Proof by Metadata

Words, words, words…

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

What distinguishes the digital format and makes it so appealing is its adaptability to electronic processing. However, this flexibility turns into a drawback when it comes to proving the data stored in this digital format. As digital storage perpetuates the ability to reprocess the data, and given the powerful capabilities of information technology, this issue needs to be approached with great caution. Such concerns are further fuelled by the instability of electronic archiving media, requiring repeated handling of the data and thus offering an ideal opportunity to manipulate them. Accordingly, some favour a solution that provides a *trace* of what happens to the data. For example, the French Standard AFNOR NF Z 42-013 defines a long list of information to be gathered each time an operation is applied to the data. In other words, this strategy involves collecting *metadata* to prove that the document they trace has not been altered. In analysing this approach, we need to assess:

- what level of probative reliability we can attain;
- how feasible it is; and
- whether or not it flouts the rules of evidence.

Is it reliable evidence?

Compiling traces when archives have to be migrated from one medium to another to preserve data integrity raises a fundamental question: on what medium should the metadata themselves be stored? Should they be recorded on a durable and irreversible medium, or on a similar medium to those that make such migrations necessary in the first place? The answer lies in the question.

If we were able to record these metadata on a durable and irreversible medium, that would imply that this type of medium is readily available, thus rendering the metadata redundant, if not suspect. If this were possible, it would beg the question as to why the digital documents themselves are not stored on such a medium, which would be a more transparent and immediate answer to this problem.

Clearly, collecting metadata for evidential purposes can only be justified where no other method of proof is available. Otherwise, there would be no reason for collecting such data in the first place. The sole reason for the existence of metadata is to compensate for the general lack of reliable media, and yet ironically these metadata can only be stored on media that are as insecure as those which give these metadata their very purpose. QED.

The implications here are particularly significant in that, once recorded on whichever type of digital medium, the metadata become … plain and simple data, just like the data they are intended to trace. They are no less "digital" and are just as vulnerable to data processing systems.
This whole issue is far too uncertain and, on the face of it, it is very difficult to see how transferring evidence of digital data onto other digital data might provide better safeguards.

Practical considerations

For the sake of discussion, let us suppose that we have stored various digital documents in an electronic archiving system and that, some five years later, we will need to migrate them to avoid the risk of erasure. We will therefore need to change the storage medium, taking care to record the detailed circumstances in which this migration occurs. Once this process is completed, we will have two types of information: the migrated data and the metadata tracing the migration.
Three years later, a change in operating system requires us to perform another migration to prevent the data from becoming obsolete. This presents no apparent difficulty, the data just needs to be transferred from one system to another, as before.
However, we need to bear in mind that we are now dealing with data transferred over from an initial migration, together with the metadata mapping that migration. Losing these metadata is out of the question because they provide a trace of the migration for evidential purposes, making it all the more important to safeguard their integrity. Consequently, we now have to transfer the already-migrated data along with the previously-collected metadata, while producing a trace of the entire process. This effectively boils down to:

- tracing the *migration of a migration*;
- tracing the metadata, i.e. *tracing the traces*; and
- storing *metadata about the metadata*.

The problem here is that this exponential process is likely to endure; firstly, because evidence needs to be retained over the long term, and secondly, because the instability of information technology and digital media is unlikely to be resolved in the foreseeable future. On top of this, there is the issue of potentially daily backups, which need to guarantee the same level of integrity and should be subject to equally stringent standards with regard to tracing. It is difficult to contemplate the unwieldy mass of data that a statutory limitation period of 20 years or more might produce.

Legal issues

We need to pit this tracing system against the rules of evidence. Here, we will be looking at:

- the rules governing pre-constituted evidence;
- the requirement for *inter partes* proceedings, allowing each party a fair hearing;
- the prohibition against "creating evidence to support one's own case".

PRE-CONSTITUTED EVIDENCE

Pre-constituting evidence is not restricted to proving a document. It constitutes evidence in the intellectual sense of the term. The document is no more than a material manifestation of the evidence. If we are looking ahead to the need for collecting traces to prove a document,

this means that the document proves nothing in itself and that the evidence is not *pre-constituted*. On the contrary, it shows that we are seeking to create evidence after the fact.
This seems contrary to all logic in general and to French law in particular. Article 1341 of the French Civil Code requires not only a written instrument by way of proof but also excludes information not directly contained in that instrument: *"A written instrument is required […] and no witness testimony will be admissible as evidence against or in addition to the content of such instruments or as evidence of what is alleged to have been said before, at the time of or after such instruments …"*. In other words, any information not initially recorded can have no legal impact on the instrument despite any connection to it. This exclusion would appear to extend to the metadata we are considering. Metadata would be supplementary information, prepared "after the fact" for evidential purposes by human beings who, in this case, are merely testifying as to what they did to the instrument in question.
To all intents and purposes, recording traces relating to a document and making a legal outcome dependent on producing a record of these traces is tantamount to adding major sections to a document after finalisation or signature. In any event, this whole issue is very unclear.

REQUIREMENT FOR 'INTER PARTES' PROCEEDINGS
Allowing both parties to argue the merits of evidence is a fundamental requirement for its validity in legal proceedings and at any time during those proceedings. It is imperative for a party to have an opportunity to discuss or contest the evidence adduced against him. This is a requirement under the principles of litigant equality and "equality of arms"[1].
Therefore, if one party relies before a court on traces collected unilaterally and of his own volition, this would deny his opponent the opportunity of discussing the traces on equal terms and raising grounds for contesting that evidence. On this basis, it is very likely that the metadata referred to here would be ruled inadmissible.

ONE CANNOT CREATE EVIDENCE TO SUPPORT ONE'S OWN CASE
Evidence created by the person seeking to rely on it is prohibited under the principle that no-one may create evidence to support his own case. It is inadmissible for a party to attempt to prove facts by means of a document that he has himself created or which alleges his own conduct.
This prohibition clearly rules out any strategy whereby proof of a document depends on metadata collected by the party himself.

An admission of inadequacy
Those who advocate proving data using metadata also recognise that their medium is unsuited to proving digital data. The long litany of information needed to comply with French Standard NF Z 42-013 speaks volumes. There are pages and pages of it! There is also no basis for this because the Standard fails to explain why a particular metadata item needs to be stored as opposed to another and, more significantly, what all these data are actually supposed to prove. In any case, the Standard admits – inadvertently no doubt – that digital media do not make good evidence.
In any event, the idea of using metadata as a means of proving data integrity fails the logic test. A trace merely explains what has been done. It falls short of showing what has been

[1] See, *inter alia*, Article 6 (1) of the European Convention on Human Rights and Fundamental Freedoms

hidden. When given his day in court, an opponent in litigation will obviously highlight that, in a case of malfeasance, the culprit will do his utmost to erase all trace of his actions, especially where he has control over what is recorded and what is not. Worse still: like the arsonist who guides fire fighters to the scene of the fire, the trace log could be seen as a list of the system flaws known to the data holder.

Reliability is out there

In the absence of any alternative solutions, it is understandable that such an over-refined method of proof needs to be cobbled together. However, durable and irreversible media, capable of storing digital documents, are clearly available. They are described and standardised in the AFNOR NF Z 43-400 (2005) and ISO 11506 (2009) Standards, which explain the reasons for their probative weight. The digital-analogue archiving system put forward by these two standards, referred to as "dual-recording", is particularly relevant because it covers all requirements and provides real guarantees for stable and durable storage.
If we consider that the purpose of an electronic archiving system is to ensure an adequate lifespan for documents, to guarantee proof of what they are, and to facilitate their electronic management, we can see that Standard ISO 11506 provides the most secure, simple and clearly-defined electronic archiving solution.

Proving proof?

We can see from this review that the strategy of proof by metadata is very tenuous. It highlights the lack of reliable digital media, raises feasibility issues, fails to meet legal requirements and, above all, is likely to be ruled inadmissible by any court.
The whole process lacks technical and logical analysis. If such a tracing scenario needs to be contemplated, it is because, intrinsically, the digital format does not constitute proof. Therefore, seeking to prove digital data using other digital data – even if we call it something else – just seems to create a vicious circle. We might even wonder whether some of us are not simply confusing "metadata" and "metaphysics"…
Adducing solid evidence is an obligation and a duty of the highest importance, meriting reliable and unequivocal methods commensurate to what is at stake legally and socially in the process.

Lucien PAULIAC
President of the Association
"*Preuve & Archivage*"