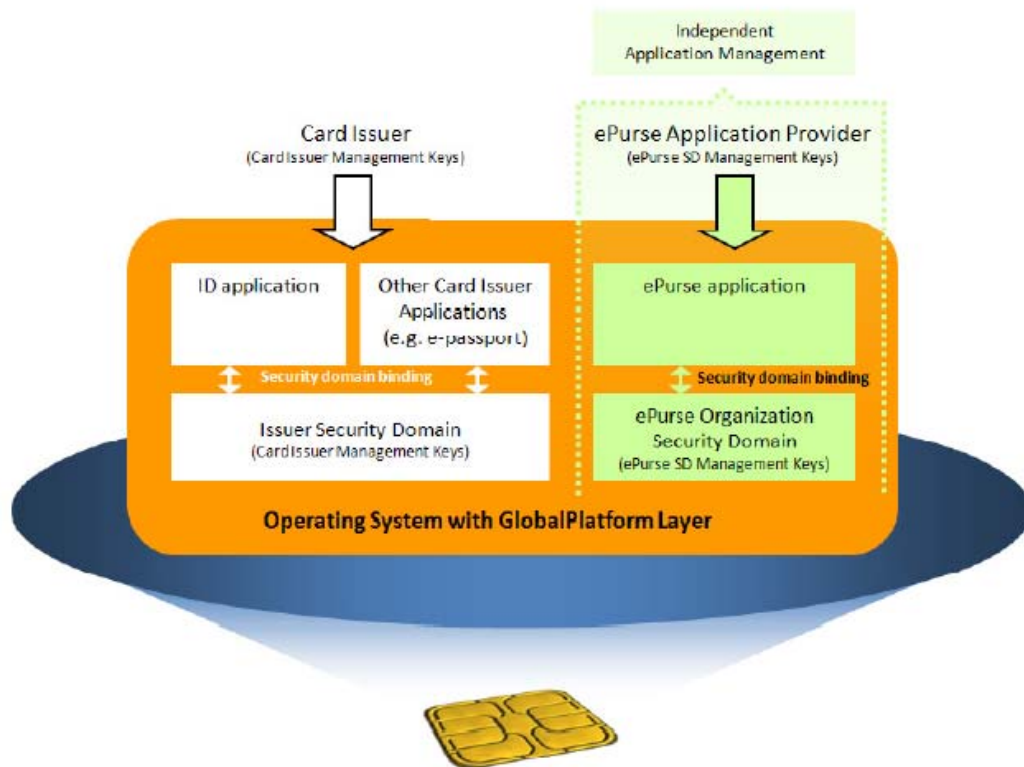


Figure 6: Service Provider Security Domains [GPIDM07].

8.3 Draft GlobalPlatform based architecture for multiple TURBINE service providers

In Figure 7, an exemplary GlobalPlatform based architecture for TURBINE applications is drafted. Here, the token stores four Service Provider applications (for two pharmacies and for two shops) and three Identity Providers' Supplementary Security Domains (for one health insurance and for two different banks). The two pharmacies use different service applications as they might have different bonus systems but have the same Identity Provider, here a health insurance. The two shops, on the other hand, work together with two different banks which have two different on-card Identity Provider SSDs.

The ISD is the security domain for the issuer (Issuer Security Domain), e.g. the company in case of an enterprise id card or the network operator in case of a Universal Integrated Circuit Card (UICC). Applications provided by the issuer can use this security domain. For additional service providers, a supplementary secured domain (Trusted SSD) is provided, which is owned by a Trusted Service Center, i.e. a Provider trusted by all Identity Providers, e.g. a governmental organisation. There could also be several Trusted SSDs / Trusted Service Centers in the case that not all IdPs trust the same Trusted Service Center (not shown in the figure).

The idea of this construct is that the issuer has no access to the Trusted SSD and hence also no access to the SSDs of the third party services. This is important for instance in the mobile telecommunication domain, where the issuer, i.e. the network operator, should not have access to data from 3rd party providers, e.g. bank account data. In some cases, for instance for an enterprise id card, the ISD and the Trusted SSD may not be differentiated.

The Trusted SSD can be used for the (remote) installation of IdP SSDs. With help of the Trusted Service Center / Trusted SSD, Identity Providers may initiate the creation of IdP SSDs as their on-card representatives. After creation, the IdP SSDs communication key sets can be exchanged with new confidential key sets by the respective off-card IdP administration servers so that even the Trusted SSD no longer has any knowledge of the used keys.

Usually, an IdP SSD should store one key set per associated SP service application so that all SP Service Applications can establish independent secure channels. SP service applications can only communicate with the respective off-card SP service application via the associated IdP SSD. Using Secure Channel Protocols with different key sets per SP service application, both the integrity and the confidentiality of the communication is protected. Such, neither external parties (e.g. by eavesdropping) nor other on-card SP service applications or IdP SSDs can track or manipulate any communication between off-card and on-card applications. Both IdP SSDs and SP Service Applications of the same and of different IdPs are all completely isolated from each other.

De facto, the key sets, managed on-card by the IdP SSDs and off-card by the IdP administration servers, uniquely separate between the different pseudo identities of all subjects. The GlobalPlatform architecture ensures the complete separation between the different SSDs and between the different SP service applications where the Trusted SSD is responsible for the secure installation and initial personalisation of the Identity Provider SSDs.

The real identities of the subjects using SP service applications on the Token are only known off-card via the associated IdP administration servers. While the communication security relies on these Secure Channel Protocols with the distinct key sets, the authentication to each SP service application could rely on derived biometrics where the respective Protection Template per SP Service Application could also be stored in the associated IdP SSD.

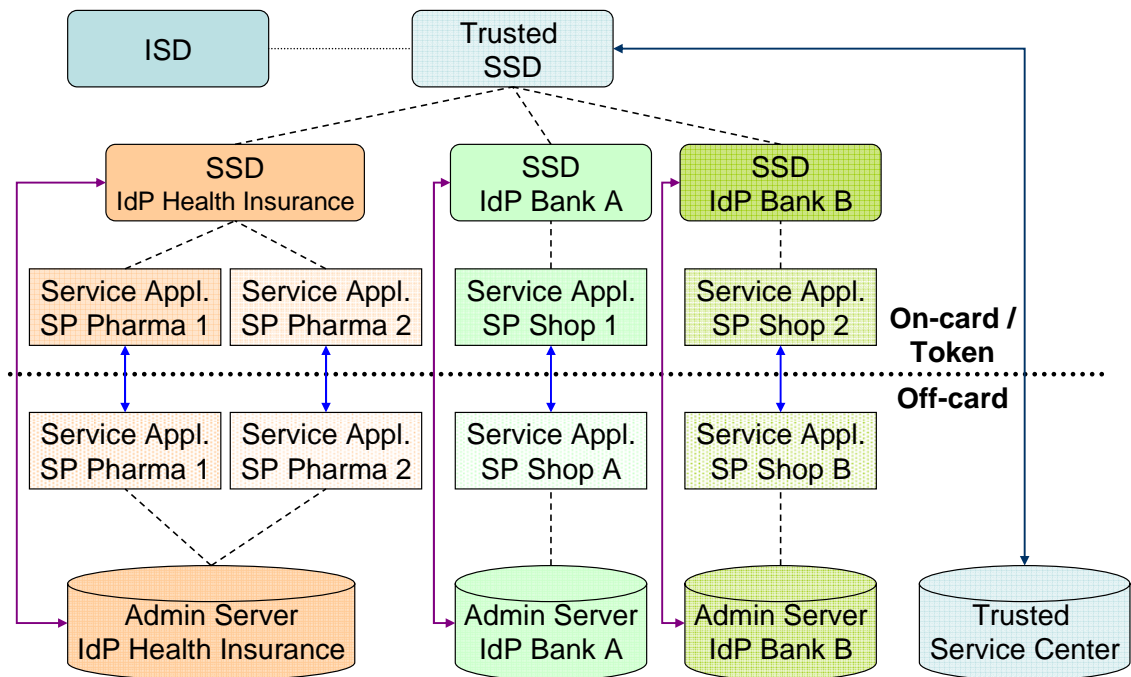


Figure 7: Architecture proposal.

9. Duplicate enrolment check scenarios

In this section, we analyse scenarios, where multiple identities should be detected such as the Duplicate Enrolment Checks (DEC) in visa or passport application processes. We also analyse the potential impact that the **TURBINE** project can provide to solve the forged passport problems.

The main issue is when an individual registers twice, using the same biometric data, for the same service, such as visa or passport application services, when it is allowed to register only once. The subject may also register two times with different amounts of biometrics. For instance, the registration requires 4 fingers, and the subject claims to have lost one finger between the 2 registrations.

There are 2 critical cases:

1. The SP policy is to give control to the PT (or PI) to the subject only. Then the SP has no option to check if the subject is already registered to the SP with another Identity.
2. The Subject is already registered with an SP and asks for registration to another SP, which is not compatible with the previous SP, with another identity. For instance, the same subject asks for a French passport with identity A and for an Italian passport with identity B, while this is not allowed.

For those two cases, the IdP should keep a database that would allow the duplication check. This database is not part of the Trust Model Architecture. It is independent of the Identity Management System and is used only by the IdP at registration step. This implies that, when the Identity Management System is specified, the responsibilities, rights and role of the IdP have to be defined. This information (responsibilities, rights and role) is subject to legal requirements, including those defined in [D1.1.1] and those specified by the Article 29 Working Party on Data Protection [A29WP]. They should also be explained to the subject when he registers to the SP.

In the following cases the duplication check can be carried out by PT comparison:

- The SP shares control of the PT with the subject,
- The subject registers to compatible SPs. The subject can register to SP₁ and SP₂ with a different identity.

For detectability of duplications, it is necessary to maintain a kind of central list of registrations with biometrics and/or identifiers data. It can be a database containing a biometric sample for all the registered users along with an associated identifier. As soon as a new subject claims for registration, the IdP can check if the corresponding biometric or identifier is already registered and take the appropriate actions in case of any conflict (for instance when a matching biometric sample is already enrolled with a different identifier). Such a procedure raises privacy issues. A way to reduce the risk is to conduct it in a secured off-line environment but some applications do not support this constraint. We describe below a solution which intends to be an alternative to a sole central database with couples of biometrics data and identities.

9.1 A solution for DEC using a “weak link scenario”

For visa and passport applications, the civil identity is required and must be accurate (used by the civil ID owner). Then the relation civil ID/user is important. The government must be sure that the individual who claims the civil ID is the actual owner of the civil ID. For this kind of application, the system may strongly benefit from the use of biometrics because it allows negative identification and revocation. Nevertheless, it may be required, for privacy protection, to separate biometric data from civil data.

In this section, a service scenario is proposed for an application that allows DEC. We consider two off-line databases:

- A database of registered civil identities
- A database of registered biometrics (captured samples) in a protected template form

The access to those databases is only authorized to the IdP. The SP does not have direct access to them.

To make sure that an individual cannot corrupt a civil identity, we suggest to store some specific data together with this identity. During the enrolment phase, a derivative of the captured biometric sample is constructed. This derivative, called there below a biometric key, is close to a message digest of the biometric sample. One of its goals is to be less discriminating than the original biometric sample. For instance, it can be computed as the least significant bits of a biometric sample. This biometric key has the following properties:

- It is created from the captured sample or the protected template of the civil identity owner.
- The key is robust to different captured samples from the same biometrics (i.e. almost reproducible from an acquisition to another one).
- The key has a weak power of discrimination among biometric keys. This means that a key stored in the database of registered identities would not allow to identify (i.e. to sort out) a single captured sample or protected template among those stored in the database of registered biometric data.
- The key is discriminating enough to avoid key corruption (guessing, retrieval without a genuine sample) except with a low probability.

We can then define the following scenarios:

Enrolment

1. The user asks for registration for a SP and presents his/her biometrics.
2. The IdP creates a protected template specific to the SP and checks the PT database if the PT (biometrics) is already used.
 - a. If the biometrics are already used, the IdP informs the user that he/she may be already registered and proposes a revocation procedure. Moreover, in parallel, the IdP computes the biometric key corresponding to the fresh biometrics and recovers the corresponding registered ID. The IdP then asks for the user civil ID and if these IDs doesn't match, the IdP asks for a new acquisition and logs after, a few unsuccessful tries, the fact that the ID may have been victim of a fraud attack.
 - b. If the biometrics are not already used, the IdP continues with step 3.
3. The IdP asks for the civil ID.
4. The user enters his/her civil ID.
5. The IdP verifies the ID and checks if the civil ID is already used.
 - a. If the civil ID is already used the IdP informs the user after several tries that the civil ID may be corrupted and keeps a trace of a possible fraud attack on this ID.
 - b. If the civil ID is not registered, the IdP continues with step 6.
6. The IdP creates a biometric key and stores the couple civil ID/biometric key in the registered civil ID database and the protected template in the corresponding and independent database. The PT may also be saved on a token, depending on the SP policy.

Verification

1. The user presents his/her civil ID and biometrics (or token and biometrics)
2. The SP checks the new PT (computed from the fresh biometrics) with the PT stored in the PT database (via the IdP) or in the token.
3. The SP validates or rejects the users' rights.

Revocation (asked by the user)

1. The user asks for revocation to the SP and presents his/her biometrics.
2. The IdP creates a protected template specific to the SP and checks the PT database if the PT (biometrics) is registered.
 - a. If the PT is not registered, the IdP informs the user that he/she is not registered.

- b. If the PT is registered, the IdP continues with step 3.
3. The IdP asks for the civil ID.
4. The user enters his/her civil ID.
5. The IdP checks if the civil ID is registered.
 - a. If the civil ID is not registered, the IdP informs the user after several tries that the civil ID is not correct.
 - b. If the civil ID is registered, the IdP continues with step 6.
6. The IdP checks the couple civil ID/biometric key with the fresh biometrics.
 - a. If the captured sample doesn't match the biometric key corresponding to the civil ID, the SP asks for a new acquisition and logs after, a few unsuccessful tries, the fact that the civil ID may have been victim of a fraud attack.
 - b. If the biometrics matches the biometric key corresponding to the civil ID, the IdP continues with step 7.
7. The identity is revoked.

10. Biometry trust management

10.1 Performance level Structure

As fingerprint biometry state-of-the-art, stronger accuracy is obtained using several fingers of an individual instead of one finger. In this section we will not consider only fingerprint technology but multi-modal biometry as well, allowing the same improvement of accuracy.

In this section, we distinguish different notions of security and performances. This will be developed in section 10.1.1.

Section 10.1.2 defines the criteria for determining the level of authentication performance required for specific applications. The criteria are generic criteria that have to be tuned depending on the service (see section 10.2). The amount of biometric data (e.g. number of fingers, biometric fusion) will be determined by those criteria.

The second section presents services that can benefit from the use of such performance levels and some specific criteria.

10.1.1 Criteria on performance

The performance level we will define in each application example will vary with the authentication system performance. The False Acceptance Rate (FAR) is the probability that a false identity claim will be accepted, thus allowing fraud. The False Rejection Rate (FRR) is the probability that a true user identity claim will be falsely rejected, thus causing inconvenience. The system performance is a trade-off between the FAR and the FRR. Additional error rates may also be estimated, e.g. Failure To Enrol, Failure To Acquire, Failure To Verify, as they have direct consequences on the efficiency of the system. If a system has various performance settings at which it can be operated, then the FAR and FRR will vary in accordance with the performance setting selected. In general and here in particular, as the FAR is reduced, the system becomes more secure against fraud, and, the FRR is increased. A consequence of a higher FRR is user inconvenience (discomfort), since successful authentication of an authorized user may require additional access attempts. For a given system, there may be only a handful of performance settings available, to increase the flexibility and the possibilities of such settings, we will use multiple fingers (see section 10.2.1) or multimodal biometrics (see section 10.2.2).

As overviewed there above, the **accuracy** or **performance** of a biometric system is associated to several error rates, of which the main ones are the FAR and the FRR. The FAR is measuring the difficulty to be falsely accepted by the system – it represents the **biometric security** of the system. In a given context, the performance is routinely estimated by the FRR only, as a biometric system will tend to have a fixed biometric security level (FAR). As we intend to handle protected biometrics templates, the biometric security level has to be differentiated from the security level of the protection mechanism which should rely mostly on cryptographic techniques. In the sequel, we are discussing only the performance and biometric security levels.

Consequently, and in general, the more there are biometric data, the better the performance and the biometric security level of a system would be. For instance using four fingers instead of one finger would allow to achieve a better performance/biometric security (FRR/FAR) trade-off.

Different performance level may be required within one application. According to the performance level, there may be two kinds of specifications:

1. The data to be acquired (fingers for instance) and the FAR are fixed, then if a subject is not able to provide all the required biometric data, the FRR may increase. The system may be more sensitive to acquisition and Captured ID may be falsely rejected more often. Then the biometric security level of the system will remain the same but the user's comfort may be deteriorated.
2. The required FAR and FRR are fixed and the amount of acquired data will depend on the subject (for instance if one of the required finger is amputee).

It is important here to note that a system will never really be able to provide a fixed FAR/FRR couple because it will depend on the acquisition quality or the data quality. The amount or biometric data cannot be completely controlled.

Moreover, if we consider, for instance, multiple finger structure, all the fingers aren't processed the same way. A thumb considered by the system as an index will not provide the same performance as an index. It is also important to note that the use of another finger (a thumb instead of a required index) to reach a certain amount of data is only allowed in the case of an amputee finger (and not in the case of a damaged or banded finger) because of the risk of multiple identities (the same subject may come one day with an index banded and acquire an index and a thumb, and come back another day with 2 indexes).

10.1.2 Technical criteria

The general scenario flows (registration, verification and revocation) defined in the previous sections can be used as reference to define the corresponding scenarios in each application example. Some issues should be studied:

- For the PT creation there may be two options: The IdP creates one Protected Template that gathered all the performance level or one Protected Template per performance level. This may depend on the template size required by the authentication system.
- In the verification scenario, the SP verifies with required biometric data the validity of the unique stored protected template, all the stored protected template (and check that at least one is valid considering the biometric data provided) or the stored protected template that corresponds to the required performance level.

In the case of multiple protected templates, the requirements are those defined in section 7 of this document.

In any case, the chosen process should not have an impact on the security of the template protection (i.e. the concrete level of this protection), which should remain constant (or at least above a fixed security bound).

10.2 Services description

In the next two sections, we specify two services that can be obtained with a Performance Level Structure. The first one is a service that takes place in a pharmacy for instance, when a subject gets his/her medicines. The second one is a service that corresponds to an access control service that may take place in an airport.

10.2.1 Prescription control

A subject registered to an eHealth service that proposes an electronic prescription service. To get his/her medicines the subject needs to authenticate at the pharmacy. The authentication process follows a Performance Level Structure. The required confidence in the subject identity depends on the medicine he/she is getting.

Health control bodies draw up lists of medicines according to the risks their take represents. For instance, in France, there are several lists of medicines:

- list I : dangerous medicines,

- list II : potentially dangerous medicines, less toxic than those in list I,
- narcotic: psychotropic medicine that is able to induce dependency and noxious effects on the mental and physical health,
- Limited prescription medicines : those reserved for hospital use, those that can only be prescribed by a hospital doctor and those that need specific watch and prescribed by a specialist,
- exceptional medicines : especially expensive medicines, they need a specific monitoring and medical justification for refunding,
- Limited and exceptional medicines.

If we consider the medicines that do not require a prescription as well, we can identify 7 performance levels. This may be too much. So we can limit the number of levels to 5:

- Level 0: no prescription required (no refunding),
- Level 1: exceptional medicines (medicines that are not in list I or II, that are not narcotic),
- Level 2: medicines from list II,
- Level 3: medicines from list I,
- Level 4: narcotic.

Only Levels 1 to 4 require the use of biometric. Table 1 gives an example of how several fingers may be used as requirements for the performance levels the pharmacy needs about the client's identity.

Since we propose an authentication system, we focus on the FAR of the system. We consider a constant FRR and decrease the FAR using more fingers.

The number of fingers allows the authentication system to have defined characteristics such as:

- The system performance (FAR/FRR),
- The capture time.

The system will adapt its behaviour according to the FAR required and the number of fingers available. If the subject has amputee fingers and cannot provide the 6 fingers required for performance level 4, the system may still be able to achieve a FAR very very low but the FRR will be higher than required.

<u>Level</u>	<u>Medicines category</u>	<u>Capture time</u>	<u>FAR⁹</u>	<u>Nb. of fingers</u>
0	No prescription required			0
1	Exceptional medicines (medicines not in list I or II, and not narcotic)	Immediate	Standard	1
2	Medicines from list II	Short	low	2
3	Medicines from list I	Long	Very low	4
4	Narcotic	Very long	Very very low	6

Table 2: Performance Levels for an eHealth service.

It is possible to introduce in this Performance Level Structure an additional level of biometric security. Instead of using fingers from one individual, it would be interesting to use fingers from two individuals (the client's and the pharmacist's in this application). Table 2 would then be Table 3:

⁹ The actual FAR values depend on the system and the performance requirements of the service.

<u>Level</u>	<u>Medicines category</u>	<u>Capture time</u>	<u>FAR¹⁰</u>	<u>Nb. of fingers</u>
0	No prescription required			0
1	Exceptional medicines (medicines not in list I or II, and not narcotic)	Immediate	Standard	1 from client
2	Medicines from list II	Short	low	1 from client 1 from pharmacist
3	Medicines from list I	Long	Very low	4 from client 2 from pharmacist
4	Narcotic	Very long	Very very low	6 from client 4 from pharmacist

Table 3: Security Levels for an eHealth service.

10.2.2 Access control

An airport has several restricted areas. The airport employees don't have the same access rights to those areas. The airport management wants to control their access¹¹ using an authentication system based on multimodal biometrics. The restricted areas require specific levels of trust in the employee identity and access rights.

There are numerous factors to take into account when we decide to merge several biometric technologies:

- Sensor price
- Acquisition environment (compromise Failure To Acquire (FTA)/ Failure To Enrol (FTE))
- Technology merging algorithm
- Interoperability
- Time of capture
- Ergonomics (user friendliness)

As for the previous application example that uses multiple fingers to define performance levels, we will consider multimodal biometrics.

Since we propose an authentication system, we focus on the FAR of the system. We consider a constant FRR and decrease the FAR using more biometric data from multimodal biometrics.

The biometric technologies used for each performance level will allow the authentication system to have defined characteristics such as:

- The system performance (compromise FAR/FRR),
- The capture time.

The number of performance levels depends on the number of restricted areas. We propose here an example of an access control for the employees of an airport to 4 different restricted areas.

We made the following hypotheses:

- there are no constraint on the price of the system,
- the acquisition environment is not controlled (lighting may change during the day for instance),

¹⁰ The actual FAR values depend on the system and the performance requirements of the service.

¹¹ On a related note, PRIME has studied logical access control private information stored and processed by airline companies. See [PRLaca08].

- the biometric algorithm are controlled (the system is built for specific biometric algorithm which is not a major constraint since the system is develop for a unique client for instance an airport).

At capture stage, several biometrics may be captured at the same time (face and fingers for instance) to decrease the capture time.

We can then define the following performance level:

<u>Level</u>	<u>Area category</u>	<u>capture time¹² (in second)</u>	<u>FAR¹³</u>	<u>Biometric technologies</u>
0	No restriction			None
1	Employees with access level I	1	Standard	One finger
2	Employees with access level II	2	low	2 fingers
3	Employees with access level III	5	Very low	2 fingers + face
4	Employees with access level IV	8	Very very low	2 iris + face

Table 4: Performance Levels for an Access Control service.

In this example, the use of multimodal biometrics may introduce a biometric security level notion, in particular against fraud.

¹² The capture time here is a rough estimate of the time required by the sensor to acquire the biometrics. This time does not take into account the ergonomics of the acquisition software. We consider that the employees are familiar with the acquisition process.

¹³ The actual FAR values depend on the system and the performance requirements of the service.

11. Conclusion

In this document the different identity management procedures (creation, verification and revocation of identities) and the relevant privacy requirements have been presented as sequences of transactions between data subjects, service providers and identity providers. They have been generalized to the specific cases of duplication enrolment checks and the management of multiple identities on the same token. The properties of such a token have been studied and a first draft architecture was proposed to deal with multiple service providers accessing the same token. A performance level structure was defined and applied on two different services: prescription control and access control. The draft architecture and the two service descriptions will serve as direct input to work package 1.3 where the functionality and the architecture of the project's demonstrators will be elaborated.

12. Bibliography

- [A29WP] ARTICLE 29 Data Protection Working Party. Working document on biometrics Working Document on Biometrics. <http://ec.europa.eu/justice-home/fsj/privacy/docs/wpdocs/2003/wp80-en.pdf> , 2003.
- [Apoteket] EC, Information Society and Media, ICT for Health Unit. Good Practice Knowledge Database. Good eHealth Case: Apoteket, Sweden – ePrescribing, <http://www.good-ehealth.org>, http://www.e-receptstockholm.se/imcms/servlet/GetDoc?meta_id=1008, http://195.227.12.109/kb_empirica/browseContent_alt.do?contentId=23&action=v3
- [Biosig08] J. Breebaart, C. Busch, J. Grave, E. Kindt. *Reference Architecture for Biometric Template Protection based on Pseudo Identities*. In Proceedings of the Special Interest Group on Biometrics and Electronic Signatures, Lecture Notes in Informatics (LNI) P-137, A. Brömme, C. Busch, and D. Hühnlein (eds.), Bonner Köllen Verlag, pp. 79-92, 2008.
- [D1.1.1] Justine Grave (ed.), *TURBINE deliverable D1.1.1. Requirements for privacy protection and trusted identity verification*, July 2008.
- [Dir9546EC] *Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data*. Official Journal of the European Communities of 23 November 1995 No L. 281 p. 31.
- [FID-D2.3] Thierry Nabeth (ed.), *FIDIS deliverable D2.3. Models (v2.0)*, <http://www.fidis.net/fileadmin/fidis/deliverables/fidis-wp2-del2.3.models.pdf>, 6 October 2005.
- [GP] GlobalPlatform organisation, <http://www.globalplatform.org>
- [GPCS06] GlobalPlatform, Card Specifications version 2.2, <http://www.globalplatform.org/specificationcard.asp> , March 2006.
- [GPIDM07] GlobalPlatform, The GlobalPlatform Value Proposition for Identity Management, http://www.globalplatform.org/uploads/GP_White-Paper_IdentityMGMT_justified.pdf, November 2007.
- [HAC1997] Alfred J. Menezes, Paul C. van Oorschot, and Scott A. Vanstone. *Handbook of Applied Cryptography*. Discrete Mathematics and Its Applications. CRC Press LLC, 2000 N.W. Corporate Blvd., Boca Raton, Florida 33431, U.S., 1997.
- [IDEMIX] IDEMIX, anonymous credential system, <http://www.zurich.ibm.com/security/idemix/>
- [ICAO06] International Civil Aviation Organization (ICAO). Doc 9303 Machine Readable Travel Documents – Part 1 Machine Readable Passports – Volume 2 Specifications for Electronically Enabled Passports with Biometric Identification Capability, 2006.
- [Ins-D1] *Inspired Deliverable 1. TPD Functional requirements (Abstract)*, <http://web.archive.org/web/20071021020009/http://www.inspiredproject.com/>
- [Ins-D8] *Inspired Deliverable 8. Security requirements for TPD (Abstract)*, <http://web.archive.org/web/20071021020009/http://www.inspiredproject.com/>
- [PRI-D14.1c] Simone Fischer-Hübner, Hans Hedbom (eds.), *PRIME deliverable D14.1.c. Framework (v3)*, https://www.prime-project.eu/prime_products/reports/fmwk/ , 17 March 2008
- [PRIaca08] Yves Deswarte, *ACAPIS - Prototype Access Control for Airline-related Private Information Storage*, PRIME closing event, 21 July 2008, Leuven – Belgium, <https://www.prime-project.eu/events/closingevent/ACAPIS%20-%20Yves%20Deswarte.pdf>
- [PRIePr08] Riccardo Serafin, *eHealth: Online Prescription*, PRIME closing event, 21 July 2008,

- Leuven – Belgium, <https://www.prime-project.eu/events/closingevent/eHealth%20-%20Riccardo%20Serafin.pdf>
- [PRlide08] Jan Camenish, *Identity Mixer Demonstration*, PRIME closing event, 21 July 2008, Leuven – Belgium, <https://www.prime-project.eu/events/closingevent/IDEMIX%20-%20Jan%20Camenish.pdf>
- [PRIMELife] PRIMELife, Deliverable 3.1.2, Project Presentation, http://www.primelife.eu/images/stories/deliverables/d3.1.2-project_presentation-june_2008-public.pdf, June 2008
- [RFC2459] R. Housley, W. Ford, W. Polk, and D. Solo. *Internet X.509 Public Key Infrastructure Certificate and CRL Profile*. RFC 2459, <http://www.ietf.org/rfc/rfc2459.txt>, January 1999.
- [RFC2560] M. Myers, R. Ankney, A. Malpani, S. Galperin, and C. Adams. *X.509 Internet Public Key Infrastructure Online Certificate Status Protocol - OCSP*. RFC 2560, <http://www.ietf.org/rfc/rfc2560.txt>, June 1999.
- [RFC3281] S. Farrell and R Housley. *An Internet Attribute Certificate Profile for Authorization*. <http://www.ietf.org/rfc/rfc3281.txt>, RFC 3281, April 2002.
- [SC27BTP] ISO/IEC JTC 1/SC 27 3rd. Working Draft 24745 – Information technology - Security techniques – Biometric template protection, June 2006
- [SC37SD2] ISO/IEC JTC1/SC 37 Standing Document 2 (SD 2) version 10, *Harmonized Biometric Vocabulary*, <http://isotc.iso.org/livelink/livelink?func=ll&objId=2299739>, July 2008
- [STORK] STORK (Secure idenTity acROss boRders linKed) <http://www.eid-stork.eu/>

Appendix A – BIOSIG 2008 Paper¹⁴

Reference Architecture for Biometric Template Protection based on Pseudo Identities

Jeroen Breebaart, Christoph Busch, Justine Grave, Els Kindt
jeroen.breebaart@philips.com; christoph.busch@hig.no;
justine.grave@sagem.com; els.kindt@law.kuleuven.be

Abstract

Biometric authentication is often considered to enhance identity verification. The use of biometrics also introduces new challenges to protect the privacy of the subjects while at the same time increasing the security of a verification system. In this paper, a set of requirements is proposed for biometric processing techniques to safeguard privacy and security. From these requirements, a reference architecture is derived that outlines processes and interfaces of biometric template protection methods in a high-level, technology-neutral way.

1. Introduction

The increasing demand for enhanced border control security, and the increasing amount of electronic transactions that are being sent across wired and wireless networks has created a strong need for more reliable identity management. In an identity management system, verifying the identity of a person is a critical task. Existing possession-based identification methods (an ID card, a token or a key) or knowledge-based methods (a PIN, or a password) can be forgotten, lost, shared or stolen, possibly resulting in identity theft or abuse.

In a financial application for example, identity theft may lead to account fraud, payment card spoofing, forgery of cheques or the use of stolen credit card numbers. In the healthcare domain, identity theft can result in access to medical records, unauthorized access to restricted areas, unauthorized use of medication or medical treatment, or health-insurance fraud. In government applications, identity theft may result in counterfeit or abused identity documents. This can have serious consequences since governmental identity documents are often used to authenticate identities for other applications as well.

According to [tw], every 3 seconds there will be a new victim of identity theft in the US and the total damage of identity theft is estimated at a USD 53 billion a year. The need for more reliable identity verification has resulted in an increased interest in biometrics to allow extension of the traditional possession and knowledge-based authentication methods. Biometrics can deliver an increased reliability for identity claim verification, while at the same time being more convenient since biometric characteristics are not easily forgotten or lost.

2. Challenges

1. **Privacy.** The use of biometrics as identity verification mechanism has also raised concerns. More specifically, the tight coupling of a biometric verification method and physical/anatomical properties allows the use of biometric measurement data for other purposes than intended, hence resulting in a privacy risk. This risk may be subdivided into four categories:

¹⁴ This section contains the text of [Biosig08]. Citations, references and numbering are relative to this section.

- Unauthorized collection: collection of biometric samples without the subject's knowledge, for example using hidden cameras.
- Unnecessary collection: biometrics that are employed in situations without or with little benefit from strong user verification.
- Unauthorized use and disclosure: use of biometrics for purposes other than approved by the subject, such as forensic usage, linking or cross-matching databases, monitoring an individual's daily activity, and alike.
- Function creep: expansion of a system into areas for which it was not originally intended, for example as occurred for national identity numbers.

Unfortunately mechanisms to minimize one risk may cause another risk to increase. If for instance a system design includes stronger mechanisms that would prevent spoofing attacks by observing fingerprint patterns and finger-vein patterns at the same time the potential risk increases for a function creep i.e. that additional health-related information that is included in vein-images could be exploited.

The Data Protection Directive 95/46/EC [EE95] on the protection of individuals with regard to the processing of personal data and on the free movement of such data which is applicable to the processing of biometric data, does not provide a clear answer to these and other privacy risks of biometrics. The Article 29 EU Advisory Body on Data Protection and Privacy therefore also underlined in its Working document on biometrics of 2003 [[A29WP]] the importance of Privacy Enhancing Technologies in order to promote biometric systems that are constructed in a privacy and data protection friendly manner, minimize the collection of data and prevent unlawful processing.

2. **Security.** Biometrics are often employed to enhance the security of an application by improving the accuracy of an identity verification mechanism. One potential caveat is that the increased security may come with a decrease in privacy [CS07]. Furthermore, the incorporation of biometrics may even result in new security risks due to vulnerabilities present in the biometric subsystem. According to [JNN08], the security risks of a biometric verification system can be subdivided into four categories:
 - Intrinsic biometric failure due to incorrect decisions made by the biometric verification system, often expressed as a probability for false acceptance and/or rejection.
 - Administration attacks as a result of improper administration policies.
 - Non-secure infrastructure resulting in vulnerabilities related to non-secure hardware, software, or communication channels.
 - Biometric overtness facilitating means to covertly acquire a biometric sample from a genuine user and use that to create artifacts or any other means to influence the result of the identity verification system.
3. **Trust.** A third factor that is important for acceptance of biometric verification systems is trust. Trust differs from objective measurements such as false acceptance rates in the sense that it is a subjective property. Trust is a prediction or reliance on an action and its consequences, based on what a subject knows about an application or technology. Examples of (a lack of) trust are concerns such as health effects induced by biometric measurements (for example frequent illumination of the retina), hygiene issues (on fingerprint sensors), or the risk of stolen body parts containing a biometric characteristic (such as fingers). Some biometric modalities may also suffer from negative associations (for example fingerprints and crime).
4. **Risk mitigation.** The persistence of biometric characteristics has important consequences for the ability to mitigate risks associated with identity theft or exploited security vulnerabilities. Once a biometric characteristic has been subject to theft or abuse, it is virtually impossible to renew this characteristic. For the biometric characteristic itself, this problem is difficult to solve. However, a significant reduction of the risks associated with stolen or abused biometric characteristics can be obtained by ensuring that the biometric templates, i.e., the representation of the biometric characteristic in an identity verification system, is *renewable*.
5. **Interoperability.** Last but not least, given the large range of biometric modalities, sensor types, feature extraction types and template formats, an interoperable scheme that supports

all technology permutations of sensor, feature and template types is difficult to realize. Interoperability is especially important for large-scale open applications (e.g., biometric passports, biometric banking cards). There are efforts for standardization [37, tCM, AI, ISOa, ISOb, 17], but their scope does currently not cover a complete, end-to-end, interoperable, biometric verification system that employs techniques to protect the privacy of the subjects.

A related issue is the risk of a vendor lock in. Many biometric verification solutions exist today that are based on proprietary sensors, template formats and comparison algorithms. Switching to another vendor hence may create substantial switching costs.

3. Requirements for biometric templates

Given the challenges described in the previous section we can derive a set of requirements for a biometric verification system to ensure that privacy and security result in a positive sum [CS07], and to allow risk mitigation.

1. **Protected templates.** The representation of biometric templates that is used in a privacy-protected verification system should satisfy the following constraints:
 - It is impossible to retrieve or decode the original biometric sample(s), features or (unprotected) template from the protected template or any derivative that reveals private information from the biometric sample (such as health data, racial or ethnic origin, and alike).
 - It is impossible to uniquely link subjects within and across databases through comparison of templates.
 - A biometric template represents identity verification data for a specific, predefined purpose or application only.

These constraints should be satisfied for storage, transmission and comparison operations on templates. If a template representation satisfies these constraints, the template is referred to as a “protected biometric template”.

2. **Revocable, renewable, and diversifiable protected templates.** Protected biometric templates should support mechanisms for revocation (for example using certificates from a certificate authority). Furthermore, the template encoding process should have means for generation of multiple independent protected templates from the same, or very similar biometric characteristics. The process of generating multiple independent protected templates from the same biometric characteristics is referred to as “diversification”. This diversification property is also required to prevent cross-matching of subjects across databases, and to prevent searching for subjects with very similar biometric characteristics.
3. **Universal approach.** The protected biometric templates approach should in principle be applicable to any biometric modality, and support combinations of biometric modalities (fusion) to obtain a high verification performance. Preferably, biometric modalities can be selected and/or combined for each enrollee individually within the same application to resolve potential problems with weak biometric characteristics for a certain group of enrollees.
4. **Interoperability.** Although interoperability in general is regarded as increasing privacy risks, the biometric protected template will not allow linking data subjects across databases or across applications (see above).

Interoperability dictates that a biometric verification system should be based on a predefined format and method that satisfies the constraint given above. This format should be compatible with a wide range of sensor types and feature extraction types. It is foreseen that such interoperability can be obtained by a two-stage approach:

- Convert a biometric sample into a modality-dependent, predefined biometric feature data format that is preferably in line with existing (and/or standardized) template formats;
- Convert the modality-dependent, predefined biometric feature data to a protected template using a predefined format and method.

In this two-step approach, the intermediate predefined biometric feature data format allows the use of technology from various vendors (including sensors and feature extraction algorithms) within one system. A good example of such intermediate feature data format is the use of fingerprint minutiae data [ISOc].

The same argument holds for the second step. If the formats of the inputs and outputs to create a protected template are standardized, and the process to create protected templates is well defined (either by describing the process itself, or by using conformance criteria), the complete chain from biometric sensor to protected template could be fully interoperable.

5. **Data minimization.** For efficient storage, transmission and matching of protected templates, and to ensure maximum privacy, the amount of binary data associated with the template should be minimized while minimizing negative effects on the identity verification performance.
6. **Intrinsic security.** The verification performance of protected templates should be preferably in line with state-of-the-art biometric verification methods. A limited amount of verification performance degradation is however expected and acceptable as long as this is balanced with the gain in privacy protection. Furthermore, the trade-off between false acceptance and false rejection rates should be adjustable on an application level, and preferably also on a personal level. The latter is especially important to prevent repeating (rejection) inconvenience for persons with weak or noisy biometric characteristics.

The degree of similarity (comparison score) that was obtained during a comparison may be derived and communicated to an application, but only if there is a strong need or benefit of such information, and only in the case of a match. For a non-match, it is preferably intrinsically impossible to derive a comparison scores to thwart certain attack types that threaten security and privacy (such as hill-climbing attacks).

7. **Seamless integration with existing verification methods.** The biometric architecture should fit seamlessly to existing 2 or 3-factor verification methods (i.e., possession and knowledge-based authentication). The combination of multiple verification methods should result in a multiplicative effect on the difficulty of a zero-effort attack. The use of both application as well as user-specific secrets should be supported. The balance between possession-based, knowledge-based, and biometric security should be controllable on an application or data subject level to assure maximum system flexibility and personal convenience.

Since data subjects have in particular circumstances the right to object against the processing of biometric data on compelling legitimate grounds relating to their particular situation [EE95] such as privacy concerns, difficulty to enroll or false rejections, an authentication system should provide alternative means for authentication that is not based on biometrics. Such means should preferably be specified for failures to enroll, failures to acquire, and false rejects.

8. **Architecture flexibility.** The template protection scheme may support both on-line verification (using a central database) as well as off-line (local) verification. The architecture may provide a mode that requires both centrally stored protected template information as well as locally stored template information for successful verification.

The approach outlined in this paper is intended to provide additional security and privacy. It should be noted that some applications offer secure storage and processing by storing data on personal tokens. Such token based systems store a biometric template in a secure environment of smart cards and compare on the card [Ber08]. These applications are currently standardized in ISO/IEC JTC1 SC17 [ISOd].

4. Reference architecture

To facilitate a common vocabulary and to outline architectural aspects to meet the requirements described in the previous sections, a reference architecture for protected templates is described below. This architecture is based on so-called "pseudo identities" (or PIs). The pseudo identity life cycle, its embedding in a reference architecture and the associated interfaces and processes will be described in the next sections. The architecture is meant to be technology neutral, i.e., it should

provide a generic framework for many existing template protection techniques that currently exist and is preferably future proof. Depending on the biometric application different technical requirements will apply. Thus for a specific implementation single functional components might be left out from this reference architecture or might need to be added to it.

4.1 Pseudo identities

Pseudo identities (cf. [DCB08]) are diversifiable, protected identity verification strings within a predefined context (i.e., the protected biometric ecosystem). A pseudo identity (PI) does not reveal any information that allows retrieval of the original biometric measurement data, biometric template or true identity of its owner by any other person than the enrolled subject. Within a protected biometric ecosystem, pseudo identities follow 4 distinct phases that are visualized in Fig. 1:

1. Creation (or renewal) of PIs from biometric reference data during an enrollment phase;
2. Verification of a PI based on a recognition sample;
3. Expiration of the validity of a PI;
4. Revocation of a PI if its validity is expired.

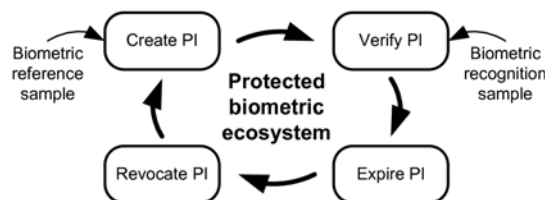


Figure 1: Pseudo identity lifecycle in a protected biometric ecosystem.

4.2 Pseudo identity creation

The Pseudo identity creation process is outlined in Fig. 2. During an enrollment phase a biometric reference is generated for an individual. In this process, a biometric capture device creates one (or more) biometric sample(s), for example in the form of an image of a fingerprint or a photo of a face. Subsequently, a feature extractor creates biometric feature data from the biometric sample. Preferably, but not necessarily, these feature data are in line with existing (standardized) template formats. Finally, a pseudo identity encoder (PIE) generates a pseudo identity and possibly additional auxiliary data (AD). Depending on the employed method and algorithms, the auxiliary data may serve the following purposes:

- It allows generation of multiple independent pseudo identities for the same individual within an application to provide renewable templates;
- it allows generation of independent pseudo identities across applications to prevent database cross-matching and linking;
- it allows generation of independent pseudo identities for subjects that have very similar biometric characteristics to prevent impersonation through spotting of biometric look-a-likes;
- it provides means for template data separation to enhance security and privacy; and
- it allows individualized comparison parameters to optimize the verification performance.

If the auxiliary data contain data elements that are associated with the diversification process, these data elements are correspondingly referred to as “diversification data”. The auxiliary data could result from various approaches that provide renewable and protected templates¹⁵. Table 1

¹⁵There may exist methods that do not strictly use auxiliary data; in that case the AD is assumed to be an empty string.

provides an overview of some existing template protection methods and their relation to pseudo identities and auxiliary data.

<u>Method</u>	<u>Pseudo identity</u>	<u>Auxiliary data</u>
Fuzzy commitment [JW99]	Hash of secret string	offset
Cancelable biometrics [RCCB07]	Transformed template	Transform parameters
Helper data systems [TAK05]	Hash of secret string	helper data
Biometric encryption [SRS98]	Cryptographic key	Filter and key link
Fuzzy vault [JW02, NJP07]	Hash of secret string	Point set P
Shielding functions [LT03]	Hash of secret string	Authentication challenge W
Fuzzy extractors [DRS04]	Hash of secret string	Public string P
Extended PIR [BCPT07]	Encrypted template	n/a

Table 1: Overview of template protection methods and their relation to pseudo identities and auxiliary data.

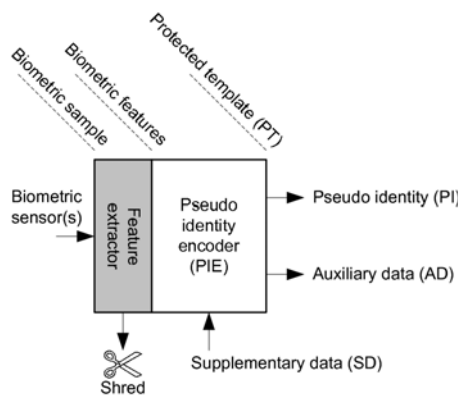


Figure 2: Creation of protected templates.

The combination of PI and AD is referred to as a protected template and hence the requirements mentioned in Sect. 3 apply to this template. Both PI and AD are stored, while the biometric sample and the biometric features are destroyed. The PI and AD can be stored in different ways that can be grouped into three categories: central storage (both PI and AD stored on a database), local storage (both PI and AD stored on a token) and hybrid storage by separating these data elements (for example by storing the AD on a token and the PI in a database). The advantages for central storage of at least one of the data elements are blacklist and audit functionalities, and simple revocation. For local storage, the advantages are the absence of security risks related to central databases and that the subject has full control over the template data. For hybrid storage, the advantages are that the subject and the provider have control over the use of the template data and that it could decrease potential security risks related to central databases (for certain attack types).

The Pseudo identity encoder takes as input some supplementary data (SD). This input can be used for various purposes, such as security enhancement by possession or knowledge-based secrets to be entered by the enrollee (cf. biometrically hardened passwords, [MRW99]); security enhancement by application or system specific secrets or signatures; limiting the scope of a PI by incorporating time and/or place-specific information for which the PI is valid and digital signature or certification of data. In any case, the supplementary data string itself is not stored with the template; it is destroyed when the PI is generated.

Some protected template technologies could be used to perform secure identification as well. This could help in a duplicate enrollment check scenario for instance.

4.3 Pseudo identity verification: PI recoder (PIR) approach

The verification process can be divided in two different classes. The first class of verification processes is based on a “pseudo identity recoder” (PIR) approach. This approach is based on the re-creation (or recoding) of a pseudo identity during the verification process, which is subsequently compared to the pseudo identity that was generated during enrollment (for example [SRS98, JW99, LT03, DRS04, TAK05, ST06, NJP07, RCCB07]). The verification is obtained by transforming a captured recognition biometric data sample to a new pseudo identity (PI*) based on the provided auxiliary data (see the left panel of Fig. 3) by a pseudo identity recoder (PIR). If supplementary data input was provided to the pseudo identity encoder during enrollment, the same supplementary data should be provided as input for the pseudo identity recoder. When the PI* is created, all input data, such as the biometric sample, the feature data and the supplementary data, are destroyed. The PI* is provided to a pseudo identity comparator (PIC) that compares both PI and PI*. Only if PI is equal to PI*, verification is successful. The advantage of this approach is that the exchange of information between the PIR (which could for example be integrated in a biometric sensor or a local terminal) and the PIC (which could reside at the application or service provider level) is in protected form (cf. [CS07]).

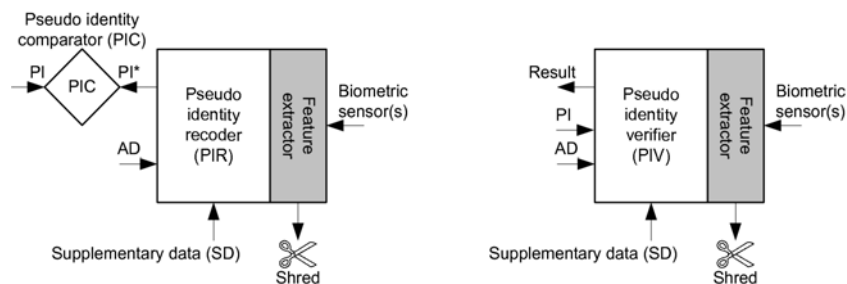


Figure 3: Verification of protected templates by PI recoding and comparison (PIR and PIC; left panel) and by direct PI verification (PIV; right panel).

4.4 Pseudo identity verification: PI verification (PIV) approach

The second class of verification methods does not rely on re-creation of a PI* during verification but rather directly verifies a PI based on a provided recognition sample (cf. [DKM07, BCI07, BCPT07]). The corresponding scheme is visualized in the right panel of Fig. 3. Given a protected template consisting of PI and AD and a sample from a biometric sensor, a pseudo identity verifier (PIV) provides the verification result. If supplementary data were provided during enrollment, the same supplementary data should be provided during the verification process to allow a successful verification. If the verification result is published, all input data is destroyed. The advantage of this approach is that no exchange or transmission of template information is required if the PIV module and the protected template are implemented on the same device, for example in a Match-On-Card solution [17, ISOd].

4.5 Pseudo Identity expiration

Pseudo identities may expire for several reasons. For example, it may have been issued for a limited period only, or may require renewal because it was compromised. Furthermore, aging effects might impact the biometric characteristic, as it is the case for the human face, which requires a renewal of the biometric reference. Validity checks and expiration can be controlled by means of watch lists. Alternatively, in some cases a validity period can be used as supplementary data for the Pseudo Identity creator resulting in an intrinsic validity check.

4.6 Pseudo Identity revocation

Depending on the implementation of a verification system, pseudo identities can be revoked by deleting the pseudo identity from a database, and/or removing the authorization to use a pseudo identity. Subsequent to revocation, re-enrollment may result in a new protected template. Depending on the architecture, this may require capturing of new biometric reference samples.

5. Architecture overview

The PI creation, storage and verification architecture is shown in Figs. 4 and 5 for the PI recoder and PI verification approach, respectively. Pseudo Identities are created during an enrollment phase. The biometric sample, biometric features and supplementary data (if these are being used by the application) are deleted when a PI is generated (or stored in a vault for later use, e.g., renewal without physical presence of the data subject). The PI and AD are published and stored on a suitable medium or different media (such as databases, smart cards, bar codes, etc). During verification in case of a recoding approach, a new PI* is generated (recoded) from the issued AD, a biometric measurement and supplementary data (if these are being used by the application). Both PI and PI* are communicated to an application to verify the claimed identity. For the PIV approach, the PIV generates a verification result without recoding a PI.

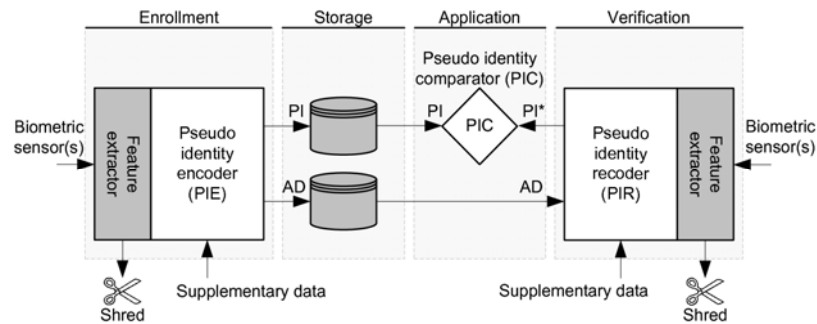


Figure 4: Reference architecture for biometric template protection based on pseudo identity recoding and comparison.

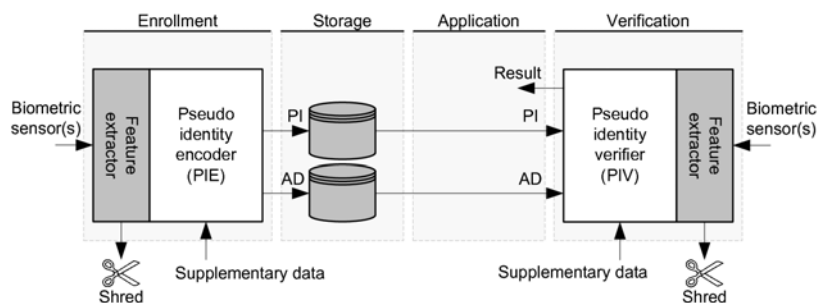


Figure 5: Reference architecture for biometric template protection based on direct pseudo identity verification.

6. Conclusions

In this paper, challenges and resulting requirements for biometric template protection methods were outlined. Based on the requirements, a reference architecture was described that describes the relevant interfaces and processes for template protection in a technology-neutral way.

7. Acknowledgments

The authors would like to thank Ton Akkermans, Julien Bringer, Jens-Petter Glittenberg, Berk Gokberk, Koen de Groot, Alty van Luijt, Johannes Midgren, Koen Simoens, Menno Treffers, Michiel van der Veen, and Bian Yang for their very helpful comments and suggestions to improve the proposed architecture and this manuscript. This work is supported by funding under the Seventh Research Framework Programme of the European Union, Project TURBINE (ICT-2007-216339). This document has been created in the context of the TURBINE project. All information is provided as is and no guarantee or warranty is given that the information is fit for any particular purpose. The user thereof uses the information at its sole risk and liability. The European Commission has no liability in respect of this document, which is merely representing the authors' view.

References

- [17] ISO/IEC JTC 1 SC 17. Application of biometrics to cards and personal identification.
- [37] ISO/IEC JTC 1 SC 37. Biometrics.
- [AI] ANSI/NIST-ITL. American national standards for biometrics. <http://fingerprint.nist.gov/standard/>.
- [BCI⁺07] J. Bringer, H. Chabanne, M. Izabachene, D. Pointcheval, Q. Tang, and S. Zimmer. An application of the Goldwasser-Micali cryptosystem to biometric authentication. In *ACISP*, 2007.
- [BCPT07] J. Bringer, H. Chabanne, D. Pointcheval, and Q. Tang. Extended private information retrieval and its application in biometrics authentications. In *CANS*, 2007.
- [Ber08] C. Bergman. Match-on-card for secure and scalable biometric authentication. In N. K. Ratha and V. Govindaraju, editors, *Advances in biometrics*. Springer, London, 2008.
- [CS07] A. Cavoukian and A. Stoianov. Biometric encryption: a positive-sum technology that achieves strong authentication, security and privacy. *Whitepaper information and Privacy Commissioner/Ontario*, 2007. available from www.ipc.on.ca.
- [DCB⁺08] N. Delvaux, H. Chabanne, J. Bringer, B. Kindarji, P. Lindeberg, J. Midgren, J. Breebaart, T. Akkermans, M. van der Veen, R. Veldhuis, E. Kindt, K. Simoens, C. Busch, P. Bours, D. Gafurov, B. Yang, J. Stern, C. Rust, B. Cucinelli, and D. Skepastianos. Pseudo identities based on fingerprint characteristics. In *IEEE 4th international conference on intelligent information hiding and multimedia signal processing (IHH-MSP)*, Harbin, China, 2008.
- [DKM⁺07] S. Draper, A. Khisti, E. Martinian, A. Vetro, and J. Yedidia. Using distributed source coding to secure fingerprint biometrics. In *Mitsubishi Electric Research Labs*, 2007.
- [DRS04] Y. Dodis, L. Reyzin, and A. Smith. Fuzzy extractors: How to generate strong keys from biometrics and other noisy data. In *Eurocrypt*, 2004.
- [EE95] European Parliament and European Council. Directive 1995/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data. <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:31995L0046:EN:HTML>, 1995. Last visited: May 24, 2008.
- [ISOa] ISO/IEC. 19092:2008 - Financial services - Biometrics - Security framework.
- [ISOb] ISO/IEC. 19794 - Information technology - biometric data interchange formats.
- [ISOc] ISO/IEC. 19794-2:2005 - Information technology - biometric data interchange formats part 2: Finger minutiae data.
- [ISOd] ISO/IEC. CD 24787 - Identification cards - On-Card matching.
- [JNN08] A. K. Jain, K. Nandakumar, and A. Nagar. Biometric template security. *EURASIP J. Advances in signal processing*, 2008. (To appear).

- [JW99] A. Juels and M. Wattenberg. A fuzzy commitment scheme. In *ACM Conference on Computer and Communications Security*, pages 28–36, 1999.
- [JW02] A. Juels and M. Wattenberg. A fuzzy vault scheme. In *Proc. IEEE Int. Symposium on Information Theory*, 2002.
- [LT03] Jean-Paul M. G. Linnartz and Pim Tuyls. New Shielding Functions to Enhance Privacy and Prevent Misuse of Biometric Templates. In *AVBPA*, pages 393–402, 2003.
- [MRW99] F. Monroe, M. K. Reiter, and R. Wetzel. Password hardening based on keystroke dynamics. In *Proc. 6th ACM CCCS*, pages 73–82, 1999.
- [NJP07] K. Nandakumar, A. K. Jain, and S. Pankanti. Fingerprint-based fuzzy vault: Implementation and performance. In *IEEE transactions on information forensics and security*, 2007.
- [Par03] ARTICLE 29 Data Protection Working Party. Working document on biometrics Working Document on Biometrics. <http://ec.europa.eu/justice-home/fsj/privacy/docs/wpdocs/2003/wp80-en.pdf>, 2003. Last visited: May 24, 2008.
- [RCCB07] N. K. Ratha, S. Chikkerur, J. H. Connell, and R. M. Bolle. Generating cancelable fingerprint templates. *IEEE Trans. pattern analysis and machine intelligence*, 29(4):561–572, 2007.
- [SRS⁺98] C. Soutar, D. Roberge, A. Stoianov, R. Gilroy, and B. V. K. Vijaya Kumar. Biometric Encryption using image processing. In *Proc. SPIE 3314*, pages 178–188, 1998.
- [ST06] B. Schoenmakers and P. Tuyls. Efficient binary conversion for Paillier encrypted values. In *Eurocrypt*, 2006.
- [TAK⁺05] P. Tuyls, A. H. M. Akkermans, T. A. M. Kevenaer, G. J. Schrijen, A. M. Bazen, and R. N. J. Veldhuis. Practical biometric authentication with template protection. In *Audio and video-based biometric person authentication*, pages 436–446. Springer, Berlin, Germany, 2005.
- [tcM] ANSI/INCITS technical committee M1. Biometrics. <http://m1.incits.org/>.
- [tw] Identity theft website. <http://www.idtheft.com>.