COMPARATIVE ANALYSIS OF INTERNAL STRUCTURE AND FUNCTIONS OF DIGITAL ARCHIVES PRESERVING COMPLEX ELECTRONIC RECORDS

Hrvoje Stancic
Faculty of Humanities and Social Sciences, University of Zagreb, Croatia
Boris Herceg
FINA – Financial Agency, Croatia
Arian Rajh
Agency for Medicinal Products and Medical Devices, Croatia

Summary

The authors conduct structural and functional analysis of four implemented digital archival systems and a reference one followed by their comparative analysis. They analyse digital archive implementations in The Federal Chamber of Architects and Engineers – BAIK (Austria), The Agency for Medicinal Products and Medical Devices – HALMED (Croatia), The Braunschweig Clinic (Germany), and Lithuanian Office of the Chief Archivist – Electronic Archive Information System – EAIS (Lithuania). A reference digital archive system is made by the Federal Office for Information Security – BSI (Germany). The digital archival systems are responsible for preservation of digitised or digital records in the land register (BAIK), classified records on medicines and medical devices granted to the market (HALMED), time-stamped electronic health records (Braunschweig), records of the public administration created as part of governmental e-services and signed by advanced electronic signatures (EAIS), and it should ideally be responsible for electronically signed records (BSI) as well. After the structural and functional analysis of each system, they are comparatively analysed according to the criteria grouped around functionalities and system processes, implemented standards and formats and available software tools for management of preserved records. The authors give conclusion on the issues regarding long term preservation of electronic records having additional preservation requirements due to the time-stamping, addition of (advanced) electronic signatures, and (qualified) certificates in the situations when it is necessary to ensure their authenticity, integrity and non-repudiation.

Key words: digital records, long-term preservation, digital archive, time-stamp, electronic signature, e-services

1. Introduction

Although, in comparison to analogue records, electronic records by themselves are complex to preserve because of the constant and rapid advancements of information technology, standards, formats etc., we refer to the records as complex if they have additional requirements during the long term preservation, e.g. preservation of integrity, authenticity or reliability, thus having additional level of complexity such as the addition of (advanced) electronic signature, (qualified) digital certificate, certificate chain, revocation status, trusted timestamp etc. Preserving such records for the long term is technically challenging if those requirements are to be met.

Before the research presented in this article we have analysed different digital archive implementations and research done so far on them. We have found out that digital archives have so far been scientifically examined on the specific topics of projects assessments, usage of available long term preservation technology and standards, conceptual and terminology related issues, and possible strategic and planning approaches as well as many others. In this article we analyse four digital archive implementations and one reference model for a digital archive. After the initial structural and functional analysis of each system, they are comparatively analysed according to the criteria grouped around functionalities and

system processes, implemented standards and formats and available software tools for management of preserved records.

From the larger pool of digital archive implementations and reference models for such systems functioning as research candidates we have selected four implemented systems and one reference model to be analysed and compared in this article. The four implementations were selected on the basis of covering important aspects of digital archival solutions and the reference one on the basis of intensively addressing aspects of long term preservation of digital records signed by electronic signatures. Therefore, in this article we analyse digital archive implementations of Austrian *Federal Chamber of Architects and Engineers* (BAIK), Croatian *Agency for Medicinal Products and Medical Devices* (HALMED), German *Braunschweig Clinic* and *Lithuanian Office of the Chief Archivist – Electronic Archive Information System* (EAIS). The analysed reference model is made by the *Federal Office for Information Security* (BSI). The analysed systems were selected because they are responsible for the records that have various levels of complexity, use various authentication mechanisms and several data and cryptographic formats. They preserve digitised or digital records in the land register (BAIK), classified records on medicines and medical devices granted to the market (HALMED), time-stamped electronic health records (Braunschweig) and records of the public administration created as part of governmental e-services and signed by advanced electronic signatures (EAIS). The reference one addresses preservation of electronically signed records (BSI).

In order to better understand formats of electronic signatures mentioned in this article a brief systematisation is needed. There are two types of electronic signatures – basic, usually referred to as an “electronic signature”, and advanced. Electronic signature “uses a cryptographic transformation of the data to allow the recipient of the data to prove the origin and integrity of the subject data” (Electronic signature, 2012) while advanced electronic signature sets higher requirements in order to establish the records’ characteristic of non-repudiation. Further, there are four main types of electronic signatures: 1. XMLDSig (XML Signature) which can be detached, enveloped or enveloping, 2. XAdES (XML Advanced Electronic Signature) with six additional, mutually nested forms: a. XAdES-BES – Basic Electronic Signature, b. XAdES-T (Timestamp), c. XAdES-C (Complete validation data), d. XAdES-X (eXtended validation data), e. XAdES-X-L (eXtended validation data incorporated for the Long term), and f. XAdES-A (Archiving validation data), 3. CAdES (CMS (Cryptographic Message Syntax) Advanced Electronic Signatures), and 4. PAdES (PDF Advanced Electronic Signature) (Brzica et al, 2013).

2. Aim of the research and research methodology

The aim and motivation for this research was to detect if the digital archive solutions have standardised their internal structure and functions, if they have developed solutions for the long term preservation of complex electronic records, and if they have based their solutions on open or proprietary standards. Method of selection was used prior to writing this article in order to choose the four representative digital archive information systems and one reference model for the digital archive system. Here, decomposition is used as the preparatory method for the method of analysis which is then used to conduct structural and functional analysis of the chosen systems. The results are comparatively analysed and synthesised in the form of conclusions.

3. Analysis of digital archive implementations

3.1. The Federal Chamber of Architects and Engineers – BAIK (Austria)

Austrian Federal Chamber of Architects and Engineers established a modern electronic documents archive (BAIK)[1], which is used to store document collection of land register. BAIK archive provides access to the data for public users and state authorities and offers an electronic archive where original documents are archived in accordance with the regulations of the archives of public bodies.

BAIK allows ingesting and querying of the data and documents. All archived electronic documents are originals in the legal sense. For this BAIK uses the technology of digital signatures which ensure that the archived documents have the status of originals. Public and private electronic documents can be ingested and stored electronically.

3.1.1. Architecture of BAIK archive

There are several key components of the BAIK archive's internal architecture. They are shown in Figure 1 and explained further in the text. The BAIK archive stores not only document collection of land register but also opinions (in the form of reports) of the civil engineers on the matters requested by the clients.

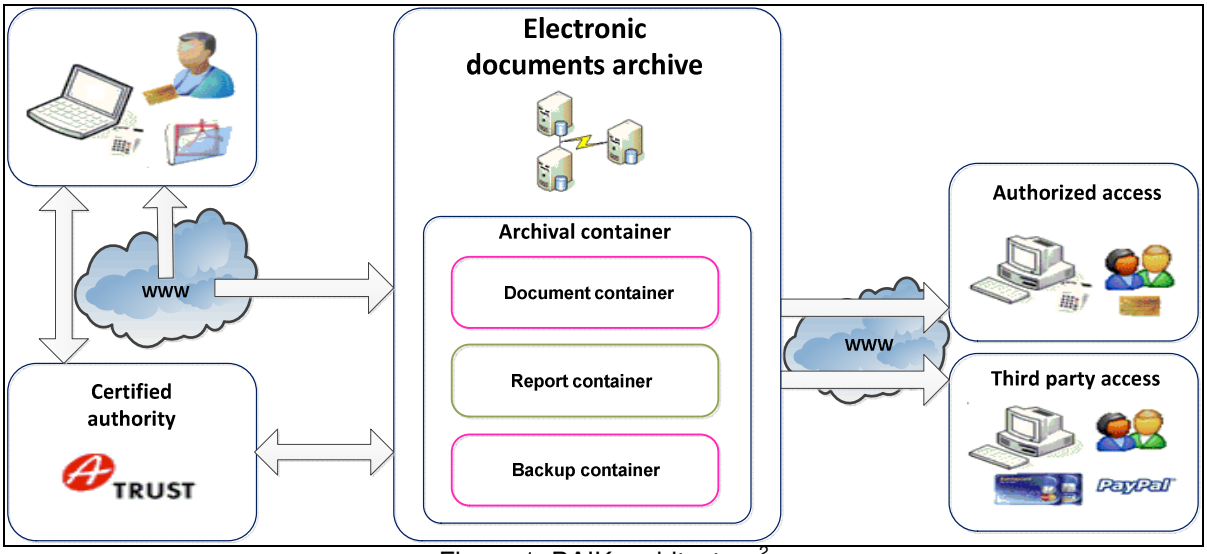

Figure 1: BAIK architecture[2]

Document container

The document container contains public electronic documents that are created by authorities, and are placed in the collection of documents of the land register, or are open to public for inspection. Public documents are stored in the PDF format (PDF/A-1b).

Report container

The report container contains private electronic documents that are stored with the consent of the client. Client enables a civil engineer electronic access to these electronic documents in order for him to create an opinion in the form of report. Customers can grant any other person electronic access to these electronic documents. Private electronic documents can be saved in any format, but the official report must be stored in the PDF/A-1b format.

Backup container

The backup container contains private electronic documents or civil engineers' electronic documents that are stored for lost prevention. In the backup container electronic documents can be saved in various types and formats.

3.1.2. Archived documents' attributes

For better identification and/or traceability of documents in the archive a standard and an additional set of metadata are implemented. The standard set provides basic information on the document and client while the additional set provides more detailed information.

Standard set
1. Document date
2. Project name
3. Client identification
4. Client address
5. Client ZIP number
6. Client place
7. Business Number (optional)
8. Subject of the document
9. Digital or scanned
10. Order Date (optional)
11. Comments (optional)

Additional geodesic set
1. Location reference
2. Street location reference (optional)
3. Local reference site (optional)
4. Gauss-Krüger coordinates (optional)

3.1.3. Technical background

All documents and accompanying files in the document container must be in the format PDF/A-1b. Only XMLDSig type of electronic signature should be used. The company A-Trust[3] is functioning as an issuer of digital certificates and as the certified authority for BAIK access and security.

3.2. The Agency for Medicinal Products and Medical Devices – HALMED (Croatia)

The Agency for Medicinal Products and Medical Devices (HALMED) is Croatian National Competent Authority entrusted for regulation of medicines, medical devices and homeopathic products. It was established by the Croatian government on 1st October 2003 as a legal successor to the Croatian Institute of Medicines Control and the Croatian Institute of Immunobiological Preparations Control. HALMED's core processes include granting marketing authorisations, performing quality control of the products, monitoring adverse reactions and vigilance of products and devices on the market and under clinical trials, drawing up national Pharmacopoeia, issuing manufacturing, wholesale and retail licences as well as issuing good manufacturing practice certificates, approving import and export of medicinal products, monitoring the consumption of medicinal products, and educating and providing public and professional information.

3.2.1. Project

HALMED has build-up its capacity through implementation of two IPA projects.[4] While IPA 2007 TWL (Twinning Light) project included arrival of experts from Spanish Agency for medicinal products to HALMED and thus it contributed to strengthening HALMED's professional capacity, IPA 2009 project "Preparations for eCTD and Implementation of Digital Archival Information System" strengthened its administrative capacity. IPA 2009 project affected positively HALMED's ability to cooperate with other EU Member States on the basis of sharing common practice like working electronically and working with eCTD[5]. Project consisted of two parts – business process analysis and implementation of enterprise content management system (ECMS). The project resulted with documented and controllable business processes, well-defined strategic plan, gained ability to manage electronic business resources (business models, records) and ability to distribute and archive electronic records using ECMS FileNet P8 platform[6].

3.2.2. DAIS ECMS

FileNet system was additionally customised for HALMED and largely aligned with ISO PDF/A and OAIS RM standards. Because of the OAIS-complying functionalities developed on FileNet P8 platform, the system was named Digital Archival Information System (DAIS). It consists of five modules. Basic module is the module for document and content management (Content navigator with ROS[7] repository). Second module is a module for records management or Enterprise records module with FPOS[8] repository of archival metadata for documents that were declared as records[9]. In addition, a module for ingest of submission information packages (SIPs), i.e. of digitised medicinal products' records, was built. Fourth module is a business process management part of DAIS system that contains models of processes, related process documentation and application for execution of processes – Process Designer. Fifth module is a workflow application for HALMED's quality management subsystem.

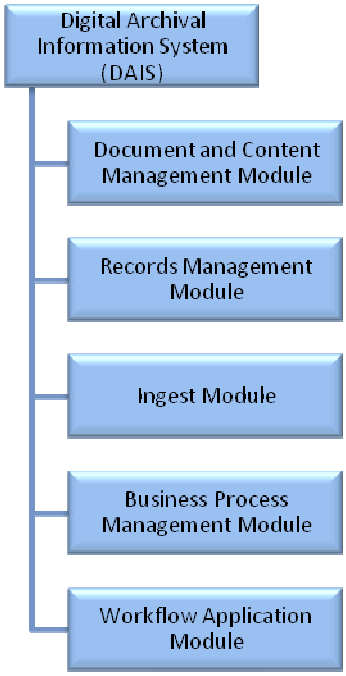

Figure 2: HALMED's DAIS architecture

During the project a number of generic methods for integration of DAIS and third parties' applications were developed. The most significant integrations, beside integration with two main business applications for medicinal products and medical and homeopathic devices, were integration with the case management subsystem and integration with the archival subsystem. Integration with the case management system enables distribution of electronic or scanned document according to administrative office's classification number and label of responsible organisational unit, as well as registration of document anywhere from DAIS or third party application into case management tool. Document in DAIS is stored in one location but it can be represented in numerous virtual directories by using file-in method. Integration with the archival tool, which includes retention schedule approved by the Croatian State Archives, enables that tool to control enterprise records and objects in FPOS repository. At the end of a business procedure (in some third party applications this is a semi-automatized function), when an electronic document needs to be archived, it is declared as a record and protected from users' modifications. Additional PDF/A version of the document is automatically created and checked-in.

Archival metadata and unique identifier is annotated in the archival application. That archival metadata and system metadata are stored in FPOS repository and represented in a directory named with the same unique identifier as in the archival tool. Records can be retrieved either from DAIS or from the archival application.

Migration module is used for transferring more complex content from the file-system into DAIS, content of previously received optical media into DAIS, and SIPs with metadata in XML files into DAIS. During migration of SIP validation procedure is executed. After SIP validation the procedure for declaration of records is initiated. In the case of digitised documents the creation and check-in of PDF/A version of a document are omitted because paper documents were already digitised as PDF/A files. Migration from DAIS with contextual metadata in XML files is also possible.

3.2.3. Records

There is a collection of numerous records stored in DAIS: digitised records originally created by marketing authorisation holders, digitally born records created by marketing authorisation holders (and eCTD records), as well as digitally born records of various classes, types and descriptions made by HALMED as the creator. HALMED's DAIS system is very flexible and seems rather resistant to obsolescence because it enables validation of conveniently created XML files used for SIPs, export of records from DAIS into other systems and integration with new business and administrative applications through the use of generic methods.

Classification system combines content-dependant and mandatory case management classification, archival classification with retention data, classification of documents types and subtypes controlled by business applications and FileNet document classes. There is one generic FileNet document class and numerous separate FileNet classes. Separate classes are usually created in business applications and sent to FileNet from business applications. Different FileNet document classes define different metadata schemes, descriptors for retrieval and document properties appearance. Metadata defined by a class can be exported in XML files along with records. An Example of metadata for OLIMP document class which covers processes of medicinal devices notifications, medical devices manufactures registration, certificates, vigilance of medical devices etc. shows the metadata structure:

1. Document title
2. Document ID

3. Case management number
4. Ordinal number of document in case
5. Status of document in business process
6. Related cases
7. Sender
8. Date of receipt
9. Type of document
10. Type of act

+ additional metadata for business(application)-related types

Medicinal (pharmaceutical) product's document can pertain to few FileNet classes: there is a class designed for digitally born documents processed in business application for medicinal products, there is a class for attachments to case files compiled by pharmaceutical industry in a special electronic format controlled by application used by majority of European medicinal agencies, and there is a class for digitised paper medicinal product documentation. HALMED uses its internal electronic signature and it has plans for implementation of legally valid electronic signature after broader implementation of electronic signature in Croatia and among European medicinal agencies.

3.3. The Braunschweig Clinic (Germany)

The period of storage of medical records may be 30 years or more. Skiagraph or CT scans (tomographic images) as records often need to be kept over the long term. Because of the legal certainty it provides, advanced electronic signatures is essential for establishing medical evidence, particularly because it is using the time stamping procedure which provides irrefutable evidence of the time of manufacturing.

When digitizing medical records in PDF/A format it is possible to scan colour documents, search the text after the OCR process, it has relatively small file size, and it is possible to embed electronic signatures in the document. PDF/A-1 is a basic standard for archiving of electronic documents, but only PDF/A-2 and PDF/A-3 formats support embedded binary content and machine-readable data (Wild, 2012).

Braunschweig clinic is using PDF/A standards to improve the process and reduce costs of filing medical records. Changes were made in order to include medical records in the hospital information system – implementation of electronic health records, electronic signatures, electronic archives. This resulted with lowering of the costs of long term archiving and communication or collaboration during hospital procedures. The findings, diagnoses, medical imagery and administrative data are stored centrally, and they can be accessed from different locations. Much of the paper documents were scanned and archived in a centralized hospital IT system where they can be easily accessed.

There are two processes (Figure 3) during the archiving of the scanned documents and those two processes, in the case of Braunschweig Clinic, harmoniously complement each other (Wild, 2012):

- convert scanned images to PDF/A format in accordance with industry standard ISO 19005,
- add electronic signatures to PDF/A thus guaranteeing the integrity, authenticity and inalterability of documents in the electronic archive.

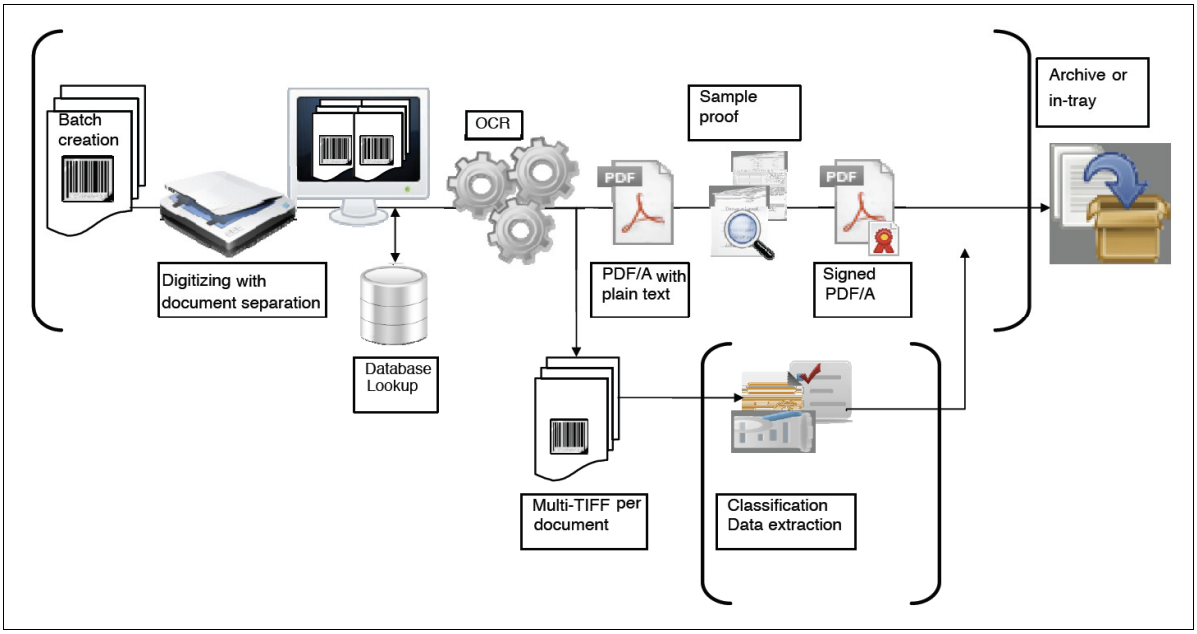

Figure 3: Scanned and signed electronic documents archive (Wild, 2012)

The example in Figure 3 represents a particularly safe and simple way to scan and sign documents in computer (batch) processing that is performed on the server. It can be secure especially if the scanning and signing procedures are either done at the same place during the same process or by the signing authority.

3.4. Lithuanian Office of the Chief Archivist – Electronic Archive Information System – EAIS (Lithuania)

Development of Electronic Archive Information System (EAIS[10]) in Lithuania was one of the steps in preparation of the Lithuanian public administration to work with the electronically signed documents. EAIS allows saving official documents signed by advanced electronic signatures thus ensuring integrity, authenticity, non-repudiation and the ability to use and store documents in the long term.

The first Lithuanian system for electronic signatures was e-Servicing System of Insurers (EDAS[11]) implemented in 2007 and its purpose was to electronically sign documents in the field of activity of State Social Insurance Fund Board of the Republic of Lithuania[12]. EDAS system uses forms based on XML, and the signature is done by using XAdES signature format. The Lithuanian model required that XAdES signatures format has possibility to sign several data objects fields in one document, and to implement several electronic signatures. Metadata could be separately signed in a subtree of the XML metadata file. The basic principle of the Lithuanian approach is that the metadata are integral part of an electronic document.

Interoperability solutions in Lithuanian public administration are defined by using two specifications for complex electronic documents throughout their life cycle. Specifications ADOC (ADOC, 2009) and MDOC have been approved by the Office of the Chief Archivist of Lithuania, Lithuanian Archives Department in 2011:

- ADOC – Specification for describing and dealing with human readable documents
- MDOC – Specification to describe and handle machine readable documents

EAIS architecture (Figure 5) consists of three main parts (Ragaisis, 2012a, p. 5):

- public portal (https://eais-pub.archyvai.lt) serving all external users,
- internal Portal (https://eais-int.archyvai.lt) serving members of government archives and the Office of the Chief Archivist of Lithuania, and
- archive of electronic documents.

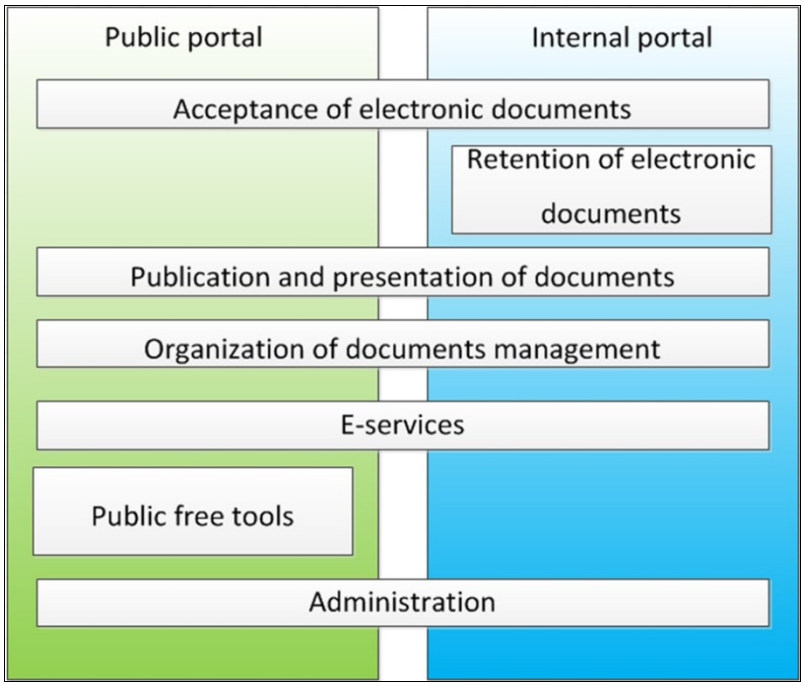

Figure 5: EAIS architecture (Ragaisis, 2012b)

As shown in the Figure 5 EAIS architecture consists of seven subsystems – five available from the public and internal portal, and two specific to only one of them:

1. Acceptance of electronic documents – this module provides the functionality of the transmission of documents to the state archives in the legally prescribed time either using packet transmission of documents through a computer network or downloading them from a physical media. Furthermore, this module checks the integrity of electronic documents, authenticity and compliance specifications with the purpose of preparing electronic documents for long term preservation. Finally, the functionality of this module is to store documents.
2. Retention of electronic documents – electronic documents include means for the physical preservation of electronic documents: backup copies, WORM storage, and funding for risk management. Retention subsystem is part of the internal portal.
3. Publication and presentation of documents.
4. Organization of documents management.
5. Electronic services.
6. Free software tools for preparation, signing, review and verification of official electronic documents – part of the public portal.
7. Administration.

Very important functionality of the EAIS system is flexible configuration of authentication. In fact, certain public bodies in Lithuania can only request an electronic signature on the entire document and the other authorities may require the use of advanced electronic signatures.

The possibility of using electronic documents on a lengthy period is provided by converting the contents of these documents in formats intended for long term storage – PDF/A formats. In addition, electronic documents are converted to the formats appropriate for previewing on the Internet – PNG and JPEG. There is possibility that the list of formats in the future will be updated (Luksaite Daiva, 2012).

XAdES-A format is used for the electronic signatures whose legal force should be preserved (integrity, authentication, non-repudiation) for the long term.

Part of the EAIS is the subsystem for retention of electronic documents and funding of risk management is a part of it, as already mentioned. EAIS identifies two types of risk. One is dealing with the risk associated with the formats of content because during the time a format may become obsolete and may no longer be supported by the current version of the software. This risk can be offset by converting the entire contents of the PDF/A format. The second risk is associated with electronic signature. A cryptographic algorithm that is used to create electronic signatures today can be sure but in the future may be rendered useless. This risk can be offset by additional authentication using time stamps on XAdES-A format.

The EAIS system is physically located at two geographically distant locations – one in Vilnius and one in Siauliai. The system was implemented so that it performs archival data replication between the primary and the secondary data centre with the ability to transfer one operation to another in case of errors or accidents. For security reasons, access to the storage space is only possible through an internal portal.

Success of the EAIS, i.e. of the application of the Lithuanian electronic public administration, has contributed to the creation of the publicly available free software tools (adapted prescribed specifications) for the preparation, signing, reviewing and verifying of the official electronic documents. Those tools are available through the web interface or as a desktop application.

Authentication of external users to EAIS is implemented through the electronic gateway of the Lithuanian public administration. Authentication service is provided for users of internet banking of all banks operating in Lithuania and owners of personal digital certificates. Certain EAIS functions are provided for unauthenticated users as well.

3.5. The Federal Office for Information Security – BSI (Germany)

German national security agency is Federal Office for Information Security (BSI)[13]. Its goal is to promote IT security in Germany. The BSI is the central IT security service provider for the federal government in Germany. BSI developed a model of long term storage of electronically signed documents which is based on international standards and the German law on archiving, German Federal Archiving Act[14], having in mind that in Germany the term “archiving” is defined by the federal and the state laws. The recommendations were adopted by the German (SAGA XOV, ArchiSafe) and international standardization initiatives (MoReq2, OAIS). Therefore, this is not an analysis of a live and implemented system, but of a model system. It would be interesting to see later on how it correlates with other analysed systems.

BSI has published a technical guideline (Preservation of Evidence, 2011) which describes the reference architecture of the system for long term storage of electronically signed documents and the technology behind it, i.e. a model system that we have decided to

analyse here and compare with other, previously explained, systems. This document defines requirements for each component of the architecture. The objectives and challenges the reference architecture for long term storage of electronically signed documents are that architecture should provide the following requirements for long term storage of digital content and metadata (Preservation of Evidence, 2011, p. 15):

- the availability and readability,
- integrity,
- authenticity,
- data protection, data security and confidentiality.

To achieve integrity and authenticity within the BSI reference architecture, the following conditions are mandatory (Preservation of Evidence, 2011, pp. 24-26):

- Electronic signature and time stamp must be created, verified, updated and stored in a safe and reliable manner. Also, they should be stored in accordance with the statutory provisions.
- Verification of data that will be needed later for verification of electronic signatures should be obtained immediately after the creation and/or verification of signatures. The verification data should be deposited together with the documents and information to be archived. Documents, data and verification data before archiving should be in a format suitable for long term preservation.
- All verification steps and verification results should be logged and be in the form allowing clear establishment of the relevant facts.
- Compatibility with standards and recommendations of BSI agencies and the Federal Network Agency15 must be ensured for the time stamps.
- Electronic signatures must be renewed before the expiration of the protective measures used in cryptographic algorithms. Renewal of the signature shall be conducted in accordance with the legal regulations and in a manner that is as much automated and cost-effective as possible.
- According to the dominant legal opinion, a document can be re-signed if it was originally electronically signed with a qualified timestamp including at least one qualified electronic signature.
- System components which display data need to be able to visualize data signatures, certificates and verification results.
- Integrity of the non-signed data transmitted to the ECM or long term storage can be secured with cryptographic security measures such as hash values or electronic signatures and qualified time stamps.

Below is a list of functional requirements of middleware in the BSI reference architecture (Preservation of Evidence, 2011, p. 28):

- archiving of signed and unsigned data,
- changes of metadata of archived data,
- requesting archived data,
- retrieval of records with evidence,
- deletion of archived data.

The BSI technical guideline also provides recommendations of document formats for long term storage. Further, it defines that metadata and verification data for long term preservation must be stored with the data documents to create archival information package within the XML syntax. Such a package is called XML formatted Archival Information Package (XAIP).

Detailed explanation of syntax and semantic rules of XAIP in appropriate formats and protocols are given in Annex TR-ESOR-F – Formats and Protocols (Annex TR-ESOR-F, 2011). Recommended data formats in architecture are XML, XSD, PDF/A, ODF, TIFF, JPG,

PNG. Recommended cryptographic formats are PKCS # 7, CMS, CAdES, XMLDSig, XAdES. As the certificates format X.509 is recommended. Certificate validation protocols recommendation is to use Online Certificate Status Protocol – OCSP (OCSP, 1999) and Server-Based Certificate Validation Protocol – SCVP (SCVP, 2007).

Annex TR-ESOR-F recommends implementation of Evidence Record Syntax – ERS standard (ERS, 2007). ERS standard is technically based on hashing value of XAIP that is organized as a hash tree. In the hash tree roots are provided or sealed with a qualified time stamps containing a qualified electronic signatures to ensure integrity.

3.5.1. IT reference architecture

The BSI's IT reference architecture used for archiving comprises of (Preservation of Evidence, 2011, p. 16) (Figure 4):

- Enterprise Content Management (ECM) – a system for long term storage that includes and manages a variety of storage media, and guarantees reliable and secure access to the storage medium when saving, accessing and deleting the archived documents and data,
- Middleware – includes cryptographic components that support the preservation of documents in accordance with the law regulations.

Application Layer

Applications and services in this layer should satisfy several important requirements (Preservation of Evidence, 2011, p. 38). One of them is creation of documents for archiving should be in a manner consistent with the standardized data formats that are recommended for long term preservation (e.g. PDF or XML). Next, they should provide the possibility of functional verification of electronic signatures. In order to ensure the verification of the validity of the signatures, even longer than the legally specified period of retention of signatures, it is recommended that the verification data on the validity of signatures should be saved immediately upon signing in order to be available along with the signature inside the archive. A "trusted viewer", i.e. a credible application component for displaying the advanced electronic signature, is needed in order to display documents and data signed with such signatures. The application layer should provide interface and functionality for archiving documents, finding and deleting documents filed, and for finding evidence records. Logging of all archival activities is mandatory.

TR-ESOR-Middleware

Modules and interfaces of TR-ESOR-Middleware are (Preservation of Evidence, 2011, pp. 43-45) ArchiSafe module, Cryptographic Module, and ArchiSig module.

The ArchiSafe module is a standardized and secure gateway that controls access of business applications to the ECM/long term archive. This module separates applications from the application layer and ECM/long term archive. Any action (write/change/delete) with specific application to the ECM/long term archive should be done through this module. The ArchiSafe module provides evidential quality of the information when electronic signatures are validated as the result of verification embedded in an XML document in a standardized form. Verification of the signature is done in the cryptographic module.

The Cryptographic module provides a variety of cryptographic functions that are necessary for preservation of evidence. This module performs the functions of creation of electronic signatures, verification of the signatures – specifically signed archival information packages, validation of certificates, calculation of hash values, and request verification of qualified time stamps.

The ArchiSig module's functions are responsible for "receipt and renewal of the probative value of electronic signatures and for the integrity of the archived information packages" (Preservation of Evidence, 2011, p. 45). The ArchiSig module is IT implementation of the previously mentioned ERS standard.

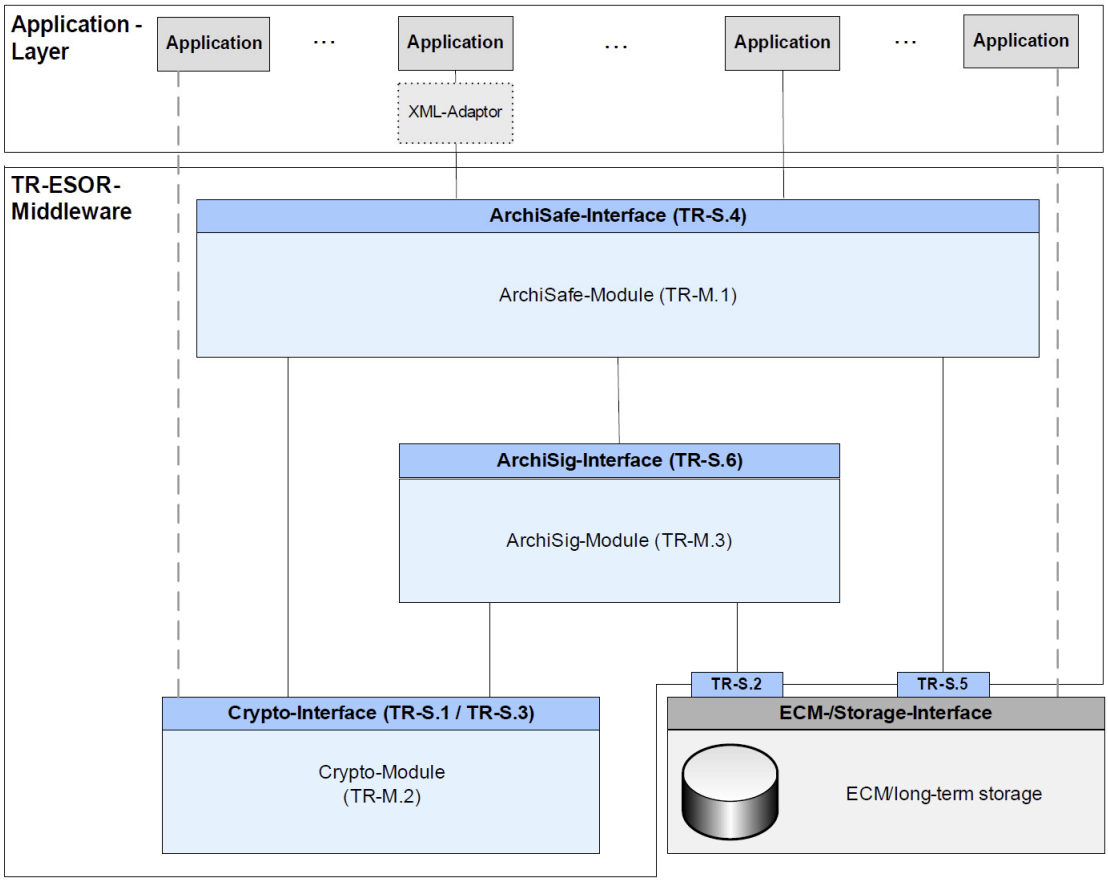

Figure 4: BSI referent architecture (Preservation of Evidence, 2011, p. 41)

4. Comparative analysis

Comparative analysis of the analysed systems is based on the criteria grouped around functionalities, implemented standards and formats, and available software tools for management of preserved records (Table 1).

From the Table 1 one can see that all analysed systems, no matter whether implemented ones or a model one, have the functionality of archiving electronic records. This was expected since the systems chosen for analysis were the digital archival systems. The results also indicate that they were chosen correctly. The second category for comparison was the category of archiving of electronic records that were signed by some kind of electronic signature. All analysed systems have this functionality either implemented or planned to be fully implemented. E.g. HALMED is currently using internal signatures only and is planning to implement the functionality of archiving electronic records with externally created e-signatures. The category of supportive modules shows the internal modules responsible for creation, management and archiving of electronically signed records. It can be seen that in every system there is a dedicated module. They are called differently but are responsible for similar activities. In the category of implemented standards we have put standards and standard file formats. It can be seen that PDF/A, or some specific version of it, e.g. PDF/A-1, PDF/A-1b or PDF/A-2, is used. Also, three out of five systems use XML and two are OAIS-based. Regarding the implemented standards for electronic signatures it can be seen that dominantly the detected types are based on XML (XMLDSig, XAdES). The technology that is used for establishing the authenticity is either digital certificates or

(qualified) time stamps. This is also in correlation with the detected usage of advanced electronic signatures (XAdES, CAdES).

Table 1: Comparative analysis

System	Archiving of electronic records functionality	Archiving of e-signed records functionality	Supportive modules	Implemented standards	Implemented standards for e-signature
BAIK	yes	yes	document container, backup container	PDF A-1b	XMLDSig with implementation of digital certificates
DAIS	yes	partially, records signed with internal signature only	FileNet ECM with migration module, enterprise records module, backup	OAIS RM, PDF/A-1 and 2, XML	integration of module for internal e-signature with DAIS system
Braunschweig Clinic	yes	yes	OCR, classification and data extraction	PDF/A	time stamp
EAIS	yes	yes	retention module, organisation of documents module	PDF/A, XML, PNG, JPEG	XAdES
BSI	yes	yes	ECM with ArchiSafe, ArchiSig and Cripto-modules	BSI Technical Guideline 03125, MoReq2, OAIS, XAIP, PDF/A, XML	XMLDSig, XAdES, CAdES, qualified time stamp

5. Conclusions

The comparative analysis showed that all examined digital archives emphasize the same basic problem – preservation of authentic content in some form. If this content is stored in a form of a record then its preservation and preservation of its metadata can be facilitated by using highly standardised file formats like PDF/A. The PDF/A file format and XML format for metadata are widely recognised solutions preferred by all digital archives considered in this analysis. One step further could be a suggestion to develop a solution compliant to the OAIS's information packages (SIP, AIP, DIP) using these two standards. Regardless whether the records archived in PDF/A are digitally signed or not we can conclude that the open

standards have been recognised by the analysed institutions and used in the implementation of their digital archival systems. Nevertheless, the process of long term preservation of complex electronic records will remain to be technically and organisationally challenging process. Basing it on the open standards instead of proprietary ones could bring stability to the process.

Future research

The authors will continue their research of digital archive solutions and cloud-based archival services. They plan to research portability, continuity and sustainability aspects of long term preservation of electronically signed records in the cloud environment.

Bibliography

ADOC (2009). <https://signa.mitsoft.lt/static/signa-web/webResources/docs/ADOC_specification_approved20090907_EN.pdf> [Accessed: 10/08/2014].
Annex TR-ESOR-F – Formats and Protocols (2011). Bonn: Federal Office for Information Security. <https://www.bsi.bund.de/SharedDocs/Downloads/EN/BSI/Publications/TechGuidelines/TG03125/TG-03125AnnexTR-ESOR-F.pdf?__blob=publicationFile> [Accessed: 09/08/2014].
Brzica Hrvoje, Herceg Boris, Stancic Hrvoje (2013). "Long-term Preservation of Validity of Electronically Signed Records". In: Gilliland Anne et al. (eds.). *Information Governance*. Zagreb: Department of Information and Communication Sciences, Faculty of Humanities and Social Sciences, University of Zagreb, pp. 147-158. <http://infoz.ffzg.hr/infuture/papers/4-03%20Brzica,%20Herceg,%20Stancic,%20LTP%20of%20Validity%20of%20Electronically%20Signed%20Records.pdf> [Accessed: 29/08/2014].
Electronic signature (2012). European Telecommunications Standards Institute. <http://www.etsi.org/index.php/technologies-clusters/technologies/security/electronic-signature> [Accessed: 29/08/2014].
ERS – Evidence Record Syntax (2007). Network Working Group. <http://tools.ietf.org/html/rfc4998> [Accessed: 09/08/2014].
Luksaite Daiva (2012). *The Life Cycle of e-Documents: Methodological and Legal Approach in Lithuania.* <http://www.rrt.lt/download/16522/5%20nb8%20archives%20lt-1.pdf> [Accessed: 10/08/2014].
OCSP – X.509 Internet Public Key Infrastructure, Online Certificate Status Protocol (1999). Network Working Group. <http://www.ietf.org/rfc/rfc2560.txt> [Accessed: 09/08/2014].
Preservation of Evidence of Cryptographically Signed Documents. BSI Technical Guideline 03125 (2011). Bonn: Federal Office for Information Security. <https://www.bsi.bund.de/SharedDocs/Downloads/EN/BSI/Publications/TechGuidelines/TG03125/TG-03125_main.pdf?__blob=publicationFile> [Accessed: 09/08/2014].

Ragaisis Saulius, Birstunas Adomas, Mitasiunas Antanas, Stockus Arunas (2012a). "Electronic Archive Information System". In: *Local Proceedings and Materials of Doctoral Consortium of the Tenth International Baltic Conference on Databases and Information Systems.* Vilnius,, vol. 924, pp. 107-114. <http://ceur-ws.org/Vol-924/paper11.pdf> [Accessed: 10/08/2014].
Ragaisis Saulius, Birstunas Adomas, Mitasiunas Antanas, Stockus Arunas (2012b). "Electronic Archive Information System". Doctoral Consortium of the Tenth International Baltic Conference on Databases and Information Systems, Vilnius. <http://www.mii.vu.lt/BalticDBIS2012/my_files/Presentations/Ragaisis_Birstunas_Mitasiunas_Stockus.ppsx> [Accessed: 10/08/2014].

SCVP – Server-Based Certificate Validation Protocol (2007). Network Working Group. <http://www.rfc-editor.org/rfc/rfc5055.txt> [Accessed: 09/08/2014].
Wild, Bernd (ed.) (2012). *PDF/A in Healthcare.* White paper, Berlin: PDF Association – PDF/A Competence Center. http://www.pdfa.org/wp-content/uploads/2012/05/WP-PDFA-in-Healthcare.pdf [Accessed: 08/08/2014].
ZHU, Wei-Dong, AITCHISON, Richard, BONNER, Eric, CASALS MENDEZ, Hector, RATHGEBER, Ron, YADAV, Amit, YESSAYAN, Harry (2009). *Understanding IBM FileNet Records Manager.* IBM. <http://www.redbooks.ibm.com/redbooks/pdfs/sg247623.pdf> [Accessed: 18/08/2014].
ZHU, Wei-Dong, BUCHANAN, Nicholas, OLAND, Michael, POGGENSEE, Thorsten, ROMERO, Pablo E., SNOW, Chuck, WOREL, Margaret (2011). *IBM FileNet P8 Platform and Architecture.* IBM. <http://www.redbooks.ibm.com/redbooks/pdfs/sg247667.pdf> [Accessed: 18/08/2014].

Notes

[1] BAIK Archive. <https://www.baik-archiv.at> [Accessed: 08/08/2014].
[2] Aufbau des Archivs (translated by authors). <https://www.baik-archiv.at/urka/img/aufbau-des-archivs_klein.gif> [Accessed: 08/08/2014].
[3] "A-Trust was founded in February 2000 and is an accredited Trust Centre in Austria issuing smartcard based qualified certificates for Austrian citizen used in e-Government. In March 2002 A-Trust has been accredited according to § 17 of the Austrian Signature Law by Telekom-Control-Kommission, the Austrian supervisory body." A-trust GMBH. <https://www.a-trust.at/ATrust/CompanyProfile.aspx> [Accessed: 08/08/2014].
[4] "HALMED is implementing one year IPA project Preparation for eCTD and Implementation of Digital Archive Information System". <http://www.halmed.hr/?ln=en&w=novosti&d=2014&id=1053&p=14> [Accessed: 06/08/2014].
[5] eCTD – electronic Common Technical Document is a standard for medicinal products' records.
[6] For more information see: ZHU, Wei-Dong et al., 2011.
[7] ROS – Records-enabled content Object Store.
[8] FPOS – File Plan Object Store.
[9] For more information on ROS and FPOS see: ZHU, Wei-Dong et al., 2009, pp. 39-40.
[10] Lit. Elektroninio archyvo informacinė systems. <http://eais-pub.archyvai.lt/eais/> [Accessed: 10/08/2014].
[11] EDAS – e-Servicing System for Insurers. <http://www.issa.int/details?uuid=245568ab-fb2d-4d1e-93a7-58e88a4bedf4> [Accessed: 27/08/2014].
[12] Lit. Södra.
[13] Ger. Bundesamt für Sicherheit in der Informationstechnik (BSI). <https://www.bsi.bund.de/EN/Home/home_node.html> [Accessed: 09/08/2014].
[14] Ger. Bundesarchivgesetz (BarchG). <http://www.bundesarchiv.de/bundesarchiv/rechtsgrundlagen/bundesarchivgesetz/index.html.en> [Accessed: 09/08/2014].
[15] Ger. Bundesnetzagentur. <http://www.bundesnetzagentur.de/EN/Home/home_node.html> [Accessed: 09/08/2014].