Jeroen Breebaart, Bian Yang, Ileana Buhan-Dulman, Christoph Busch

Biometric Template Protection

The need for open standards

Biometric-enabled applications have the potential to become wide spread. A joint strategy is required to materialize the associated benefits of cost reduction, improved user authentication and increased convenience to prevent these benefits from being overshadowed by security and privacy threats and vendor lock-ins. ISO is supporting this strategy with a new standard on privacy compliant biometric systems.

Jeroen Breebaart

Jeroen Breebaart
Philips Research,
The Netherlands

E-Mail: jeroen.breebaart@philips.com

Bian Yang

Gjøvik University
College, Norway

E-Mail: bian.yang@hig.no

Ileana Buhan-Dulman

Philips Research,
The Netherlands

E-Mail: ileana.buhan@philips.com

Christoph Busch

Hochschule
Darmstadt, Germany
/ Gjøvik University
College, Norway

E-Mail:
christoph.busch@igd.fraunhofer.de

Abstract

The objective of this paper is to outline the potential threats to security and privacy that are associated with biometric-enabled applications, to summarize the resulting requirements to ensure secure and private handling of personal data, and to explain why standardization in this area is required. The currently ongoing standardization efforts in ISO/IEC in the area of biometric template protection are described.

1 Introduction

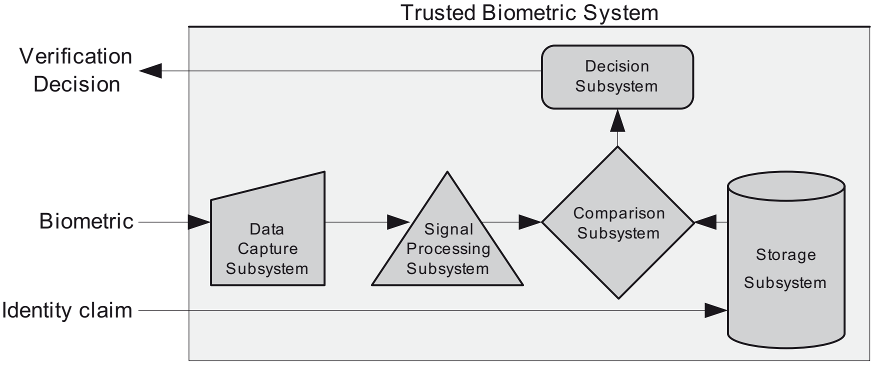
The use of biometrics for authentication or identification of individuals has been subject to extensive research during the last decades. A wide variety of biometric modalities has been investigated (e.g., iris, fingerprint, voice, gait, face, and alike). Aspects such as the recognition performance, robustness, and persistence are well documented and the limitations of specific modalities and applications have been described in various publications. The current maturity level of biometrics as authentication means has facilitated the use of biometrics in a range of applications. The most well-known applications are the use of biometric-enabled passports (for example as currently being introduced in the EU), the new EU Visa Information System (VIS) and the US-VISIT program in the USA. In the near future, the use of biometric-enabled applications may be much more profound. One could foresee the introduction of government applications (for example for municipal services or administration), healthcare (to identify beneficiaries in an undisputable way), and banking (to avoid money laundering and to improve convenience for transaction authentication).

As an example consider the use of biometrics in automated teller machines (ATMs) and payment authentication terminals in shops. Currently customers need to enter a personal identification number (PIN) as an authentication means. The fact that the PIN is not so much bound to the user but predominantly to a card makes the system vulnerable to skimming attacks. Furthermore, the PIN is inconvenient for a subgroup of users that have difficulty in remembering a PIN. The system security and the user convenience could be greatly enhanced by biometric authentication, for example by using biometric authentication only for low transaction amounts, and using combinations of biometrics and PIN for high amounts. Identity theft and withdrawals using stolen or replicated payment cards could be minimized. Additional benefits would come from shorter transaction durations and a reduced cost for banks to reset forgotten PINs.

The example of PIN authentication is of particular relevance here because the process has important similarities to biometric authentication. Both biometric and PIN authentication rely on an enrolment phase, where the subject provides a biometric sample or PIN, respectively, which is stored centrally as a reference.

During verification, a biometric probe or PIN is entered and transmitted to the central authentication facility to be com-

Figure 1 | Trusted Biometric System (TBS) model for biometric verification.



pared against the data created during enrolment. Interestingly, the PIN itself is not stored – only a non-invertible but reproducible derivative of the PIN is maintained in the central database (the PIN Verification Value, or PVV)[1]. This ensures that for security reasons, the PIN is *only* known by the legitimate owner. In the same line of reasoning, one would expect that in a biometric-enabled payment system, biometric references themselves are not stored but only a non-invertible but reproducible derivative is maintained. In fact, for biometric data such non-invertible representation is even *more* important than for PINs because biometric characteristics (1) are not renewable, and (2) introduce new privacy threats.

Before the various security and privacy threats can be described in more detail, one first needs to define what security and privacy mean in the context of biometrics. To formalize the concepts of privacy and security, a *Trusted Biometric System* (TBS) is defined. The TBS takes as inputs a biometric characteristic and an identity claim, and as outcome produces the verification decision. Hence the TBS represents the ideal biometric system, where for example all the components function as expected and the various components inside the TBS are not accessible to fraudulent attackers. The security of a TBS can be understood as the *difficulty to obtain a false accept*. Similarly, privacy can be understood as the level of protection against an attacker that tries to *obtain any other information than a verification decision from a provided biometric characteristic and a claimed identity.* For the purpose of biometric authentication, the TBS contains (1) a biometric sensor which captures biometric samples (i.e., a data capture subsystem), (2) algorithms that convert the captured information into a biometric template (often referred to as 'signal processing subsystem'), (3) algorithms that compare templates ('comparison subsystem'), and (4) a database or personal token (the 'storage subsystem') that stores the biometric data for later comparison by means of a 'comparison subsystem'. Finally, (5) a 'decision subsystem' will determine whether a capture subject is authenticated based on a comparison score generated by the comparison subsystem. The various subsystems and the TBS are depicted in Figure 1.

2 Privacy threats

Several major privacy threats related to biometrics have been described extensively. The first comprises the ability to cross match data subjects across different services or applications by comparing biometric references. The persistence and uniqueness of biometric characteristics allow a malicious person to link users between different databases. For example, an attacker could link different financial service records across different banks' databases to one specific customer to illegally obtain the customer's financial condition or investment plan.

The second threat to privacy is the possibility to extract sensitive information from the stored biometric data, such as a subject's ethnic background or (the probability for) certain diseases. Such data could in principle be abused by healthcare insurance providers (for example biometric references that are intended for patient identification in a hospital could be used to differentiate in insurance premiums). Similar threats could occur in other application areas as well.

A further privacy risk associated with biometrics is often referred to as function creep. If the application scope of biometric technology is not well defined and restricted, its use may expand to other applications or services. For example, an application initially intended to prevent misuse of municipal services may gradually be extended to rights to buy property, to travel, or the right to vote. As a consequence, data subjects that would agree to use biometrics for the initial application would be forced to use biometrics for other applications.

These privacy issues are covered by legislation. The processing of sensitive data is prohibited under EU jurisdiction. For example Article 29 EU Advisory Body on Data Protection and Privacy underlined in its Working Document on biometrics of 2003[2] the importance of privacy protection for biometric systems to prevent unlawful processing of collected data.

3 Security threats

The various security threats can be described by potential attacks on the aforementioned subsystems, as shown in Figure 1.

Data capture subsystem attacks

The most serious threat on an input device is presenting a fake biometric characteristic. The fabrication and presentation of a fake physical biometric characteristic is called a *sensor spoofing attack*. It is known that some characteristics are harder to forge (such as the iris, a retinal scan, or a face thermogram) while others are easier to forge (voice, face, hand written signature or fingerprint).

Since biometric information cannot be regarded as secret, the original biometric characteristic can be obtained with or without the permission or cooperation of the data subject. For some modalities, like fingerprints, extensive literature on how to spoof sensors is readily available. The biometric system and embedded liveness detection methods must thus be able to prove that the captured biometric sample came from the correct subject at the time of verification. The sensor spoofing attack

[1] The PVV is often also stored on the magnetic stripe of a credit or debit card for offline authentication.

[2] ARTICLE 29 Data Protection Working Party. Working document on biometrics, 2003.

can be implemented as a *coercive* or *impersonation attack* depending on the amount of tampering applied to the capture subsystem. A *coercive attack* is an attack where the authorized data subject's biometric data is presented in an illegitimate scenario. For example, the attacker may physically force a genuine subject to present his/her biometric characteristic in an authentication setting. System designers have to consider how to counter such attacks, for example by installing security cameras at ATMs. An *impersonation attack* involves changing one's appearance so that the measured biometric data matches an authorized individual. This attack could be conducted for biometric modalities such as face,- voice- or signature recognition. The use of gummy fingers as artificial biometric, and the reactivation of latent fingerprints present on a sensor are also well-known examples. Multi-modal biometric systems reduce the exposure to an impersonation attack (assuming the system is checking for consistency between the modalities). Furthermore, challenge-response biometric protocols may be employed.

Signal processing subsystem attacks
Signal processing subsystems are vulnerable to the intrusion of imposter data during processing for example by means of a *trojan horse* attack. Care must be taken during the employment of the system to avoid these threats, for example by using approved or certified algorithms and components.

Comparison subsystem attacks
Similar to the signal processing subsystem attacks, algorithms or components can be modified to produce other results than intended. For example, attacks could focus on changing the comparison scores resulted from the comparison subsystem.

Storage subsystem attacks
Various security threats are related to tampering with the storage subsystem. Biometric references can be accessed illegally, or could be replaced or changed. The unauthorized access or modification of biometric references may not only lead to security threats, but may also extend to threats in the privacy domain, such as cross matching of databases. Therefore, protection of the storage subsystems is key for security and privacy reasons.

Decision subsystem attacks
The decision subsystem is potentially vulnerable to *hill climbing* and *threshold manipulation* attacks. In the first category, an attacker may present an initial biometric characteristic and observe the corresponding comparison score. Depending on the value of the score, the presented biometric characteristic is modified and the resulting scores are monitored. This allows attackers to iteratively change the biometric input until a successful verification is obtained. In a threshold manipulation attack, the attacker is able to modify the comparison threshold to enforce a "correct" verification.

Transmission and other attacks
Biometric data, comparison scores and decisions are often transferred in open and distributed systems between various subsystems. A complete biometric system including data transmission is potentially vulnerable to *eavesdropping, replay, brute force,* and *man-in-the-middle attacks*. A more extensive overview of these and additional attacks is given by Buhan (2008)[3].

4 Requirements

4.1 Privacy requirements

Safeguarding the privacy of individuals in a biometric context is challenging. A set of requirements follows below.

Identity privacy
Storage of biometric references accompanied by other identity data results in significant privacy risks. The binding between biometric and other identity data allows malicious persons to link data subjects to applications beyond those using biometrics. Hence it is crucial that the binding between biometric and other identity data is securely protected.

Irreversibility
To prevent the use of biometric data for any other purpose than originally intended, the biometric data should be transformed in such a way that the biometric sample cannot be retrieved from the transformed representation. In other words, such transform should be irreversible, preferably without compromising the biometric verification performance. Irreversibility should hold even when several biometric references are accessible from different applications, services or databases.

Unlinkability
Tracking and tracing subjects across applications should be eliminated by ensuring that biometric references used in various applications are unlinkable. This guarantees that no adversary has a significant advantage over random guessing in determining whether two biometric references are related or not – meaning that they were generated from the same source.

4.2 Security requirements

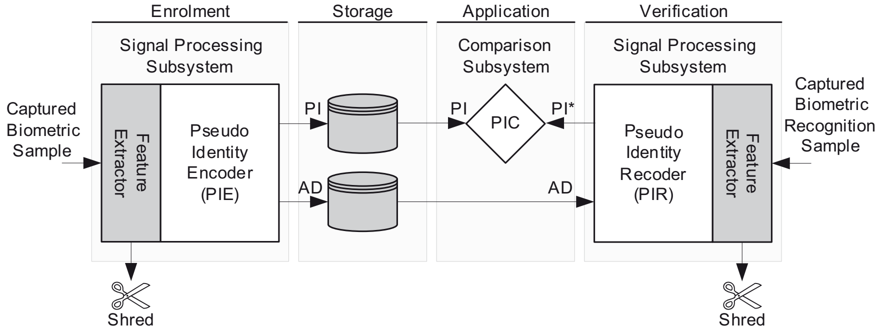
Confidentiality
Confidentiality ensures that information is not disclosed to unauthorized entities. In a biometric system, biometric data is stored and transmitted between various subsystems. Both storage and transmission of data should be protected against eavesdropping, unauthorized disclosure, or modification of the data. This requires cryptographic techniques such as symmetric or asymmetric ciphers.

Integrity
Integrity, in general, is the property of safeguarding the accuracy and completeness of assets. If the integrity of a biometric reference or the result of the various processing subsystems are untrustworthy, the verification outcome will also be untrustworthy. Therefore, cryptographic means to protect the integrity of the data (such as signatures or authenticated encryption, possibly extended with time stamping) are required.

Renewability and revocability
A strong security concern for biometric system relates to renewability and revocability of biometric references. Individuals have a limited number of irises and fingers; identity theft renders corresponding biometric references as unusable for future use. Due to the persistence of biometric characteristics, a biometric reference that is compromised once is compromised forever. The risk of compromised biometric references can be mitigated for certain types of attacks by providing methods to allow renewable biometric references. If various different biometric references can be extracted from the same or similar bio-

[3] Buhan, I. (2008) Cryptographic keys from noisy data – Theory and applications. PhD thesis University of Twente, The Netherlands.

Figure 2 | Architecture for renewable biometric references consisting of pseudo identities (PIs) and auxiliary data (AD). The data capture subsystems are omitted for clarity. PIs are created during enrolment using a pseudo identity encoder (PIE) based on extracted features from a captured biometric sample. During verification, a new PI* is created in a pseudo identity recoder (PIR) based on a captured biometric recognition sample.


Enrolment
Storage
Application
Verification
Signal Processing Subsystem
Comparison Subsystem
Signal Processing Subsystem
Captured Biometric Sample
Feature Extractor
Pseudo Identity Encoder (PIE)
PI
AD
PI
PIC
PI*
AD
Pseudo Identity Recoder (PIR)
Feature Extractor
Captured Biometric Recognition Sample
Shred
Shred

metric characteristic, the biometric reference can be revoked and renewed in case it has become subject to identity theft.

5 The need for an open standard

As outlined in the previous sections, the procedures, mathematical transforms, data elements and storage/transmission aspects for biometric data can be quite complex and elaborate to meet all requirements for security and privacy. Several proposals for such data transformations have been published in the literature. These proposals are often termed “biometric template protection” or “biometric encryption” schemes and typically address one or more of the privacy and security requirements described in the previous section. Interestingly, from a high-level point of view there are many similarities in the structure of these proposals, which are often difficult to recognize because most of them employ their own terminology. Therefore there is a need to describe template protection schemes using a harmonized terminology and structure. Such “standardized” structure would also help in translating the privacy and security requirements into technical properties of the various involved processes and data elements, and allow verification of the extent to which the various requirements are met.

The most convincing argument however for defining a standard on biometric template protection is related to *interoperability*. From a technical perspective it is virtually impossible to authenticate a biometric measurement using template protection method B if the reference data were created using method A. Consider the exemplary case of payment authentication in which a client enrols at a local bank. To protect privacy and enhance security, template protection method A is being used to generate a protected biometric reference. This means that all payment terminals that exist in that service context should support the same template protection method A. Of course one could consider to support a certain set of template protection schemes at each payment terminal, but it is virtually impossible to support a large number of methods due to technical difficulties and required effort for such an implementation, and to acquire all licenses of the various technology elements. If an interoperable system for biometric payment authentication (or any other large-scale biometrics-enabled system) would materialize, it is preferred to develop an *open standard* to ensure that the various technology elements can be properly designed, implemented and reviewed, to ensure vendor neutrality to prevent vendor lock-in and the associated switching costs.

6 ISO/IEC WD 24745

A first step towards the creation of an open standard for biometric template protection is currently being taken by ISO JTC1 subcommittee 27 (SC27) workgroup 5. This workgroup deals with identity management and privacy technologies in the area of IT security techniques and is developing a standard for cryptographic guidance to protect biometric data. As of February 2009 a completed working draft (WD) *ISO/IEC 24745 – Information technology – Security techniques – Biometric template protection* has been issued. This document aims at describing the potential threats and requirements with respect to data confidentiality, integrity, availability and renewability of biometric references during storage and transmission. Furthermore, the binding between biometric data and other personally identifiable information (such as identity data, contact information, account numbers, and alike) is described and the associated privacy requirements are formulated.

Two important aspects that are described in WD 24745 will be outlined in this paper. These comprise:

1. A framework for renewable templates that describes the requirements for irreversibility and unlinkability; and
2. Means for encryption and signatures to ensure confidentiality and integrity of biometric data.

6.1 Renewable, unlinkable and irreversible biometric references

ISO/IEC WD 24745 describes the framework for renewable, unlinkable and irreversible biometric references based on the concept of pseudo identities (PIs), which are anonymous and renewable biometric identity verification strings within a predefined context[4]. A pseudo identity does not reveal any information that allows retrieval of the original biometric measurement data, biometric template or true identity of its owner (i.e., it is irreversible and unlinkable). Furthermore, it is renewable – a very large number of independent PIs can be generated from the *same* biometric measurement. In essence, a pseudo identity has a similar function as the PVV in PIN authentication: it is renewable, it does not reveal the PIN itself, and can be revoked.

A pseudo identity is created during a biometric enrolment process. The elementary parts of this process are visualized in the left part of Figure 2. One or more captured biometric samples are processed by a feature extractor to generate a set of features having discriminative properties. This part is in essence the same as a con-

[4] Breebaart, J., Busch, C., Grave, J., Kindt, E. (2008). A reference architecture for biometric template protection based on pseudo identities. In: Gesellschaft für Informatik (GI): BIOSIG 2008. Proceedings of the Special Interest Group on Biometrics and Electronic Signatures. Editor: Brömme, A. Bonn: Gesellschaft für Informatik, 2008, pp 25-37.

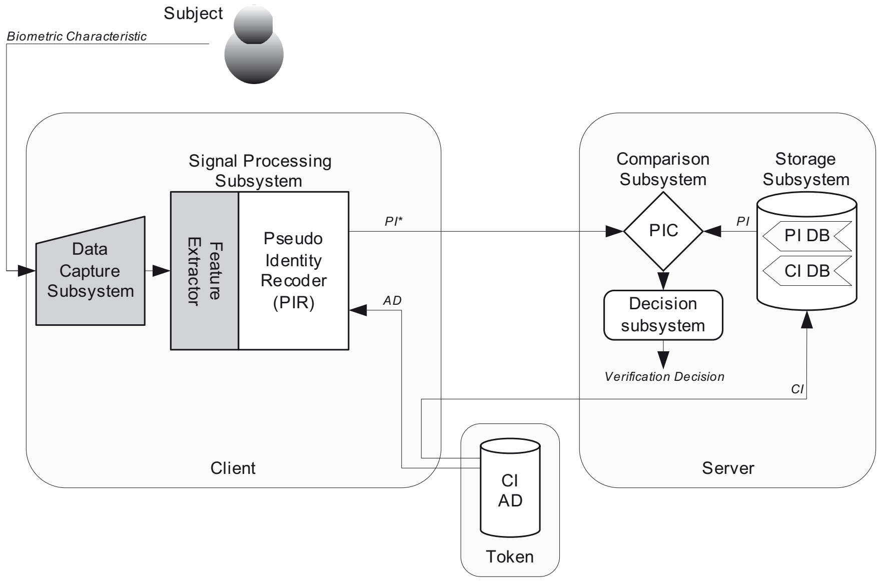
ventional biometric enrolment process. Subsequently, a pseudo identity encoder (PIE) generates the renewable biometric reference comprising a PI and auxiliary data (AD). Note that the captured biometric sample and the extracted features thereof can be discarded as soon as the PI and AD data elements are created.

The PI and AD elements are stored for later use. Storage can be implemented in a central database or on a token carried by the user. Data separation by storing the PI in a central database and AD on a token is feasible as well. The large advantage of such data separation approach is that both the data subject (carrying the token containing AD) and the service provider (which has access to the PI) should cooperate to produce a successful biometric verification. This can be explained by considering the verification stage of Figure 2. During this process, a captured biometric recognition sample is processed by a signal processing subsystem, consisting of a feature extractor and a pseudo identity recoder (PIR). The PIR generates a new PI* based on the extracted features and the AD that was created during enrolment. Only if the correct AD and biometric characteristic are presented, the reconstructed PI* will match the PI generated during enrolment. In all other cases authentication will fail.

In some verification scenarios, the pseudo identity recorder (PIR) can be combined with the pseudo identity comparator (PIC) into one logical or physical component called pseudo identity verifier (PIV), which directly outputs a binary verification outcome without need to send a reconstructed PI* to a PIC.

Depending on the architecture of the biometric system, the storage, signal processing and comparison subsystems may be divided across a token or smartcard, a client and a server. WD 24745 describes a set of *models,* where each model targets a different system layout that fits one or more applications. One (simplified) example of such a model is shown in Figure 3. In this model, the signal processing subsystem is situated in a client, the comparison and decision subsystems are positioned on a server, and storage is available in both a smartcard/token (for AD) and the server (for PI). In this model, the client can be a kiosk for border control or registered traveler checkpoint in an airport, or it could be a payment terminal in the public sector. When authentication is initiated, the AD and a common identifier (CI) that are stored on the token/smartcard are transferred to the client. The CI could for example be a traveler number or an account number. Secondly, the data subject presents the biometric characteristic to the data capture subsystem. The signal processing subsystem generates a PI* based on the provided AD and biometric sample. The PI*, accompanied by the CI are transmitted to the server. The CI is used to query a database with PIs. The PI associated with the CI during enrolment will then be compared to the PI* generated by the client. If these match, the verification is successful.

Figure 3 | Example model for biometric verification using a token, a client and a server.



WD 24745 emphasizes the renewability of the protected biometric references. Renewable biometric references should support mechanisms for revocation and generation of multiple independent references from the same or very similar biometric characteristics. The renewability is often based on the generation of a secret key during enrolment that can be randomly generated for every application. Both PI and AD depend on this key, but do not reveal the key itself. More specifically, to meet the privacy and security requirements, the following properties should apply to protected templates:

- Sufficient entropy in the generated data. This requirement is needed to allow a sufficient number of diversifications for a single person biometric characteristic, and to prevent brute-force attacks.
- Low mutual information between the biometric features and the protected biometric reference, to prevent information leakage about biometric characteristics. This guarantees the irreversibility of the protected biometric and prevents key-inversion attacks.
- Low mutual information between protected templates derived from equal or very similar biometric features, which is required to prevent cross-comparison of subjects across applications and databases, and to prevent searching for subjects with very similar biometric characteristics.

6.2 Confidentiality and integrity

As described in Section 4, the unlinkability, renewability and irreversibility requirements should be extended with data confidentiality and integrity to ensure that stored and transmitted data is not accessible to unauthorized persons, and that the data cannot be tampered with. These requirements can be met using existing cryptographic techniques. To provide confidentiality, symmetric, asymmetric, block or stream ciphers can be used. These techniques are used as an additional layer on top of the methods to provide unlinkability, renewability and irreversibility. More specifically, the PI and AD are preferably encrypted when stored and transmitted from one subsystem to another. However, encryption introduces the problem of key management which is one of the most difficult problems in cryptography and refers to generation, exchange, storage, safeguarding, use, and replacement of keys. Various solutions can be implemented that differ in in terms of key management. For example if a personalized key is used to protect the PI and AD of an individual and the key is controled by the data subject, he or she has to present the key along with the biometric characteristic during authentication; if on the other hand the key is controled by the authority who authenticates the data subject, measures have to be taken to store each individual's key securely. Data integrity can be provided by means of a signature or a Message Authentication Code (MAC). Alternatively, authenticated encryption can be used as standardized in ISO/IEC 19772[5].

7 Conclusions

Biometrics improve the security of multifactor authentication systems but impose several threats to data privacy. More specifically, biometric systems can be subject to various types of attacks by malicious persons. To prevent these from overshadowing the benefits of improved authentication and user convenience, it is crucial that methods for biometric data storage and processing satisfy the requirements of renewability, irreversibility, unlinkability, confidentiality and integrity. These requirements in turn result in a strong need for standardization of architectures and technology in this area.

8 Acknowledgements

The authors would like to acknowledge their colleagues Koen de Groot and Emile Kelkboom for their helpful comments to improve this manuscript. This work is in part supported by funding under the Seventh Research Framework Programme of the European Union, Project TURBINE (ICT-2007-216339). 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.

[5] ISO/IEC 19772: Information technology – Security techniques – Authenticated encryption.