

# E-Government and Digital Preservation

*Jos Dumortier*

*K.U.Leuven – ICRI*

---

## I INTRODUCTION

In today's world, electronic document management and electronic information transmission constitute already an extensive part of commercial and administrative activities. It is expected that the use of digital data will be generalized in the coming years and that it will gradually replace traditional paper-based methods of information processing. Paper documents will of course not completely disappear but the paper form will no longer be the core of the document management system. Its role will be reduced to one of the output formats of a system that is essentially based on the processing of digital information. This tendency is already apparent in advanced administrative environments, for instance in the banking or in the insurance sector, and it will sooner or later without any doubt also invade government administrations, parliaments and courts of justice.

Despite this undeniable trend, the use of electronic information still faces some skepticism and reluctance. When dealing with crucial information such as important contracts or decisive administrative documents people still often fall back on the use of paper. One of the reasons – though certainly not the only one – is the lack of security about the possibilities to store electronic documents on a longer term. Computer hardware and software are undergoing constant and rapid changes and nobody can foresee how electronic information will be processed twenty or thirty years from now. How can we guarantee that the electronic documents that are being stored today will still be readable by the computers and programs that will be used in the future? How can we protect electronic information with its volatile and easily alterable nature, from being modified or deleted?

Professional archivists are still discussing about possible solutions for this problem and in these discussions two basic strategies have been proposed. Following a first strategy, the archivist should try to guarantee the usability of electronic data over long periods of time by storing the data in their original format and by making sure that the necessary hardware and software environment enabling the use of these data can always be made available afterwards. This approach is generally called the "emulation" strategy.<sup>1</sup> If, for instance, a particular document is produced and archived in a current version of a specific word processor on a currently used operating system, the archivist will make sure that this document will remain readable over time by "emulating" the word processing environment in which the document was originally stored. In order to make this possible, the archivist must of course be able to keep a complete set of all the hardware and software needed to use all electronic data

---

<sup>1</sup> Jeff Rothenburg, *An Experiment in Using Emulation to Preserve Digital Publications*, Amsterdam, National Library of the Netherlands, 2000, 74 pages, <http://www.kb.nl/coop/nedlib/results/NEDLIBemulation.pdf>

formats stored in the archive.<sup>2</sup> Experiments in archiving institutions have demonstrated that this strategy not only requires important efforts and investments but also bears considerable risks.

Therefore other archivists propose an alternative solution putting a stronger emphasis on “migration”. Following this second strategy the archivist should not primarily try to keep the electronic document in its original format. The role of the archivist should, on the contrary, be to restore the information contained in the documents that have been archived. To enable this, the archivist may need to convert the document into another format, for instance in order to keep the document readable on a new hardware and software platform. At the end of the archival chain, the user will not necessarily find the original document as such. Possibly the document will have been adapted to keep it usable but the archivist will guarantee that the information contained in the document is correctly restored.

There has been much debate about both of these strategies and exponents of one or the other have argued their relative merits.<sup>3</sup> Recent research, for example in the context of the “Cedars”<sup>4</sup> and the “CAMiLEON”<sup>5</sup> projects suggest a combination of these strategies and one which has the potential to overcome the major disadvantages associated with either. One approach is to preserve both the original bitstream as well as detailed metadata enabling it to be interpreted in the future. The combination will hopefully sidestep the major technical difficulties commonly associated with adopting either migration (loss of information through successive migrations) or emulation (risking that the attempt to recreate a particular environment will be successful). Little by little the discussion about long-term preservation of digital information is leaving the “emulation versus migration debate” and proposes more sophisticated and open-ended ways to solve the problem.<sup>6</sup>

---

<sup>2</sup> Stewart Granger, Emulation as a Digital Preservation Strategy, D-Lib Magazine 2000, <http://www.dlib.org/dlib/october00/granger/10granger.html>

<sup>3</sup> For example: Harrison Eiteljorg, Preservation for the Future? – with emulation or migration?, CSA Newsletter, 1999, Vol. XII, n° 1, <http://www.csanet.org/newsletter/spring99/nls9906.html> ;

<sup>4</sup> “Cedars” (CURL Exemplars in Digital Archives) is a digital preservation project in the context of eLib phase 3. The Cedars project began in April 1998 and was initially funded for three years. It began as a collaboration between three CURL institutions, the universities of Leeds, Cambridge and Oxford. For more information about “Cedars”: <http://www.leeds.ac.uk/cedars/index.html>

<sup>5</sup> CAMiLEON stands for Creative Archiving at Michigan & Leeds: Emulating the Old on the New. The aim of the project is developing and evaluating a range of technical strategies for the long term preservation of digital materials. The project is a joint undertaking between the Universities of Michigan (USA) and Leeds (UK) and is funded by JISC and NSF. For more information we refer to the project’s website: <http://www.si.umich.edu/CAMILEON/>

<sup>6</sup> See for example the findings of the InterPARES project, The Long-Term Preservation of Authentic Electronic Records, <http://www.interpares.org/book/index.cfm> ; also: Kenneth Thibodeau, Overview of Technological Approaches to Digital Preservation and Challenges in Coming Years, in: The State of Digital Preservation: An International Perspective, Conference Proceedings, <http://www.clir.org/pubs/reports/pub107/thibodeau.html>

## I DIGITAL SIGNATURES

One technology that is often referred to in this context is the “digital signature”. This cryptography-based technique allows authenticating electronic information in such a way that the originator of the information, as well as the integrity of the information, can be verified.<sup>7</sup>

The basic characteristic of digital signatures is that electronic information can be “signed” by using a secret cryptography key. This key must be kept private at all times by the signatory. The signature can only be verified with the associated public key of the author.

The idea behind this authentication is the confirmation of identity by proving the possession of a secret key. The author encrypts the information or a part of it with his secret key. The recipient of the information can check the identity of the author by decrypting the information with the public key of the presumed author. If the decryption is not successful the recipient will not validate the message. This process of authentication relies on the public keys of the users that are accessible to all the communication partners and on a trusted relationship between the identity of the users and their public key.

The authentication procedure is based on the presumption that the public key really belongs to the signer. This presumption is, however, not self-evident. The risk exists that somebody creates a key-pair, places the public key in a public directory under somebody else’s name and thus signs electronic messages in the name of somebody else. Furthermore, a public and private key pair has no inherent association with any identity because it is simply a pair of numbers. Therefore, the assurance should exist that the public key really belongs to the claimed identity.

The answer is to rely on third parties to certify public keys. A third party will guarantee the relationship between the identity and the public key. This association is achieved in a digital certificate that binds the public key to an identity. The third parties are known as Certification Authorities and must be accepted by all users as impartial and trustworthy. In addition, the process of key certification must be foolproof and should be afforded the highest level of security. A Certification Authority will, by issuing a digital certificate, certify the identity of the user and guarantee that the public key really belongs to the claimed user.

Digital signature technology can be used wherever there is a need to keep track of the origin and the integrity of computer data. Therefore it has been adopted as a privileged electronic substitute for the handwritten signature, for instance in the European Directive 1999/93/EC<sup>8</sup> dealing with electronic signatures. According to this Directive, where the use of electronic documents is legally permitted, so-called “qualified electronic signatures” must receive a status that is equivalent to the legal status that handwritten signatures normally have in relation to paper documents.

---

<sup>7</sup> For a more detailed but accessible explanation on digital signature technology and public key cryptography, we refer to <http://developer.netscape.com/docs/manuals/security/pkin/contents.htm>

<sup>8</sup> Directive 1999/93/EC of the European Parliament and of the Council of 13 December 1999 on a Community framework for electronic signatures, OJ, 19 January 2000, L13/12; For more details: Jos Dumortier, Directive 1999/93/EC on a Community framework for electronic signatures, published in: Lodder, A.R., Kaspersen, H.W.K.,: eDirectives: Guide to European Union Law on E-Commerce. Commentary on the Directives on Distance Selling, Electronic Signatures, Electronic Commerce, Copyright in the Information Society, and Data protection., Kluwer Law International, p.. 33-65, <http://www.icri.be/publications>

The technique of the digital signature plays an important role in this new legal framework. From the current state of the law in Europe results that only digital signature technology can bring forth so-called “qualified” electronic signatures. As a result of this new legal framework, archivists are increasingly challenged to deal with digital signatures as an organic part of electronic documents.

## I USING DIGITAL SIGNATURES FOR ARCHIVING

Although digital signatures are known best as a substitute for handwritten signatures with legal value (= electronic signatures), the technique of the digital signature has many other applications. It can be used in all cases where the origin and the integrity of electronic data have to be guaranteed.<sup>9</sup> These qualities are very important for documents that are stored in archives. A digital signature added to the (signed) record by the archivist, allows the verifier of the signature to check the identity and the authority of the archivist. That is how the authenticity of a record “as a record” can be checked in a network environment, the future work area of archivists. The presence of the digital signature of the archivist in the metadata of a record indicates that this record has the status of an archived record. The use of the digital signature technique also creates the opportunity for checking the integrity of electronic records. When used in this manner, the digital signature functions as a “seal”.<sup>10</sup> By creating and archiving an encrypted, and thus inaccessible hash code, it can be noticed at all times when the plain text has been tampered with.<sup>11</sup>

Nevertheless there exists a lot of resistance in the archival community against the preservation of digital signatures. This is well-illustrated by the report of the InterPARES Authenticity Task Force, entrusted with the task of identifying “conceptual requirements for assessing and maintaining the authenticity of electronic records”.<sup>12</sup> The Task Force adopted an unequivocal position with regard to the role of digital signature technologies and PKI as a means of ensuring the authenticity of records:

“Digital signature and public key infrastructure (PKI) are examples of technologies that have been developed and implemented as a means of authentication for electronic records that are transmitted

---

<sup>9</sup> The possible use of digital signatures for the preservation and authentication of records through time has been analyzed in the framework of the DAVID-project (which stands for Digital Archiving in Flemish Administrations and Institutions, <http://www.antwerpen.be/david> ). See also: Sofie Van Den Eynde., *The OAI Reference Model as starting point in search of the role of Public Key Infrastructure for electronic archives*, Leuven, Interdisciplinary Centre for Law and Information Technology, August 2001, 63 p. (in Dutch only).

<sup>10</sup> As opposed to the digital signature used as an electronic signature with legal value in the sense of the European e-Signature Directive.

<sup>11</sup> The possibilities of digital signature technology must not be overestimated though. To guarantee integrity, we probably must combine this technology with carriers of the ‘Write Once Read Many’ type.

<sup>12</sup> J.P Blanchette, ‘Dematerializing’ Written Proof: French Evidence Law, Cryptography and the Global Politics of Authenticity, Doctoral Dissertation submitted to the Department of Science and Technology of the Rensselaer Polytechnic Institute, 2001, p.308, writes: “The fundamental premise of the InterPARES project is that authenticity is not primarily a function of technology, but rather, of institutions. Archivists have historically been entrusted with the task of providing this function, within either private or public institutions, and they remain the most appropriate, professionally organized, socially recognized, historically legitimate profession to accomplish similar functions in the electronic environment.”

across space. Although record-keepers and information technology personnel place their trust in authentication technologies to ensure the authenticity of records, these technologies were never intended to be, and are not currently viable as a means of ensuring the authenticity of electronic records over time.”<sup>13</sup>

Skepticism appeared as soon as it became clear that, when using digital signatures, control of the integrity is only possible if the electronic data remain completely unchanged at the bit-level. This raises a problem when archivists want to migrate electronic data to new formats or software platforms in order to keep them accessible and legible. Some people have immediately concluded that digital signatures are therefore not useful and hence not relevant for archival purposes.

But is it not possible to avoid the need for migration by storing the digital data in a standardized open format that remains stable over a very long period of time? This is the reason why people refer in this discussion to the development of hardware- and software-independent document formats, such as XML. XML (eXtensible Markup Language) is nowadays the most popular standard for structured information exchange. However, the XML 1.0 Recommendation defines multiple syntactic methods for expressing the same information. That is why XML applications tend to represent the same content in different ways. Therefore, XML “canonicalization” was designed.<sup>14</sup> The canonicalization method uses an algorithm to generate the canonical form of a given XML document. The canonical form is the common denominator so to speak for all possible syntactic representations of a given content. A digital signature over the canonical form of an XML document allows the hash calculations to be oblivious to changes in the original document’s physical representation.

It would be naïve however to believe that XML will solve the problem of electronic documents and digital signatures becoming obsolete. To begin with, the canonicalization method developed for XML 1.0 may not be applicable to future versions of XML without some modifications. The transfer of an XML document to this newer version will invalidate the signature, since the canonical form cannot be carried indefinitely into the future. At the moment, software companies are implementing XML in their products. The multiple use of XML and its vendor independent character give XML the status of de facto standard. But it is not very likely that XML will be maintained as a common format forever. IT will keep evolving and it is unthinkable that there will never be a better alternative for XML. A canonical form that takes all current and future formats into account is unfortunately still IT science fiction. Many archivists therefore believe that there will always remain a need for migration.

## ARCHIVING ELECTRONIC SIGNATURES

At this point of the discussion we come logically to a following question and that is whether or not it is possible to avoid long-term preservation of digitally signed data? Is migration, in other words, acceptable for all kinds of documents or will there always be a need to keep the original document intact?

---

<sup>13</sup> See the draft final report of the InterPARES Authenticity Task Force, [http://www.interpares.org/documents/atf\\_draft\\_final\\_report.pdf](http://www.interpares.org/documents/atf_draft_final_report.pdf), p. 8

<sup>14</sup> Canonical XML, Version 1.0, W3C Recommendation, 15 March 2001 <http://www.w3c.org/TR/2001/REC-xml-c14n-20010315>

In a traditional paper-based environment, some documents contain handwritten signatures and as we have seen earlier, digital signature technology is being used more and more as an electronic substitute for such handwritten signatures. Although it can be expected that electronic signatures will be needed less frequently than handwritten signatures, some important contracts or administrative documents will require an electronic signature in the future.

The question arises how to deal with these electronic signatures if the related documents have to be migrated for preservation purposes.

American government administrations have suggested the following solution:<sup>15</sup>

“To ensure continuity of record integrity, you should perform the following sequence of procedures:

- Just prior to performing the electronic record migration a trusted third party from outside of the organization that has some responsibility for the electronic record verifies the digital signature using the old system methods;
- Under supervision of the above trusted third party, the signed electronic record is migrated to the new system; and,
- The above trusted third party then applies a new digital signature (using technologies appropriate to the new system) to the migrated electronic record. The same third party also prepares and applies a digital signature to a new separate electronic record (or to an addition to the migrated electronic record) that explains the migration. In this situation, although you would no longer be able to verify the old digital signature directly, you should nonetheless be able to demonstrate continuity of record integrity by verifying the newly digitally signed migrated electronic record and explanatory statement.”

Is the procedure proposed by the American government administrations acceptable for all documents with an electronic signature? Or is it in some circumstances necessary to keep the original document with the electronic signature intact?

From a legal point of view, in order for signed documents to keep their value over time, it could often be important that the original electronic signature remains present. Signatures could be needed for non-repudiation purposes in an evidential context, for example. Many European countries require for proof that non-commercial transactions are embodied in a signed document.<sup>16</sup> Recent developments in the context of e-government have also made clear that signed electronic communication with the government must be archived. The government that picks the lowest bidding firm in the context of a public contract conducted by electronic means, will want to be able to prove before court that this firm is bound by its price offer. In a traditional paper-based context the original document with the handwritten signature is often needed to avoid all possible disputes in these circumstances. Will it be accepted to replace the original signature in an electronic environment by presenting a declaration of a trusted third party?

---

<sup>15</sup> US FDA et al: Guidance for Industry 21 CFR Part 11; Electronic Records; Electronic Signatures: Maintenance of Electronic Records (July 2002), <http://www.fda.gov/OHRMS/DOCKETS/98fr/00d-1539-gdl0001.pdf>

<sup>16</sup> KÖTZ, H., European Contract Law: Formation, Validity and Content of Contracts, Contract and Third Parties, Oxford, Clarendon, 1998, 78.

In the paper world the content of the document and the signature are one indivisible artifact. A traditional signature has all the characteristics of a classical one-way function: it is easy to process in one direction but very difficult to reverse the process, i.e. the signature is easy to affix but difficult to remove. This is not the case with electronic signatures: an electronically signed document is not different from an electronic document that has not been signed except that it has appended to it another series of bits that can be used to identify the signatory and verify the integrity of the document. Thus, an electronic signature can very easily be stripped from a document for fraudulent purposes without leaving a trace.

Although they have the same functions from a legal viewpoint, traditional signatures and electronic signatures are two very different concepts that need to be treated differently. Never before in the history of written communication a signatory has had to worry about how the signature will be linked to the content of the document that he is signing. When using electronic signatures, this becomes now a very relevant issue.

### **LONG TERM VALIDATION OF ELECTRONIC SIGNATURES**

In the context of the development of the European regulatory framework for electronic signatures the current opinion is that there is a need to keep electronic documents in their original form. This is particularly clear in the standardization initiatives concerning the long-term validation of electronic signatures.

The European Commission took the view that the requirements identified by the e-Signature Directive needed to be supported by detailed standards and open specifications so that products and services supporting electronic signatures can be known to provide legally valid signatures. A mandate was issued to European standardization bodies, CEN/ISSS and ETSI, to analyse the future needs for standardization activities. Under the auspices of the European ICT Standardization Board the European Electronic Signature Standardization Initiative (EESSI) was launched. The first result of this initiative was an expert report about future standardization requirements. This report affirms that trusted archival services could play an important role in supporting electronic signatures that may need to be used in evidence long after they were created and identifies it as a topic requiring further study since no standards exist yet for the use of such services in support of electronic signatures.<sup>17</sup>

In the mean time, ETSI has published a standard “Electronic Signature Formats” defining all the elements necessary to prove the validity of a signature long after the normal lifetime of the critical elements of an electronic signature.<sup>18</sup> This so-called validation chain has to be archived.

Thus, it is not enough that just the electronic signature and the content of the document are present in the archives when a signed document is needed years later. In order to perform validation, the certificate used by the signatory must be obtained, and its validity at the time of signature creation must be proofed. It is possible that the certificate was valid at the time of signature creation, but had

---

<sup>17</sup> NILSSON, H., VAN EECKE, P., MEDINA, M., PINKAS, D. and POPE, N., *European Electronic Signature Standardization Initiative*, Final Report of the EESSI Expert Team, 20 July 2000, 69, available at: <http://www.ict.etsi.fr/eessi/Documents/Final-Report.pdf>

<sup>18</sup> *Electronic Signature Formats*, ETSI TS 101 733 v.1.3.1 (2002-02).  
[http://webapp.etsi.org/exchangefolder/es\\_201733v010103p.pdf](http://webapp.etsi.org/exchangefolder/es_201733v010103p.pdf)

expired or had been revoked or suspended some time later. By consequence, the certificate status information must be archived as well.<sup>19</sup> Signature validation must be performed immediately after, or at least as soon as possible after signature creation time, and not only at archival time, in order to obtain certificate status information that was issued by the CA as closely as possible to the moment of signature creation.

Only the moment of signature creation has an archival value. A signature that has been found to be valid at signature creation time shall continue to be so for the same document months or years later. Evidence must be provided that the document was signed before the certificate became invalid. Thus, the time of signature creation must also be determined and archived.

A time stamp can provide for such evidence. A time stamp is a set of computer data, consisting of the hash code of the digital signature and the time of stamping, signed by a trusted third party. It proves that the digital signature was formed before the certificate became invalid. Anyone who wants to make sure that he can rely on a signed electronic document for proof years later, must obtain a time stamp before the certificate becomes invalid. The sooner after the creation of the signature the time stamp is obtained the better it is for legal certainty.

The solution put forward in the EESSI standardization project is that the content of the document and the digital signature should be concatenated and the hash-code of the concatenation should be lodged with an independent entity that would time stamp the hash-code.<sup>20</sup> The hash-code establishes the bond between signature and content. The time stamp must be included in the metadata of the document.

The only possibility in this view is the archival of the original binary representation of the document or in other words a preservation strategy based on “emulation”.<sup>21</sup> A trusted third party must guarantee that it will still be possible to validate an archived document years after the initial archival date, even if the applications that have been used at signature creation time are no longer in use. In other words, the third party should maintain a set of applications (viewers as well as signature validation applications) together with the corresponding platforms (hardware, operating systems) or at least an emulator of such applications and/or environment in order to guarantee that the signature of the document can still be validated years later.

## **TRUSTED ARCHIVAL SERVICES**

It is striking that the intervention of trusted third parties and digital signature technology is being proposed in both of the proposed solutions. In the context of preservation based on migration the trusted third party is needed to keep track of the migration process and to make sure that the resulting document at the end of the migration chain keeps being trusted. If one opts for a solution based on emulation, the trusted third party is even more essential. The costs and expertise required for this

---

<sup>19</sup> It is the responsibility of each Certification Authority (CA) to make available in repositories on the Internet all the information needed to validate any signature that was created by means of a certificate issued by that CA. This includes making public at a regular basis information about the time a certificate expired, or was revoked or suspended.

<sup>20</sup> McCULLAGH, A. et al., ‘Signature Stripping: a digital dilemma’, *Journal of Information, Law and Technology*, 2001/1, <http://elj.warwick.ac.uk/jilt/01-1/mccullagh.html>

<sup>21</sup> European Commission, August 2000, 37.

solution, requires that the task of archiving digital data will be appointed to an independent third party. Although contractual freedom also applies for the manner in which contracts are archived, private persons will not always be able to securely keep signed documents in their own possession.

If our conclusion is that, whatever the ultimate solution for digital preservation will be, specialized trusted third parties – commonly called “trusted archival service providers” or “TAS” – will play a central role, a further question is in which framework these service providers will operate.

A TAS should be able to present and validate digital data years after their initial date of archival. As it was already indicated in the final report of the EESSI expert team, standards must be developed for the use of trusted archival services also in support of electronic signatures. A clear Community framework regarding the conditions applying to TASs will strengthen confidence in and general acceptance of this kind of services.

A legal framework could, for instance, determine:

- ?? that Member States must ensure that by accepting data for archival, a TAS is liable for damage caused to an entity or a legal or natural person who relies on its services. Breach of this “obligation de résultat” should mean that liability is indisputable. A TAS should not be admitted to proof that it has not acted negligently since the loss of evidence is irreversible. Therefore, a TAS must obtain appropriate insurance to bear the risk of liability for damages.
- ?? that the archives of a TAS can never be destroyed. For the case where a TAS ceases its activities, procedures must be drafted to steer the transfer of the archives to another TAS. In order to prohibit that a TAS goes into failure, a very strict investigation regarding the financial situation and prospects should be carried out prior to the start of his activities.
- ?? that a TAS must employ personnel who possess the expert knowledge, experience and qualifications necessary for the archival services provided.
- ?? that a TAS must use trustworthy systems to store the documents, the signatures and the validation chains so that only authorized persons can make entries and changes.
- ?? that a TAS, before entering in a contractual relation with a person who wants to archive a document, must inform that person of the precise terms and conditions of the storage, such as the term of storage and the accepted file formats. Such information, which may be transmitted electronically, must be in writing and in understandable language. Relevant parts of this information must also be made available on request to third parties relying on the archived document for proof.

The European electronic signatures directive contains a very wide definition of “certification service providers”. Trusted archival services are under the scope of this definition. It is somewhat strange that the European legislator, despite the very wide scope of services included in the definition, has exclusively focused on certificate issuers. It seems unavoidable that this will have to be corrected in the future to be better adapted to the challenges of the upcoming information society.