GDPR and Privacy-by-Design

By Alty van Luijt
Executive Director, R&D

Thinking about data privacy and its implications when developing products and solutions is extremely important in today's increasingly digital world. In Europe, the General Data Protection Regulation (GDPR) became operational in 2018. It has made Privacy by Design a legal requirement, underscoring its prime importance for data privacy and protection. The African Union has also taken the initiative to develop future-proof identification ecosystems that enable quality service delivery in Africa, very much on similar lines as GDPR. This means it is critical to engineer Privacy by Design features into ID schemes from the start. With virtually all ID schemes in Africa now involving a kind of biometric modality for enrolment and identity verification, considerations on how to protect biometric data will be key. This article highlights advances in biometric data protection and some useful applications.

Why Biometrics?

Biometrics is a powerful and convenient way of identity verification. It is the only way to confirm physical presence of a person. But it has several properties that make it vulnerable, more so than alternative ways of authentication. Biometrics represent some intrinsic properties of the individual, so it is to be considered as PII (Personal Identifiable Information). Furthermore, it is a finite resource, as you have only one face, two eyes and ten fingers, and basically there is the issue: Once compromised, compromised forever.

In this note, we'll use the example of fingerprints, as this is the most commonly used biometric modality. The traditional approach to fingerprint matching uses minutiae templates. These are typically stored and transmitted in an encrypted form, but matching can only be done in cleartext, so the templates will have to be decrypted just before being matched, causing an inevitable vulnerability. Software exists that can re-synthesize a credible fingerprint image from a minutiae template. So, the best embodiment of this traditional matching process is Match-on-Card, where storage, decryption and matching all take place inside a well-protected Smart Card. Unfortunately for some use cases the use of a high-end Smart Card can be cost-prohibitive, whereas more and more use cases bypass the use of cards altogether, as in mobile authentication.

Privacy by Design

One of the basic tenets of Privacy-by-Design is that protection needs to be intrinsic, not bolted on after-the-fact, as in the example above. In 2018, the EU mandated the use of Privacy-by-Design principles for Biometric data in the GDPR (General Data Protection Regulation). Thus, we are looking for a completely different approach to fingerprint matching, one where there is never a cleartext image or minutiae template exposed.

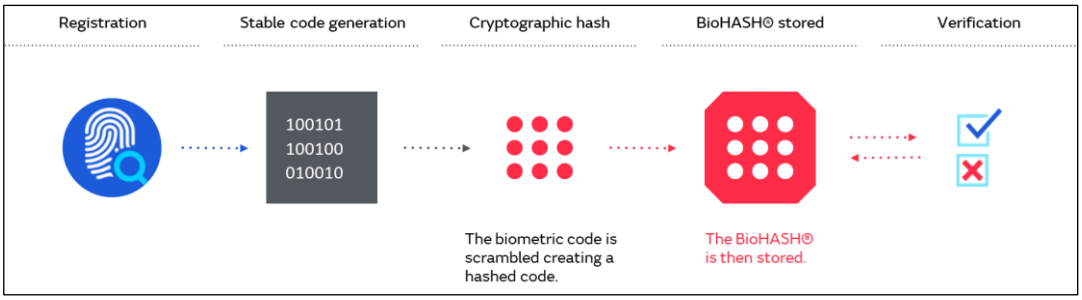
At first sight this would seem like a mission-impossible, but fortunately there is a branch of mathematics that has some fascinating properties when it comes to secure processing: cryptography.

We'll revert to the specific tools that are available in cryptography in a later section of this note, but they all share a common property: they need stable data as input for processing.

Feature vector processing

If we want to apply cryptography to fingerprints, we have a problem: every specific imprint of a finger is slightly different. It depends on the type of sensor, the pressure on the sensor, placement of the finger, the humidity, etc. In scientific terms, the fingerprint data is described as "noisy".

The challenge that this presents to the development of Privacy-by-Design schemes for biometric data, is to derive a stable and reproducible number from the noisy biometric measurements. This can be done by extracting so-called features from the fingerprint images that contain the key differentiating properties. Starting from a large set of features, these are then processed and transformed into a compact template that contains all the differentiating



information in a condensed form. After the feature extraction, the original fingerprint images can be discarded, since they are no longer required in the rest of the process. The technology to create them borrows heavily from digital image processing, signal processing, pattern recognition, information theory, error correction and statistics. The net result is that the resulting feature vector has a constant value, even when the input was a "noisy" biometric fingerprint image. BioHASH© is the GenKey name for this family of technologies.

Applications

Having the ability to derive a constant number from biometric input opens a whole arsenal of applications, using cryptography as the main tool. Here are some examples:

- *Hashing* is a cryptographic one-way function. This is used as an integrity check when downloading software modules, or when validating passwords or PIN codes in a computer system. The comparison is then done in the hashed domain, so not even the system administrator of the computer system has access to the cleartext password. The biometric equivalent of this is called BioHASH©. It is a mature technology by now, standardised in ISO 24745 (ISO standard for biometric information protection), and deployed in millions of cards.

Ghana's National Health Insurance Authority (NHIA) and Kenya's National Hospital Insurance Fund (NHIF) have been enrolling citizens using GenKey's hashing technology since 2013. More than 18 million cards with no biometric data stored on them have been issued so far. The BioHASH© template, a hash of individuals' fingerprint is stored instead, thus protecting the privacy and security of their identities.

More than 18 million cards with no biometric data stored on them have been issued so far. The BioHASH© template, a hash of individuals' fingerprint is stored instead, thus protecting the privacy and security of their identities.

- *Digital signatures.* By digitally signing the hash of a document with BioHASH©, a document can be validated by a digital signature. In the Healthcare example, we generate a Claim Verification Code (CVC) that gets copied to the claim form that goes to the insurance entity. It contains info on the healthcare provider, and the time and date where the client verified himself on the terminal in the provider's office. Because the CVC can only be generated if the member is present, this digital signature provides a safeguard mechanism against patient

impersonation as well as phantom claims by the health provider.

A key aspect of Ghana's revitalized National Health Insurance Scheme (NHIS) implementation is biometric verification at healthcare providers. When a cardholding member visits a healthcare provider, their fingerprint data is captured and matched against the fingerprint template stored on their Insurance card during enrolment. In case of a biometric match, a unique, irreversible claim verification code (CVC) is generated using GenKey's unique BioHASH© technology and added to the claim form as a biometric signature. Through this implementation, the NHIS has saved several millions of dollars in payment of fraudulent claims by health service providers.

- ***Private keys in asymmetric cryptographic protocols.*** This is based on private/public key pairs, where the private key is derived from the biometrics. It is never stored but regenerated on-the-fly when a fresh fingerprint sample is being processed. In the public/private authentication protocol, the server can be assured that the user has presented his biometrics, so this is a very appropriate solution for mobile applications that want to have access to on-line services.

- ***Self-Sovereign Identity (SSI) management of Biometrics.*** SSI uses the concept of a wallet of identity attributes that is under total control of the individual. BioHASH© offers the possibility to also use biometric information with the identity wallet. Pointers to certain public keys related to the wallet can be stored in an immutable public ledger, also known as Blockchain. SSI gives the power to the user to control what he discloses to a verifier.

  Self-Sovereign Identity is the future of digital identity management. Its objective is to take control away from the large corporates that collect personal data on a massive scale, and to return control to the individual by emulating the way we use credentials in the physical world. Rather than trying to create a description here, we'd like to refer you to the white paper and video that can be found here: www.sovrin.org

Conclusion

BioHASH© comes from the family of Biometrics/Cryptographic tools that enables true Privacy-by-Design. It allows the creation of truly "GDPR-compliant" solutions that handle biometrics in a secure and privacy-preserving way. For more details and expert advice, please check www.genkey.com.