Preliminary communication

Digitally Signed Records – Friend or Foe?

Boris Herceg, Hrvoje Brzica
Financial Agency – FINA
Ulica grada Vukovara 70, Zagreb, Croatia
boris.herceg@fina.hr, hrvoje.brzica@fina.hr

Hrvoje Stančić
Department of Information and Communication Sciences
Faculty of Humanities and Social Sciences
Ivana Lučića 3, Zagreb, Croatia
hstancic@ffzg.hr

Summary

Long-term preservation of digitally signed records may be a challenging task. Digital signatures expire and digital certificates may be revoked, thus influencing the trustworthiness of archived digital records. The authors firstly explain the problems and then proceed to a description of the planned research. The research will be conducted on a sample of digitally signed and archived electronic forms stored in the PDF file format, originating from the period from 2006 to 2009.

Keywords: electronic forms, digital records, digital signatures, PDF, file format, long-term preservation

Introduction

Thibodeau (2002) said that "the preservation of digital objects involves a variety of challenges, including policy questions, institutional roles and relationships, legal issues, intellectual property rights, and metadata" but also that "the variety and complexity of digital information objects engender a basic criterion for the evaluation of possible digital preservation methods, namely that they must address this variety and complexity." Becker, Kulovits, and Rauber (2010) further explain that since "a digital object needs the correct environment in order to function, we can either recreate the original environment (emulation) or transform the object to work in different environments (migration)". However, the digital object could also be converted to a newer version of the same file format, or to a different file format (conversion). All those changes could impact the trustworthiness of a digital record, i.e. influence its authenticity, reliability, accuracy, integrity, and/or usability. In order to prevent any of the possible unwanted changes, Duranti (1999) explicated that "irrespective of the long-term solution for the preservation of authentic electronic records, it is quite clear that there will not be much worth preserving for the future if serious measures

are not taken by records creators to guarantee the trustworthiness of electronic records (in both meanings – trustworthiness of content and trustworthiness of the record as a record) since the moment of creation." This could be a complex task if the records being preserved are digitally born records signed with (advanced) digital signatures which depend on the (qualified) digital certificates and (trusted) timestamping process, and are entrusted to the cloud (Stančić, Rajh, & Brzica, 2015). This article will focus on the issues dealing with the long-term preservation of digitally signed records, specifically on the records in the .PDF file format, and explain the planned research .

The problem

Digitally signed records, although well preserved over the long-term period, could lose their legal validity if the digital signature cannot be validated, or if it loses its characteristic of non-repudiation. Adobe confirms that "the mere existence of a digital signature is not an adequate assurance that a document is what it appears to be". If the validity check of a digital signature results in an error, e.g. no proof of existence, the record's trustworthiness may be compromised.

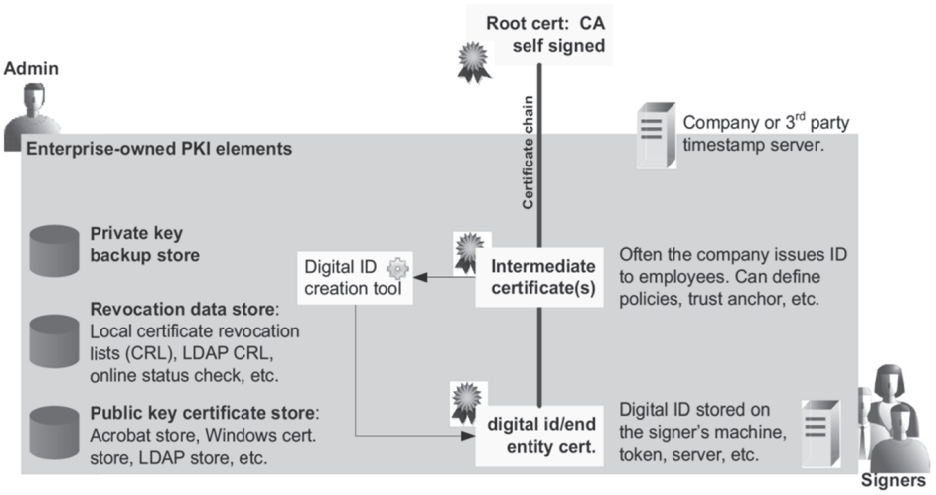

Figure 1. Common PKI elements in signature workflows (Adobe, p. 3)

The problem is that digital signatures have a valid time span, and that their validation requires a connection with the certification authority relying on the Public Key Infrastructure (PKI) (Figure 1). If any of the elements in this chain fails, the validity check will fail too. This is even more important if the records with advanced digital signatures, i.e. digital signatures using digital certificates, are being preserved. Digital certificates expire, but they also may be revoked. Should the historic information on the revocation lists be preserved along with the digitally signed records? Is that technically possible? However, this problem

should be appropriately addressed before it appears. In the case of this planned research, the digitally signed records are already archived, and it yet remains to be investigated if the passage of time will influce/influences their trustworthiness. The results of this preliminary research will provide some answers relevant to the more crucial questions recognised above, and will help prepare a later stage of the research in which those questions will be addressed in full.

Research

The research will be conducted on electronic forms of the public administration services created from 2006 up until 2009. The electronic forms subject of research will be the electronic form of the Croatian Pension Insurance and the electronic form of online registration to the Court Register, both created as PDF documents and digitally signed. The working hypothesis is that the technological progress has no effect on the long-term preservation of the content and the key elements of electronic records, and that both the content and the key elements of electronic records are fully preserved. Of course, this hypothesis can prove to be true, partially true or false.

Methodology

The methods of collection, sampling, parametrisation, testing, comparison, analysis, synthetisation, and abstraction will be used in the course of the study.

The plan is to organise the research in several steps as follows:

1. Definition of the research parameters and the testing environment.
2. Collection of samples of electronic records (.PDF files) for analysis from the production system of each electronic service, i.e. from its archives.
3. Organisation of all samples of electronic records of one electronic service into a single directory in the test environment.
4. Duplication – a copy of the sample will be made before the testing in order to ensure that the testing has no impact on the original sample.
5. Testing will be done by opening the sample PDFs in the version of the reader from the time of record creation, following by opening the same sample in the consecutive, newer versions of the reader. During the testing, it will be necessary to examine whether the individual Reader version contains a documented bug. In that case it will be necessary to register these findings and take the next stable version of the Reader in which a particular bug was corrected. The characteristics of the records will be investigated.
6. Data analysis.
7. Synthetisation of the findings.
8. Writing of the final report and recommendations.

For the purpose of this research the following characteristics of the archived electronic forms of the public administration services in the .PDF format will be investigated:

1. Readability of the content.
2. Validity of digital signatures.
3. Functionality of digital signatures' visualization.
4. Display of digital signatures' elements.
5. Size of digital signatures.
6. Size of electronic records.
7. Legal usability of the electronic records.

Other characteristics may prove to be relevant as well.

The testing environment should include:

1. Possibility of installation of all versions of the reader using the latest OS environment.
2. Ability to recognize extensions and attributes of electronic records.
3. The possibility to duplicate the testing sample without losing the attributes of the sample.

Expected results

The researchers hope to prove the hypothesis, i.e. to find out that the technological progress has no effect on the long-term preservation of the digitally signed PDFs, and that both the content and the key elements relevant to the concept of trustworthiness of electronic records are fully preserved. These expectations are based on the facts that the investigated records are stored in the stable and widely accepted file format and that the vendor takes into the consideration the backward compatibility of their products.

Results of this research will be used in the next stage in which the preservation of the historic information on the revocation lists, along with the digitally signed records, as well as the application of timestamps to the records with the expiring certificates, will be investigated.

References

Adobe. (n.d.). Digital Signatures in a PDF. Retrieved September 15, 2015, from https://www.adobe.com/devnet-docs/acrobatetk/tools/DigSig/Acrobat_DigitalSignatures_in_PDF.pdf

Becker, C., Kulovits, H., & Rauber, A. (2010, 1). Trustworthy Preservation Planning with Plato. (P. Kunz, Ed.) *ERCIM News - Special theme: Digital Preservation*, pp. 24-25.

Duranti, L. (1999). Concepts and principles for the management of electronic records, or records management theory is archival diplomatics. *Records Management Journal, 9*(3), 149-171.

Stančić, H., Rajh, A., & Brzica, H. (2015, June). Archival Cloud Services: Portability, Continuity, and Sustainability Aspects of Long-term Preservation of Electronically Signed Records. *The Canadian Journal of Information and Library Science (CJILS), 39*(2), 210-227.

Thibodeau, K. (2002). Overview of Technological Approaches to Digital Preservation and Challenges in Coming Years. In *The State of Digital Preservation: An International Perspective* (pp. 4-31). Washington, D.C.: Council on Library and Information Resources (CLIR).