Electronic signature and certification models in health care.

F. De Meyer, G. J. E. De Moor
Dept. of Medical Informatics, University Hospital of Ghent, Belgium

1 The legal context

Electronic signature legislation has been introduced for the first time in Belgian history in the social security Royal Decree of 16th October 1998. This decree was put in place to allow the use of the so-called SIS card or social security card. The goal of the introduction of the SIS card is to reduce the administrative formalities for the socially insured, the employers, care-providers and the administrations of the social security services. The card is intended to be multifunctional within the domain of social security. The card serves mainly as a means to unambiguously identify the socially insured in an electronic way and as a proof and indication of the social insurance status of the holder of the card. It is not a smartcard because it has no processing capabilities. It is a storage card. Nevertheless, the SAM card that is used to access the SIS card is a smartcard but is not a health professional card as such. Moreover, use of the card is limited to social security purposes.
On a European level, the directive on a community framework for electronic signatures is a reference document that all European countries have to implement.[1]
The bill on the operation of certification service providers for the application of electronic signatures (Belgian Chamber of Representatives, Doc 050322/001) dating from 16.12.99 is one of the documents aiming to implement the European directive.
Moreover, Belgian civil law has been adapted so that article 1322 expands the meaning of the term 'signature' to include a digital collection of data that can be created electronically.
All legislation concerning electronic signatures is built on two pillars:

- the assimilation of the terminology of electronic signatures into existing legislation and handwritten signatures.
- the creation of a legal framework for certification.

The terminology used in legislation is 'electronic signature' in order to stay as generic as possible and to avoid adaptations in legislation as technology evolves.
It should be stressed that the European directive does not make accreditation of certification authorities compulsory. Electronic signatures will have to be accepted in litigations, but their value of proof is still subject to evaluation. Public sector applications, such as health care, can however dictate specific signature and certification measures.

2 Background information on digital signatures

2.1 Cryptography basics

The term ‘electronic signature’ covers a very wide area and designates all kinds of signatures in electronic form.[2] Currently however, a subgroup called ‘digital signatures’ is most widely used. Digital signatures are based upon the principle of asymmetric cryptography. Cryptography is based upon an algorithm that transforms clear text into cipher text on the sending side and vice versa on the receiving end. The same algorithm is used on both sides but with different keys. A cryptographic key is a string of bits controlling the behaviour of the algorithm. Two different keys are needed in asymmetric encryption: one for the encryption and one for the decryption. The keys are different but not independent. They are the outcome of the same generation round. One of the keys, called the private key, has to be given to the key holder only and the other key, called the public key has to be made available to the community of users that wants to share cryptographic functions with the key holder.
Digital signatures are made with the private key of the key holder. The result is a string of bits representing the signature. This bit string, representing the signature, is the outcome of the signature function that has two input values:

- The message or text that needs to be signed.
- The private key of the signing entity.

The receiving end can check the signature when it uses the same cryptographic algorithm and has the following two elements:

- The public key that is associated with the holder of the private key that was used to create the signature that is under investigation.
- The message that was used to generate the signature.

Besides cryptographic algorithms, digital signatures use a volume-reducing algorithm, called ‘hash function’, that reduces an input file of arbitrary length to a string of a predetermined length.

2.2 Objectives of an electronic signature

Electronic signatures are used on documents and messages for the following purposes[3]:

- To ensure the integrity of the content of the document or the message. Electronic signatures do not prevent changes to the content but any change will be detected automatically when the message or the document is opened. These changes can be due to human intervention (falsification) or to errors.
- To identify and authenticate the author of a document or the sender of a message.
- To express the consent of the author with the content of the document. The consent objective is specifically mentioned in electronic signature legislation.

The above elements can be found back in all tutorials on electronic signatures. However there is another element that is important when interpreting and attributing value to a signature: the role of the person or entity that is signing. A role can in its simplest form reflect a title or rank held in an organisation but can also implicitly, through reference to the policy of the organisation reflect capabilities and constraints (e.g. a limit to sign contracts up to a certain amount).

2.3 Trustworthiness of digital signatures

A first group of elements that determine trust and assurance in digital signatures is intrinsically determined by the choice of algorithm, hash function and key length.

- Digital signatures are based upon the principle of asymmetric key encryption. There are various algorithms that can be used for digital signatures. The algorithm should be sufficiently robust to withstand all sorts of attacks.
- The time needed to break a signature with a brute force attack (i.e. by trying all combinations) increases with the key length. Therefore, sufficiently long key lengths should be used, especially when the signature should last for several years.

A second group of elements depends upon the infrastructure, methods and policy used to generate and check signatures.

- The cryptographic process producing the electronic signature should be executed in a highly secure environment. Smartcards provide such an environment. A human readable and keyable pincode is used to unlock functions and data inside the card. The string to be signed is sent into the card where it is signed and sent back to the calling process on the PC where the application requesting the signature is running.
- The key holder who puts his signature on an electronic document must have certainty about what he is signing.
- The key holder must be aware in what capacity and for what purposes he is signing; He must also be made aware of the consequences of the deliberate misuse of roles.

A third group of elements on which trust depends has to do with key and certificate handling.

- A person verifying a signature must be sure that the public key he is using belongs to the key holder that has created the signature.
- A private key should only be known by the holder of the key. Since such a key is not in a form that allows human input via a keypad, it is stored inside a smartcard.
- When the holder of a key suspects that the secure use of his key pair has been compromised (e.g. his card has been stolen, a certificate containing the public key has been falsified) he must be able to revoke trust in his certificate(s).
- A third instance (e.g. the employer or the organisation that has issued the card) can revoke the user's certificate(s) if he is no longer associated to the third instance.
- Usually certificates are automatically revoked after a predefined period of time.

3 Public key infrastructure

This section describes the functions that are needed in a PKI (public key infrastructure). PKI services are not limited to digital signature keys but deliver services to all cryptographic functions that require keys. All services listed below (except perhaps the directory services) are provided by Trusted Third Parties (TTPs).

3.1 Registration

A distinction should be made between the registration of attributes (e.g. a role or quality) and the registration of cryptographic keys.
When a user already has an electronic identity and cryptographic keys associated to that identity, issuing attribute certificates stating e.g. roles is less complicated. The (attribute) Registration Authority (RA) only has to keep a list with attributes and the identity they are attributed to.
In some cases, the health professional has to go to the (attribute) registration authority himself to present his letters of credential (e.g. a paper based diploma), but if a professional registration body already has such a registry, it can issue attribute certificates based on that register. It should be noted that according to European legislation, keys or certificates should not be created without at least informing the holder of the certificates.

When however, an electronic identity still has to be established and cryptographic keys need to be generated, the user requesting the registration must present himself physically at the registration office with his letters of credential, and ID card or passport as required.
The registration authority then checks the credentials. When approved, the request is passed to the certification authority (CA).

3.2 Certification

A certification authority (CA) binds attributes and/or keys to identities. The verification process that justifies the certification has already been carried out by the registration authority. However, in practice, certification and registration functions are sometimes done by the same organisation.
An electronic certificate is an electronic document containing a number of statements and is signed by the certification authority. Standards exist that define the form and content of electronic certificates. The best know is the ITU-T X.509 certificate standard, which itself comes in three versions, that are backward compatible.
An example of a certificate is a public key certificate that binds a public key to the identity of the key holder.
A unique identifier expresses the identity. Each certificate itself has a certificate ID.

3.3 Key generation

Before cryptographic keys can be issued, they have to be generated. The key generation is usually done by the certification authority, but another Trusted Third Party can deliver this service as well. Keys are generated in a highly secure environment. Cryptographic keys are generated in pairs. The public key part of the key pair has to be bound to the ID of the key holder in a public key certificate, delivered by the CA. The private key part has to be stored on a smartcard. The smartcard itself is protected by a pincode.

3.4 Directory services

Certificates contain a statement that is vouched for by the issuing CA. It binds either a public key or an attribute to the identity of the holder. Certificates are needed at verification time of the attribute or of the signature made with the private key part.
Apart from perhaps the availability and up-to-date aspect, directory service provision is not a trust service. Once generated by the CA, a certificate cannot be tampered with anymore. But the certificate has to be electronically available over a network so that users can check the validity of a certificate and download the content of the certificate.
As explained in the section on ‘certificate revocation’, the directory services also play a role in certificate revocation by providing certificate revocation lists.

3.5 Card issuing and personalisation

Card issuing and personalisation is a service that is often delivered by the CA, especially when the CA does the key-pair generation as well. A smartcard contains a file structure and the private cryptographic keys are written is a protected part of the smartcard. Additionally but not necessarily, a copy of the certificate(s) can be written on the card as well.
The upper surface of the smartcard usually contains information as well. That information is textual but usually also contains graphical information such as a picture of the holder and a logo.

3.6 Certificate revocation

Every certificate has a validity period stating the start and end dates of its lifetime. However due to a variety of reasons a certificate can be revoked before its validity end date. Certificate management is a function that is normally provided by the certification authority. There are various methods for propagation of revocation information. One method is to reflect the validity in the certificate itself, another method is through the use of certification revocation lists that are published on the directory services.

4 PKI models

The previous section deals with a number of PKI functions. However, there are various choices to allocate these functions.

4.1 Centralised ID card model with decentralised role attribution

The starting point for the centralised ID card model is that an electronic national identity card is issued to all citizens and residents. This ID card has the form factor and functionality that can be expected from smartcards and includes at least:

- Signature function.
- Encryption function.
- Authentication function.

Such an ID card contains at least the name, a national identification number, gender, date of birth and nationality of the card holder and identification of the issuer. On the outside, it should contain a clear picture of the holder. A Belgian ID card does not necessarily mean that the holder has the Belgian nationality. It can, for instance, be a permit that allows a person to stay in the country for a specific reason and a predetermined period of time.
Signatures created with this card have nationwide (and possibly even cross border) legal value. The card is used whenever the cardholder has to authenticate himself or sign a document, be it in the private or public domain.
The certificate that contains the public key part associated with the user is put on one or more directory servers that can be accessed publicly and without charge by any verifying identity.
The certificate only contains the identity and key related information of the cardholder.
The meaning of an electronic signature in that context is equivalent to the handwritten signature that is currently used. It proves an identity and not a role.

The model assumes that each health professional in Belgium holds a Belgian ID card. This card can be used for signature and authentication purposes. Authentication means that the holder can prove his identity electronically by means of the card.
Health professional specific functions can be built on top of this platform. For health professional contexts at least the following elements are needed:

- Certification and statement of the professional or qualities of the holder. For instance: “Mr. X. is a Medial Doctor, specialised in cardiology”.
- License to practice and by whom it is issued.
- One or more roles.

The first two elements are generic and usually invariable for a long time. The administration can be done nationally or regionally by e.g. the order of physicians or a governmental institution. The requirement is that these organisations maintain (or obtain) registration information about all health professionals under their control.
The role of a health professional may have significance over a large geographic area but is usually assigned by a local organisation. A few examples of roles are:

- A physician employed a by an insurance organisation.
- A physician that is head of a hospital department.
- A self employed nurse.
- A self employed physiotherapist or dentist.
- …

The granting and revoking of role certificates is a function that should be done locally by the organisation that assigns the role to the health professional. A role is expressed in a certificate. The main elements of a role certificate are:

- The identity of the certificate holder.
- The identity of the issuing certificate organisation.
- The signature and identity of the person within the local organisation that is responsible for the issuing of role certificates.
- Proof or reference (i.e. a certificate or certificate identifier) that the issuing organisation is allowed to issue role certificates.
- The role that is attributed to the certificate holder.
- The validity period of the role certificate.

All these elements need not be very long, so that the size of a certificate can be kept rather small. The role can be given by a role identifier that references a policy document that describes in extenso the details and constraints of the role.
Authorisation to local sources is a local matter. The signature and authentication functions of the health professional card support the authorisation process but the card does not contain authorisation information as such.

4.2 Decentralised key issuing model

If health professionals do not already hold a smartcard (issued by a central organisation) that contain signature and authentication keys, the decentralised issuing organisations have to deliver smartcards that are personalised and that contain the necessary keys. On top of that, role certificates have to be issued and maintained as well, depending on the variant of the decentralised model.
Two variants on this model are possible:

- Role, identity and signature key are combined
- Identity and keys are kept separate from the roles.

In both variants, a complete PKI infrastructure has to be set up by the decentralised organisations for the issuing of smartcards to health professionals. Moreover, the identity related information only has significance within the issuing domain, in casu health care. In the case where ID and role are combined a health professional will produce a different signature for each role he is in. This can make the process of verification and role management more complicated.

5 Some practical scenarios

This section contains some scenarios that relate to various aspects of the deployment and use of digital signatures. Some parts in the scenarios are specific to digital signatures; others are generic and apply to any type of electronic signature. The scenarios are based upon the assumption that a public key infrastructure, based on smartcards, is used. This document concentrates on signatures but large parts are equally valid for keys used to encrypt for confidentiality, authentication, …
The scenarios will highlight differences between the various model described above.

5.1 Obtaining a signature key pair

Before a person can digitally sign a document, he has to have a smartcard containing at least a private signature key. A smartcard is a tamperproof environment in which the cryptographic key used to sign is securely stored and where the actual encrypting part of the signature is executed. The public key part of the signature key pair has to be associated to the identity of the key holder, which is done in the public key certificate, and made publicly available.
If a health professional already has an ID card that enables him to sign electronically, he already has a signature key pair and this step can be skipped. This reduces the complexity of signature support significantly since all effort can be concentrated on the roles of the cardholder. This scenario further deals with the model where the health professional already has a centrally issued electronic ID card.

If not, the health professional has to go to the registration authority (or to the certification authority if both functions are combined into one organisation) and present his credentials. What credentials are needed, including form and content requirements need to be stated in a written policy document. The policy statement can e.g. contain the following requirements:

- Handover of an authentic birth certificate.
- Presentation of the paper based ID card and handover of a copy of it.
- A handwritten signature of the health professional on a document stating that he has read and understands the policy requirements for the use and handling of the smartcard that he is applying for.

5.2 Obtaining an attribute certificate

Once a user has an electronic ID and a smartcard that can be used for signatures, certificate issuing and application becomes quite simple. Some examples:

1. A health professional has finished his studies and has obtained the diploma of physician. He submits this paper document to the order of physicians. The order of physicians issues a registration number. Supposing that the applicant already has a nationally registered electronic ID and that the order of physicians has access tot a trusted source for verification of the electronic ID and the ID on the diploma, it can automatically issue an electronic certificate stating the qualifications of the health professional. The certificate is sent by e-mail to the health professional and is optionally published on a directory server.
2. A health professional is employed by a hospital. The following procedure is followed.

- He inserts his electronic ID card into a card reader of the personnel department. He signs (i.e. he pushes an 'accept' button and inputs his pincode) an electronic statement saying that he is about to receive a role certificate of the hospital and that he has read and understood the role certificate policy stating the use and constraints of the role certificate.
- He withdraws his ID smartcard.
- The hospital checks his certificate that states his type of health professional (e.g. a medical doctor, nurse,…)
- He receives a diskette with a copy of his certificate(s), a paper copy of the policy and perhaps some practical information like his e-mail address at the organisation.

5.3 Signing a document

Once a health professional has received his smartcard he can put his electronic signature on an electronic document or on a message, provided of course that the applications he is using have signing capabilities.
Signing a document consists of the following steps:

- Either the user clicks a 'sign' button or the application prompts the user automatically to sign.
- The user selects the role(s) he wants to include in the signature.
- He verifies what he is about to sign.
- He inserts his smartcard (if not already done) and inputs his pin code via the keyboard of the pc or via a separate keypad.

5.4 Verifying a signature on a document

When an application is signature enabled, it will automatically check all incoming documents for signatures.

- When a document has not been signed or when the signature is invalid, it will issue a warning and state the reason. Warnings are also displayed if a role or key certificate is not valid anymore.
- When the signature is valid it will display information about the author.

The most critical element in the verification process of signatures and roles are the key and role certificates. The policy of the organisation should specify the acceptable expiry limits that certificates and revocation lists can be held locally without refreshing (i.e. downloading up-to-date information from a directory service).

5.5 Revoking a key pair

- A health professional cannot find his smartcard anymore. He informs the personnel department and requests a new card. The policy requires that the card and certificates are revoked from that moment on and new ones are issued.
- A health professional quits the hospital he is working for. From that moment on, the personnel department revokes his locally issued role certificates.

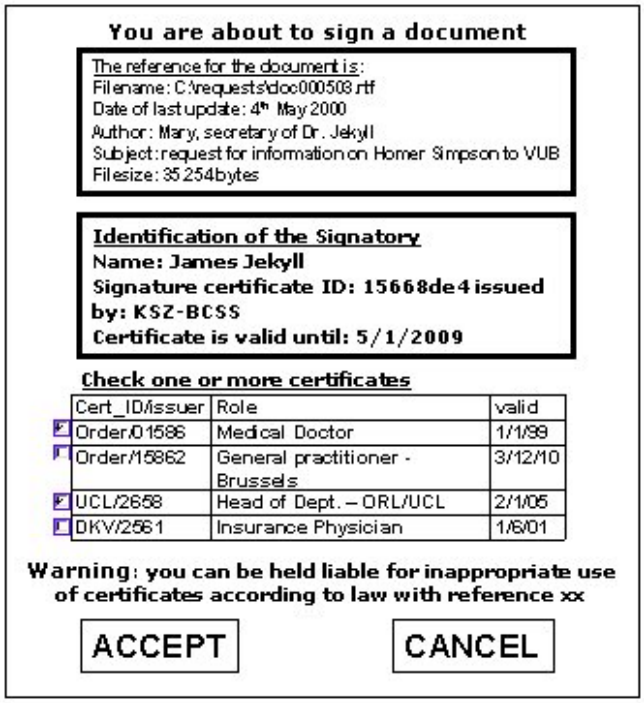
It is only at verification time that can be discovered if a certificate was revoked at the time the signature was made. Therefore a regular updating of certification revocation lists should be defined in the verification policy.

There are schemes that allow a stepwise revocation in which the certificate can be put on-hold in an early step and where the revocation can still be undone as long the final revocation has not been done.
When key and attribute certificates are distinct, it is possible to revoke specific attribute certificates without having to revoke the other certificates as well. This can be particularly interesting when certificates contain relatively volatile information, such as employment data.

5.6 User interaction with security services

Security services should be as transparent as possible to the users. The following figure shows how an interface for signing an electronic document can look like.

The minimum requirements that a signing interface has to offer are the following:



- The user has to know what he is signing. That can be done by showing the document or message he is signing, but also by reference to the document itself (name, size, date of latest edit, …).
- The user needs to know 'who' he is. That is expressed by his identification and by information on the certificates his identity are based upon.
- The user has to be given the choice in which capacities he is going to sign. When a conflict of roles can arise, he must be aware of the ethical rules and legislation that exist.

Generating a signature should always be a conscious act. Misuse of mandates cannot be prevented technically (for this would limit correct use in other situations) but the user has to be aware that he is liable for misuse and that all misuse is considered deliberate.

6 Conclusions

Health care is only one domain where electronic signatures are introduced. Health care has its specific requirements and procedures for the signing of documents. Electronic signatures will only provide sufficient proof when they are supported by an adequate certification infrastructure.[4] It is also necessary to develop message and document formats that allow the use of efficient and open structures for including attribute certificates and a flexible attachment of signatures. It is recommended that the electronic world follows the traditional paper based world in its method of signing and communication roles. The person signing should be able to choose in which role(s) he is signing. Legislation is needed as an incentive to prevent misuse of role certificates by health professionals when signing a document or message.

A costly infrastructure can be saved and flexibility gained if health care specific applications can build on top of already existing PKI services, e.g. nationally provided. In that case, health care could concentrate on the specifics of its domain, dealing with roles and authentication of health professionals. Rules for authorisation should be locally administered by the healthcare enterprise the health professional is associated with. Local authorisation and certification applications however could make use of the electronic ID smartcard of the health professional.

Electronic signatures are only trustworthy when the certificates needed to verify the signatures and the roles have been provided through services delivered by Trusted Third Parties.

7 References

[1] Directive 1999/93/EC of the European Parliament and of the council of 13 December 1999 on a Community framework for electronic signatures, Official Journal of the European Communities, 19.1.2000.

[2] Draft TR 101 xxx v.0.4.2(1998-11), Telecommunications Security ; Electronic signatures standardization report, ETSI, 1998

[3] Social Security Royal Decree of 16.10.98

[4] Spécifications techniques de l'infrastructure de sécurité AGORA, Partie III, 30/6/98