The Hidden Truth About Electronic Signatures

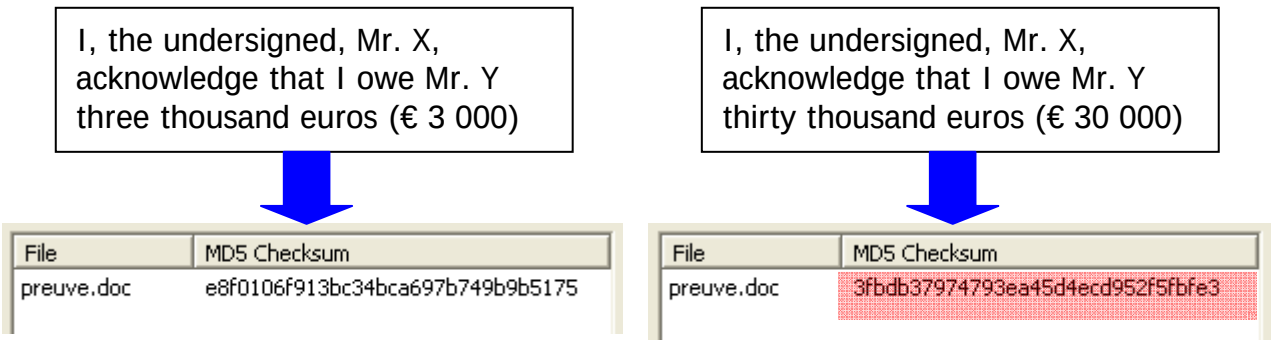
by Lucien Pauliac
President of the Association "Preuve & Archivage"

Translation by Nicholas Allen
Lawyer translator
allen@club-internet.fr

The electronic signature[1] was granted legal recognition in France for the first time on 13 March 2000, in Article 1316 (4) of the Civil Code. Subsequent descriptions of this signature seem to afford it considerable legal significance. An electronically-signed document is deemed to be:

- ascribable to the signatory;
- faithful;
- tamper-proof;
- irrevocable.

Ten years on, we have the distinct impression that there is little understanding of the mechanisms behind these electronic signatures. Hence, we felt it would be useful to examine how they work and assess the technical and legal impact of their current usage. We have selected the following example:

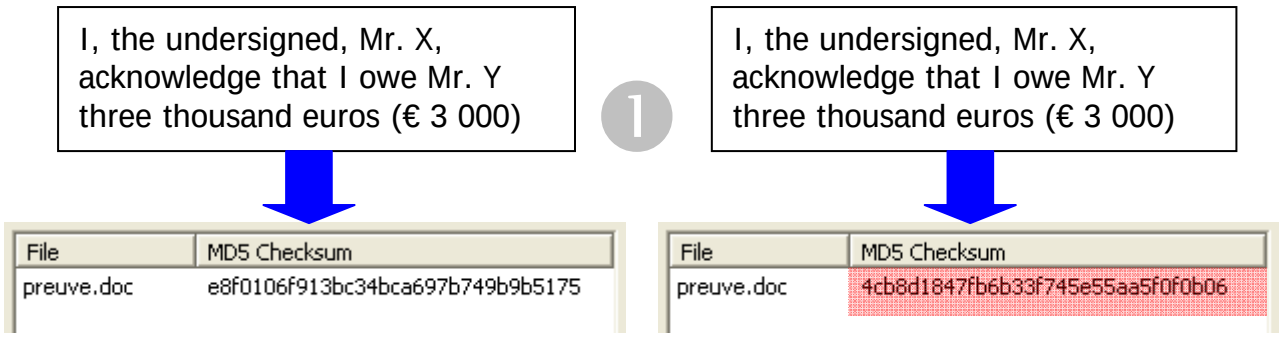


This example shows an effective use of an electronic signature. It illustrates a straightforward case in which Mr. Y, holder of the electronic debt acknowledgement issued by Mr. X, has altered the amount of the debt to his advantage. The attempted fraud is revealed by the blatant

[1] Briefly, electronic signatures rely on cryptography techniques, which calculate a "digital fingerprint" (sometimes called a "hash" or "checksum") that is unique to the document in question. This hash is a digital summary of a file, generated by applying a "hashing" algorithm (the most common being MD5 and SHA1) which produces an unpredictable character string. The hash calculation has to be "one way", i.e. it should not be possible to recreate the document from its hash value. Furthermore, a single document should only be able to produce a single hash, and it should not be possible for two different documents to produce the same hash (simply moving a comma should result in a radically different hash value). Therefore, we consider that a document has been altered if the hash code it produces does not match its original hash code. In the full process, the hash itself is encrypted using an encryption key that is unique to one person, then deciphered using the public portion of the key intended to identify the signatory.

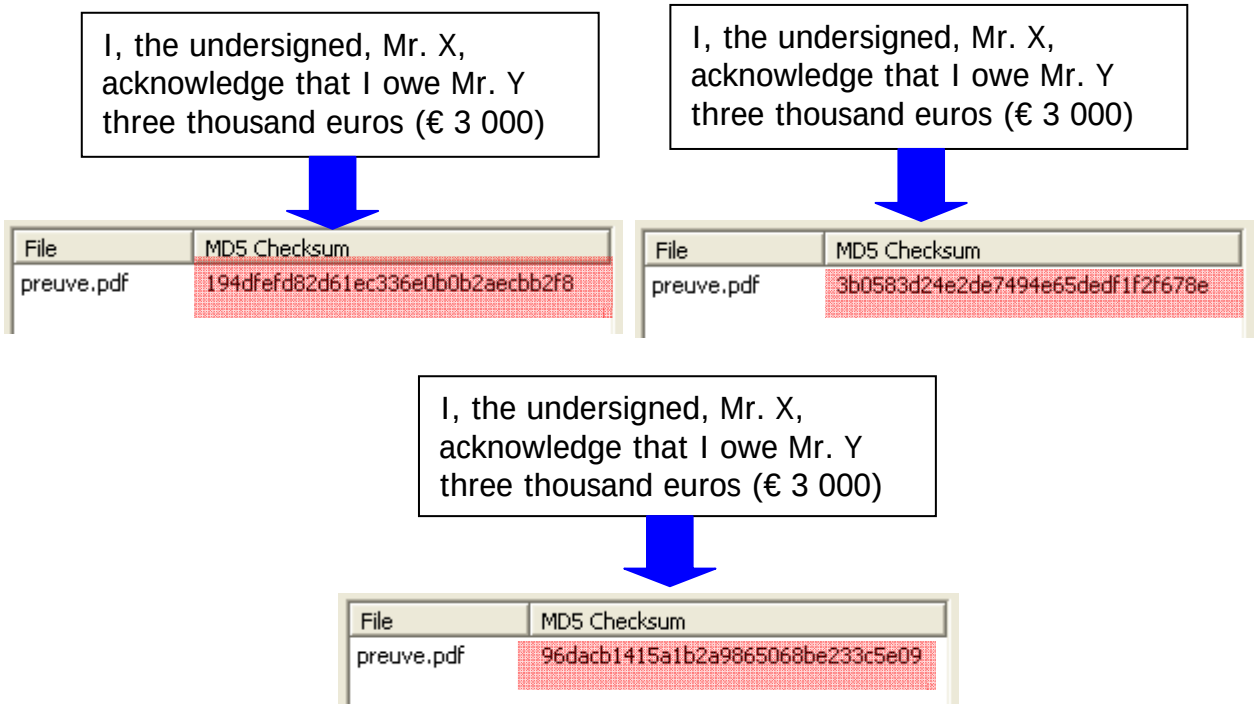
difference in the hash code (MD5 Checksum[2]). Fortunately, the dishonest Mr. Y will not be allowed to rely on this document in any lawsuit, as it has been proven to be a forgery by its electronic signature. His fraudulent intentions have been thwarted and all is well.

However, let us suppose that Mr. Y is honest and has not changed the debt acknowledgement. This raises the following issues:



In this case, digital verification tells us that the message on the right has been altered because its digital fingerprint does not match the one on the left. The problem lies in the fact that the original message is still intact, but is identified as a "forgery" by the electronic signature.

This is problematic because a document deemed to be a forgery is obviously legally unenforceable, despite the fact that, in our example, it has not been tampered with in any way. This raises a legal problem that is just as serious as if the electronic signature had been "broken". And this is far from being an isolated case, as we will see below:



As we can see, three identical PDF versions of the same document also produce different digital fingerprints. All of them will therefore be suspected of being forgeries.

[2] Demonstrated using "mst MD5" freeware

Why is this so? Has there been some form of malfunction, error or bug?

Explanation

These differences are, in fact, perfectly normal from a technical standpoint.

In example ❶, the "preuve.doc" file on the left was created using Microsoft Word 97, while the "preuve.doc" file on the right was converted using MS Word 2003. As regards the three PDF documents:

- the first is a "standard" PDF created using Acrobat 8;
- the second is also a "standard" PDF, created using PDFCreator;
- the third is a PDF/A document, produced using Acrobat 8.

As the example shows, each document has a different hash code – and the reasons for this difference are unavoidable.

The hashing algorithm ignores the material representation or semantic content of the document – it is applied to the document's digital data. Whenever a file is converted into another format, its data structure changes. For example, converting a Word 97 file into a Word 2003 file does not modify the document itself, but alters its file structure[3]. As the data have changed, so will their digital fingerprint. This outcome is inevitable.

In other words, any digital format conversion (or "migration" as we now call it), however insignificant (e.g. converting a standard PDF into PDF/A format), changes the electronic signature, discrediting the document entirely and producing what appears to be a forgery.

The problem is that, if we want to store documents in digital format, there is no way of avoiding these migrations. This is clearly stated in French Standard AFNOR NF Z 42-013 (2009), which explains that, for electronic archiving to survive, files have to be migrated from time to time for a variety of reasons[4].

This is hardly news

The worst of it is that this is hardly a new discovery. This structural flaw in electronic signatures was revealed publicly on 1 December 2005 in a recommendation by the French Internet Rights Forum which included the following statements:

> *Accordingly, the same document saved in two different formats, e.g. ".doc" and ".pdf", will have two completely different cryptographic signatures, even if these two digital documents contain exactly the same information and produce totally identical paper documents.*
> *…/…*
> *An [electronic] signature loses all validity as soon as we change the document encoding format (or the version of a given format). To safeguard a cryptographic signature, the encoding format must not be migrated under any circumstances, despite the fact that migrations are generally necessary to ensure the long-term readability of a document.*

[3] See "The Issue of Integrity": http://www.megapreuve.org/cariboost_files/Issue_of_integrity.pdf

[4] In passing, we should point out that the standard advises using a hash function and electronic signature to protect data but also recommends regular migrations, paying no heed to the incompatibility of these two processes and the legal implications of such incompatibility.

In other words, this failing has already been documented. Yet, it is surprising that its impact has not been greater and that so many descriptions of electronic signatures still gloss over the issue.

Legally speaking, the situation is one of stalemate. A person having proof of his rights in electronic form will face a choice between:

- losing everything due to the inevitable obsolescence of his document; or
- the certainty of seeing his evidence discredited as a "forgery" by its digital fingerprint if the document undergoes the slightest change required to preserve it.

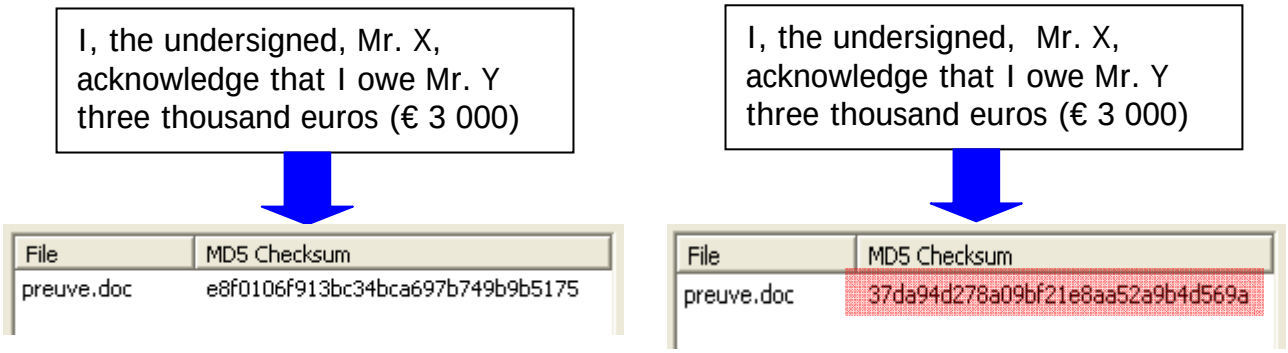
This is critical because the essential role of documents of proof is to carry forward into the future the rights and obligations to which they bear witness. And "future" is not just an empty word here. Legislation defining statutory limitation periods requires documents of proof to be archived for at least 20 years[5], without prejudice to regulations stipulating longer periods, as required in the case of real estate title documents and certain types of health records, for example. It is clear that, over such periods, several data format migrations will be needed.

Faced with these established facts, we can plainly see that electronic signatures cannot be archived as such because they do not survive conversion, and migration is essential if documents are to be preserved in digital format.

Pandora's Box

There seems to be no way out of this situation. Admitting that the fingerprint of a digital document may be distorted by a mere variation in data format would open up Pandora's Box. Such a loophole would naturally be exploited. If we look back at our example, Mr. Y could pursue his dishonest intent for little cost: he could change the substantive content of the document and then perform a simple format migration as an alibi for his fraud.

At worst, an interpretational error, based on a document's digital fingerprint, may just as easily be caused by an insignificant change to its contents, as in this new example:



[5] See *La Preuve Numérique Face à la Réalité* (Digital Evidence and Reality) http://www.megapreuve.org/cariboost_files/Prescription_20extinctive.pdf

Something has been changed in this instance: a space has been added between "I, the undersigned" and "Mr. X". The hash code indicates that an alteration has been made but there is no "forgery". Article 441 (1) of the Penal Code[6] provides that forgery is "*any fraudulent alteration of the truth liable to cause harm* ..." In this example, the truth has not been altered and the addition of a space in this position will harm no one. Nevertheless, this document will still be deemed a forgery because of the change in its hash code. Clearly, in this case, the detection of an "invalid" digital fingerprint is purely a red herring and ultimately blurs the search for the truth.

The Right to Reliable Evidence

This analysis demonstrates that electronic signatures do not afford data the degree of legal certainty with which they are credited. Potentially, such signatures create more confusion than certainty, labelling as "false" what is actually "true" after the inevitable conversion processes, and failing to differentiate clearly between what is important and what is insignificant.

However, the worst is undoubtedly the pernicious nature of the current situation. At this very moment, people acting in good faith and applying a published French standard may be converting the format of their documents of proof with the sole aim of maintaining up-to-date files, not realising that they are acting to their detriment if their documents bear a digital fingerprint.

The very principle behind electronic signatures requires their process to be beyond our understanding. Such a system is therefore only workable if we are willing to trust it blindly and accept this unknown element. However, it now seems clear that this confidence should be tempered with caution.

Reliable evidence plays a pivotal role in court rulings and, as it has already been written[7], "*the practical downfall and loss of credibility of means of proof are the harbingers of social and economic disaster because the means of settling disputes are deprived of their efficacy in consequence*."

Paris, 27 September 2010

6 Article 441 (1) of the Penal Code: "*any fraudulent alteration of the truth liable to cause harm made by any means in a document or other medium of expression, the purpose or effect of which is to provide evidence of a right or situation carrying legal consequences, constitutes forgery.*
Uttering or using forged instruments is punishable by up to three years' imprisonment and a fine of 45 000 euros."

7 In *Archivage des Documents Numériques à Vocation Probatoire* (Archiving Digital Documents as Proof), Groupe PragmArchive http://www.pragmarchive.org/PA_002.PDF