Erica Hellmer

Supervisor: Göran Samuelsson

Master thesis 15 credits

Spring 2015

Institution for Archives and Information Science

Mid Sweden University

Authenticity in Electronic Archives:

Securing Digital Records

Abstrakt

Den snabba tekniska utvecklingen har lett till ökad produktion av digitala handlingar/records och transaktioner. Inom e-arkivens domäner, har detta medfört ett ökat tryck på det tekniska området för att kunna garantera autenticitet av bevarad information över tid.

Tidsstämplingstekniker är ett sätt att säkerställa digital information för en särskild tidpunkt och kan användas för att garantera autenticitet av digital information över tid. Denna magisteruppsats undersöker funktionen tidsstämpling inom e-arkivens domäner genom en kvalitativ metod med fem semistrukturerade intervjuer av fem svenska organisationer. I undersökningen ingår också en fallstudie av det svenska innovationsföretaget Enigio Time AB för att ge en förståelse av tidsstämplingstekniker och av deras tjänst *time:stamp*.

Den teoretiska utgångspunkten i denna magisteruppsats är *records continuum model* med dels dess syn på kontinuitet i handlingar/records och dels dess belysande av skapelseögonblicket av handlingar/records. Den internationella standarden OAIS (Open Archival Information System) används för att hantera digital information och är, i denna uppsats, använd för att analysera hanteringen i dokumenthanteringssystem inom de fem organisationerna.

Resultaten visar att denna snabba tekniska utveckling har komplicerat dels hanteringen i att säkerställa att digitala handlingar/records inte kan bli manipulerade eller ändrade och dels för att fortsatt kunna garantera autenticitet i ett långtidsperspektiv.

Undersökningen visar även att organisationerna kan garantera att det bevarade materialet inte kommer att ändras när det väl är inne i arkivet men dess autenticitet, innan de kommer innanför arkivens trösklar, kan aldrig bekräftas. I och med att den moderna tiden producerar och bevarar enorma mängder data så bör detta starta med skapelsen av ett digitalt record tillsammans med bättre strukturerad metadatascheman som är säkrade i tid.

Nyckelord

E-arkiv, Autenticitet, Records Continuum Model, OAIS, tidsstämpling, time:stamp

Abstract

The rapid ongoing technical development has led to increased production of digital records and transactions. In the domain of electronic archives, this has put pressure on the technical area in order to guarantee authenticity of preserved information over time.

Time stamping techniques are one way to secure digital information at a certain point in time and can be employed to guarantee authenticity of digital information over time. This study examines the function of time stamping within the domain of electronic archives and is conducted with a qualitative method using semi-structured interviews on five Swedish organisations. The study is complemented with a case-study of the Swedish innovation company Enigio Time AB in order to gain understanding in time stamping techniques and their service *time:stamp*.

The theoretical framework in this study is the *records continuum model* with the view of the continuity in records and its illumination of the point of creation. The international standard OAIS (Open Archival Information System) is used by several electronic models to manage digital information and is, in this study, used to analyse the management in record keeping within the five organisations.

The conclusion is that this rapid technological development has complicated the management of secure digital records from manipulation and guaranteeing the authenticity in a long term perspective.

The study further shows that organisations may guarantee that records will not change once they are received into the archive but the authenticity of them, before they were delivered, can never be confirmed. Since the modern information era produces and preserves enormous amounts of data, this has to start with the creation of the digital record where better structured metadata schemes are secured in time.

Key words

Electronic archives, Authenticity, Records Continuum Model, OAIS, time stamping, time:stamp

Table of Contents

1. Introduction……..5
1.1 Purpose……6
1.2 Research questions……..6
1.3 Definition of terms……..7
2. Methodology……8
2.1 Interview procedure……...9
2.2 Interview analysis……..10
3. Theoretical framework……11
3.1 Records Continuum Model……...11
4. Related research on key concepts……14
4.1 Electronic archives, standards and regulations……14
4.2 Authenticity……...17
4.3 Records Continuum Model and importance of metadata……..18
4.4 OAIS……..19
4.5 Time stamping……...22
4.6 Time:stamp……22
5. Investigation……24
5.1 About the archives examined……24
5.1.1 The city archive of Västerås……24
5.1.2 The county council of Sörmland……..24
5.1.3 The Swedish Transport Administration……...25
5.1.4 The Swedish Tax Agency……25
5.1.5 The National Library of Sweden……..25
6. Results……..27
6.1 Electronic Archive……..27
6.2 Recordkeeping Model……....29
6.3 Authenticity / Traceability……33
6.4 Interview with Enigio Time AB……....42
7. Conclusion……...46
7.1 Further research……48

Acknowledgement
Reference list
Appendices
Appendix 1. Interview questions
Appendix 2. Interview questions Enigio

1. Introduction

The rapid development in the ongoing electronic era produces more digital records than ever before. But, even if, more records are produced and more metadata are being preserved, the reliability and accessibility of the records has decreased because of the difficulty in securing digital records from manipulation. In the long term perspective, this puts pressure on the technical domain of the electronic archives, since hard- and software systems have short life spans and specific file formats may not be readable over time. Within the electronic archives, there has to be methods which ensure that the preserved information is authentic and reliable even in the long term perspective.[1]

To guarantee provenance in a digital document by saving the original document is expensive and ineffective as it requires physical space and maintenance. Instead, a more cost-effective approach is to save unchanged patterns of binary data. However, the archival records' patterns change through transactions, and as a consequence, these patterns are not guarantees *per se*. The responsibility of the archives has long been to preserve the reliability of the archived document but not to *guarantee* the reliability of it.[2]

It has become more difficult to secure authenticity of electronic information because of increasing production and the media carrying the information. The rapid development in digital technology has led to increased business transactions in difference domains, e.g. the financial sector in online banking services and security trading services. These transactions are making it more difficult to secure information and to uncover alterations. This has led to a need to certify the creation and use of digital records and a way to do that is by using a *time stamping technique*. Time stamping is used to guarantee the time for creation, to identify actions performed on the digital record or to ensure authenticity of the digital information when it was received.[3]

Electronic records change and through the Records Continuum Model the continuity of records over time is elucidated, and the purpose of the records as memory and evidence is maintained.[4] Seen through the continuum, a record is never definite; and despite transactions, it is always usable and can be referred to because of its connection to actions there is a context connected to these actions.[5] To have a satisfying recordkeeping system that gives

[1] Ruusalepp (2005) p. 1
[2] Hänström (2007) p. 84-85
[3] Masashi (2001) p. 1
[4] Upward (1996)
[5] McKemmish (2001) p. 336

authenticity, reliability, and traceability to records, demands a pro-active holistic approach in accordance with the records continuum model.[6]

1.1 Purpose

The aim of this study is to contribute to the field of authentication of electronic records in electronic archives by examining the function of time stamping within the domain of electronic archiving (or archives). This study will analyse how organisations guarantee authenticity within their recordkeeping model on the basis of the records continuum model.

In addition, this thesis sheds light on a new perspective on securing authenticity regarding digital preservation in electronic archives. The digital era has simplified recordkeeping and lessened the physical storage space. However, soft- and hardware age rapidly and the importance of capturing metadata has been elucidated to make records usable over long time-periods. Archival science has developed along with the technical development and has contributed to new research developments e.g. InterPARES Trust. This multi-national research project aims to i.e. generate persistent digital memory with theoretical and methodological frameworks.[7] Since Nils Nilsson defined archival science in 1973, a definition that still holds today, the means and environment has changed fundamentally: contemporary archives deal with electronic records rather than physical records. According to Nilsson, archival science contributes to developing methods to make the archives function.[8] This thesis contributes to illuminate the importance of guaranteeing authenticity within modern recordkeeping in order to maintain the function and purpose of electronic archives.

1.2 Research questions

Is time stamping technology applicable in electronic archives; and does it, within the context of electronic archives, provide the additional authenticity to warrant implementation as such? In what field and in what kind of document or record within electronic archives, could a time stamping technique be useful?

[6] Sundberg, Wallin (2007) p. 37
[7] https://interparestrust.org/
[8] Nilsson (1973) p. 11

1.3 Definition of terms

The use of the word *document* in the current work is defined as "noted information or object which can be managed as an entity".[9]

Information becomes a record when it is part of a transaction.[10] There are numerous definitions of the term records, but here, the definition is that a record is bound to an object and an activity. A record is based on a context and a process and is therefore built on content, structure and a context.[11] There are differences between the English word records and the Swedish word *handling*. "Handling" is used within archival theory, but has a legal definition in which evidential value is based on the existence of the "handling" within the public domain rather than its inherent characteristic.[12] In this thesis, the word *record* is used since no corresponding translation of the word 'handling' exists.

Metadata is, in this thesis, defined as "data describing context, content and structure of records and their management through time".[13]

[9] SS-ISO 30300:2011 p. 8
[10] Bearman (1993) p. 20
[11] McKemmish (2005) p. 101-104
[12] Hänström (2007) p. 76-78
[13] ISO 15489-1 3.12 (2001)

2. Methodology

In this masters' thesis, I use a qualitative research method which can be defined as a method that investigates *how* something is structured or *how* to characterise something. This method is often used to investigate or interpret how a group of people experience their own world, their lifeworld. In other words, this method is used to comprehend an individual or a group of people's lifeworld.[14]

Examples of qualitative research methods are interviews and observations to interpret behaviour and from which certain theories can be developed.[15]

Moreover, this exam has a hermeneutic approach which means that a researcher explores a person's lifeworld and the meaning that person has in that perceived world. This lifeworld is the object when using a hermeneutic approach which gives an understanding of the perceived lifeworld, a way for a researcher to interpret the lifeworld.[16]

The aim, using this approach, is to discover the difference between perceived lifeworld's of different people which in this masters' thesis is the different investigation subjects and the archivists in these different domains.

In the hermeneutic approach, there is a concept that an investigated group of people from similar domains have similar conceptions. This corresponds to the positivistic theory, which argues that there are assumptions of what a domain consists of. The explicit difference between positivism and hermeneutics is the different realities that they examine; a hermeneutic approach examines how people perceive their world - not the world's actual nature. It is also holistic, i.e. a researcher needs to examine the whole concept of a perceived world and not interpret the concepts individually.[17]

Through a context, the researcher is given the opportunity to gain understanding of the overall picture i.e. the hermeneutic circle.[18] When interpreting parts of a study, these are connected to the interpretation of the entire study. Each part depends on the holistic perspective and the holistic perspective depends of the different parts.[19]

Anthony Giddens presented the concept *double hermeneutic* which is the interpretation of an already interpreted world of the participants. It is difficult to interpret a person's lifeworld without understanding how that person interprets the world, whether that

[14] Hartman (1998) p. 238-239
[15] Ibid p. 167
[16] Ibid p. 162
[17] Ibid p. 163-164
[18] Alvesson (2008) p. 193-194
[19] Gilje & Grimen (2007) p. 187

person's interpretation is correct or not. The researcher also needs to have certain presuppositions about the study to know the direction in what you are examining.[20]

In this master's thesis the double hermeneutic was used to analyse the data and to interpret the participant's interpretation.

2.1 Interview procedure

An interview is a good method when questions require complicated and reflective answers. Pickard has quoted Lincoln and Guba (1985,p.273) which summarizes the purpose with an interview:

> *"A major advantage of the interview is that it permits the respondent to move back and forth in time – to reconstruct the past, interpret the present, and predict the future, all without leaving a comfortable armchair"*.[21]

This study is performed using a qualitative method with semi-structured interviews on four organisations. One agency could not attend an interview and answered the questions via correspondence. Hence, the interviews were conducted both in person and via correspondence.

Data were collected from five organisations: One municipality; the city archive of Västerås (CAV), one county council; the county council of Sörmland (CCS), two Swedish government authorities; the Swedish Transport Administration (STA) and the Swedish Tax Agency (Tax Agency), and the National Library of Sweden (NLS). The reason for selecting these were to examine the situation of authenticity both in archives that have not established an electronic archive yet and within already established electronic archives. The city archive of Västerås did a pilot study in 2014 regarding the implementation of an electronic archive which will be introduced in 2015. Several of the examined archives have had established electronic archives for years, the Swedish Tax Agency 2006, the county council of Sörmland 2009, the National Library of Sweden and the Swedish transport administration 2012.

The interviews of the five organisations are complemented with a case-study of Enigio Time AB. The Swedish-owned innovation company Enigio was chosen for this study because of their considerable experience on building electronic archives. Their service time:stamp is a time stamp technique that has not yet been implemented as a part of an electronic archive solution. This case-study has facilitated an in-depth and up-to-date understanding of

[20] Gilje & Grimen (2007) p. 175-179
[21] Pickard (2013) p. 196

their service time:stamp. Their service differs from other time stamping techniques since it is not based on public keys in order to authenticate digital information.

Before the interviews took place the informants received an information letter describing the study and also examples of questions that would be used in the interview. Total length of each interview was 30 minutes -1,5 hours. Four interviews where recorded with the participant's informed consent.

The reason for collecting data with interviews is that interviews allow follow-up questions which often contribute to the research as opposed to questionnaires. In this study, four organisations were interviewed and one answered the questions via correspondence, the last being more difficult to analyse.

2.2 Interview analysis

A qualitative analysis of collected data occurs concurrently i.e. data informing analysis and vice versa. One way to analyse data is to use constant comparative analysis and what characterises this strategy is that the researcher does a comparison between one part of collected data with other parts to see the relationship between them. This was developed through the grounded theory by Glaser and Strauss, a method that is inductive, meaning that the conclusion is based on experiences. Using constant comparative analysis demands raw data and not data collected *a priori*.[22]

To be able to compare the archives, the questions for the interviews were divided into three categories; *Electronic archive, Recordkeeping model* and *Authenticity / Traceability.* With this approach, differences and similarities of the interviewed archives, views on definitions of electronic archives, regulations and standards, recordkeeping models, the guarantee of authenticity and their views of time stamping could be analysed.

[22] Pickard (2013) p. 267-269

3. Theoretical framework

The collected data in this study have been analysed from the records continuum model perspective. The model is a theoretical basis for recordkeeping in archives that is used in both archival science and in existing archives. The model is useful in managing electronic archives since its documents and records are considered to have a continued value and are always "in a process of becoming".[23]

3.1 Records Continuum Model

> *"Models are ways of seeing things. Their acceptance or otherwise in an area like records management depends upon how much contact they make with the practical consciousness of those who undertake tasks to be part of that activity."*[24]

The records continuum model has had success within recordkeeping since it illuminates the point of creation within different groups, conducts process analysis, systems analysis including structuring data about records with the purpose to facilitate initial task analysis of the evidential requirements for recordkeeping.[25]

The model was created by Frank Upward through Anthony Giddens structuration theory with a post–custodial perspective, showing the meaning of continuity in electronic recordkeeping processes.[26] The continuum was a reaction to the *life cycle model* which was based on an organic and linear approach i.e. a record is born, lives and dies or it is created, used and preserved or destructed.[27]

Frank Upward constructed structural principles for the continuum:

> *"1. A concept of records which is inclusive of records of continuing value (=archives), which stresses their uses for transactional, evidentiary and memory purposes, and which unifies approaches to archiving/recordkeeping whether records are kept for a slip second or a millennium.*
> *2. A focus on records as logical rather than physical entities, regardless of whether they are in paper or electronic form.*

[23] McKemmish (2001) p. 334
[24] Upward (2000) p. 1
[25] Upward (2000) p. 1-2
[26] Upward (1997)
[27] Eastwood et al. (2010) p. 140-141

3. *Institutionalization of the recordkeeping profession's role requires a particular emphasis on the need to integrate recordkeeping into business and societal processes and purposes."*[28]

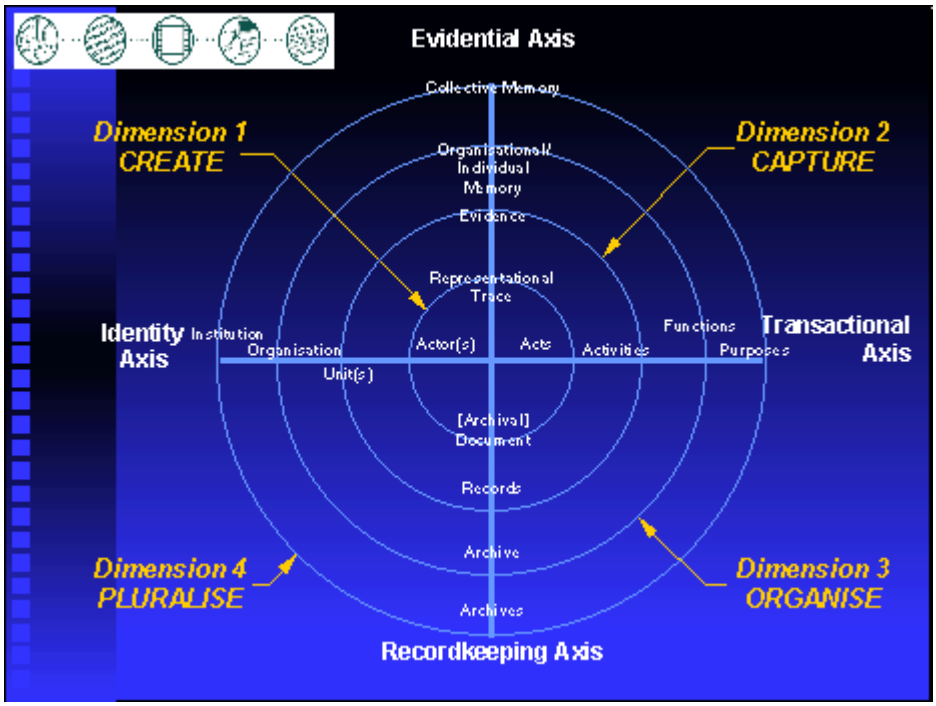

Figure 1. The records continuum model. Source: Upward 1996

Records continuum model used in recordkeeping systems can be used in all kinds of organisations as a supporting process[29] e.g. within businesses and organisations where it can elucidate the dissemination of recorded information. It also illuminates the events that result in the creation and capturing of records from an individual to a global level.[30]

The model consists of four dimensions: *create, capture, organize* and *pluralise*. In the first dimension, *create*, documents or records are created. These are now traces of actions being created. It is the second dimension, *capture*, that shows the relations between documents, and the traces of the transactions are now the evidence of the actions or events. In this dimension records or documents can be seen and understood by others in a company or organisation. Metadata is added to the document to make context. These first two dimensions are part of the initial process of information management in an organisation. In the third dimension, *organize*,

[28] Upward (1996)
[29] Sundberg, Wallin (2007) p. 38
[30] Troselius, Sundqvist (2012) p. 9

all records from the organisation are connected to a collective memory. The fourth dimension is *pluralise*, which is where the records gets the guarantee that they can be seen and reused outside the organisation. (See figure 1.)[31]

The difference between life cycle and continuum model is the time aspect where records with a life cycle approach travels from creation to accessibility to preservation, and in the continuum approach it is created, made accessible and is provided with information over time, a continuum.[32]

In this thesis, the continuum model will be seen as a supporting process within the recordkeeping process of today where metadata is added to records continuously.

[31] Reed (2005) p. 20-21
[32] Sjöberg (2014) p. 15

4. Related research on key concepts

In this chapter I discuss the related research of the key concepts relevant to this study. The purpose by showing different definitions is to see the overall picture in if and how authenticity is guaranteed within electronic archives that follows the OAIS model, by using metadata structures, time stamping or by having the record continuum model as a supporting process.

4.1 Electronic Archives, standards and regulations

The related research regarding electronic archives in Sweden is quite narrow and the reason for that might be that it has not been introduced in every municipality or government authority yet.[33] Several pilot studies concerning the implementation of electronic archives have been carried out but not much research within the area has been published. The different definitions of electronic archives have been presented in several master's theses.

In the master's thesis *Digital city archives. Using a transfer solution,* the author Annika Sjöberg analysed five cities' way of organising information in the domain of electronic archives and transfer archive solutions. A transfer archive solution is when organisations transfer information to an archive but retain the responsibility of the information. The theoretical approach was the records continuum model versus the life cycle model and the study showed that even though the city archives embraced the continuum approach, the archivists, in order to secure authenticity, considered the long term preservation as linear i.e. as the life cycle approach.[34]One important conclusion from Annika Sjöberg's study was that a sustainable digital preservation can be achieved by adding and capturing metadata early in the process.[35]

The occurrences of electronic archives in organisations are increasing which has challenged the idea of preservation. A digital record, versus a paper record, can be managed by several users at the same time with equivalent "performance". In the article *An Approach to the Preservation of Digital Records,* Heslop et al. developed a performance model for the National Archives of Australia's digital preservation program based on the international standard (AS) ISO 15489. The writers argue that the original record is not as important as capturing and recreating the performance of the record since every viewing creates new records of itself. As a result Heslop et al. observed that the preservation process needs to be the initial part of the process and require minimal effort for future researchers in order to learn new software applications. They argue that if a document can be rendered and recreated there is no essential

[33] Sjöberg (2014) p. 26
[34] Sjöberg (2014) p. 4-7
[35] Sjöberg (2014) p. 59

need for structuring data to the performance of a record. Instead, according to the writers, archivists need to determine, in the beginning, which elements of performance are essential to retain the meaning of the record.[36]

Electronic archives have seen a recent increase in both occurrence and popularity and were defined in 2008, by the National Archives of Sweden as follows:

> *"Electronic archives (e-arkiv) are constituted by authorities' electronic records despite their form, with belonging documentation that are considered archived according to the Swedish archival law Arkivförordningen and the regulations of the National Archives."*[37]

To promote efficient recordkeeping the National Archives of Sweden has certain regulations (RA-FS) regarding the management of electronic records. Regulation 2009:1 exists of advises on handling electronic records and chapter 6, *Informationssäkerhet* (information security), 1§ describes how electronic records shall be protected from "harm, manipulation, unauthorized access, and theft". This can be ensured through the standards SS-ISO/IEC 27001:2006 and SS-ISO/IEC 27001:2005 and LIS (Ledningssystem för informationssäkerhet). Many archives adhere to the requirement regarding preservation format of electronic documents that will be preserved in PDF/A in a long term perspective.[38]

The technical requirements regarding electronic records are described in regulation RA-FS 2009:2. Signed electronic documents follow IETF RFC 2315, which is a type of public key cryptography internet standard.[39]

A common standard within records management is ISO 15489. This involves guidelines "to ensure that adequate records are created, captured and managed". The characteristics of a record, according to the ISO 15489-1 is authenticity, reliability, integrity and useability.[40]

ISO/TR 15489-2 contains guidelines regarding the implementation of created or received records, in any format, by public or private organisations.[41]

---

[36] Heslop et al. (2002) p. 5-14
[37] The National Archives (2008) p. 15, translated by the author.
[38] RA-FS 2009:1
[39] RA-FS 2009:2
[40] ISO 15489-1:2001(E) p. 7
[41] ISO/TR 15489-2:2001(E) p. vi

Other standards useful within record management systems are the ISO 30300 series, *Information and documentation – Management systems for records*, which can be used by organisations to implement management systems as a reference point.[42]

To control the authenticity of records in electronic archives, metadata is added with the record i.e. more metadata connected to a record contributes to the long term preservation.[43] One metadata standard used within records management is ISO 23081 – *Records Management Process – Metadata for records*. It is a three-part standard and ISO 23081-1 complies with the requirements of the standard ISO 15489 and gives an understanding of the implementation of metadata. The purpose of metadata is to "identify, authenticate and contextualize records and the people, processes and systems that create, manage, maintain and use them and the policies that govern them." Chapter 8.5, *Metadata structures*, describes the importance of structuring metadata in schemes to make them meaningful.[44]

Another metadata-standard, for the purpose of useability and preservation over time is PREMIS, *PREservation Metadata: Implementation Strategies*. The PREMIS *Data Dictionary* contains guidelines for implementation for preservation metadata and is built on the OAIS model (ISO 14721).[45]

Governmental authorities and others who preserve electronic records must secure records from unauthorised access or manipulation within this preservation. This information security can be maintained with the SS-ISO/IEC 27001:2006 and SS-ISO/IEC 27002:2005. According to the National Archives, the OAIS model, ISO 14721:2002), may be used within electronic archives since it is in line with their definition of an electronic archive. MoReq (*MOdel REQuirements for the management of electronic records*), which is a requirement specification and can be used to set the needs for organisations within the electronic record management systems. MoReq is suitable for systems such as ERMS *(Electronic record management system*) and EDMS (*Electronic document management system*).[46]

[42] SS-ISO 30300:2011 p. 1-2
[43] Ruusalepp (2005) p. 9
[44] SS-ISO 23081-1:2006 p. 1-9
[45] PREMIS Data Dictionary for Preservation Metadata (2012) p. 2-3
[46]National Archive of Sweden (2008) p. 8-11

4.2 Authenticity

In the ongoing digital era where information is born digitally, and should therefore be preserved digitally, there needs to be a high standard for guaranteeing the authenticity and integrity of the digital information. One example of new technology that addresses the issue of guaranteeing authenticity and integrity is cryptographic technology.[47] Many electronic archives deal with the problem of guaranteeing authenticity that is preserving electronic records over time without losing information along the way.

There are several definitions of the concept of authenticity and a common definition is that an authentic record means that it is real and trustworthy and has not been manipulated. Authenticity is one of the four requirements that the standard ISO 15489 has on archival records. The definition in ISO 15489-1 is:

> *"An authentic record is one that can be proven to be what it purports to be, to have been created or sent by the person purported to have created it or sent it, and to have been created or sent at the time purported.*[48]

According to David Lynch in the article *Authenticity and Integrity in the Digital Environment: An Exploratory Analysis of the Central Role of Trust,* the definition of an authentic object is "that an object is indeed what it claims to be, or what it is claimed to be (by external metadata)."The claims of creation is the verification and can be given dates, an authorised identifier, or public keys, which are linked to the object. To validate the authenticity of information is to seek the provenance, i.e. information or metadata that points to the origin together with the chain of custody. The term provenance is closely related to authenticity since provenance may show descriptive metadata concerning relationships to other object.[49]

According to Luciana Duranti in the article *Archives as a place,* there is a close relationship between reliability and authenticity, however,

> *"*[…] *reliability is linked to creation, authenticity is linked to transmission and preservation." A document is therefore authentic if it has not been altered since it was transmitted. But when a document is delivered into an archive the*

[47] Lynch (2000) p. 33
[48] ISO 15489-1:2001 (E) 7.2.2 p.7
[49] Lynch (2000) p. 39-42

responsibility of preserving that authenticity requires that the authenticity cannot be questioned."[50]

In the article [*Authenticity in a digital world – long term preservation of records*], Kenneth Hänström argues that the guarantee of provenance in a digital document by saving the original document, is expensive and ineffective – simply because it requires physical space and maintenance. Instead, if binary data or unchanged patterns could be saved it would be a more cost-effective approach. However, these patterns are not an automatic guarantee since the patterns of records change through transactions.[51]

4.3 Records Continuum Model and the importance of metadata

Studies have shown, as a result of the increased digitalisation, that structuring metadata is becoming increasingly important. In the article *Item Level Control and Electronic Recordkeeping,* David Bearman argues that the evidential value of records will not be given by how records are organised on any storage device but through the contemporaneously created metadata connected to the record together over time. This, combined with records interactions, evidence of the use of records, and business processes, can be achieved.[52]

In a records continuum model perspective the records are viewed as a process rather than objects, with continuity in that process. Electronic records are transferred and used in different places at the same point in time so the actual position is of less importance in guaranteeing the authenticity of it. The authenticity is guaranteed by visualising relevant information with metadata throughout the existence of the record.[53]

In the study *A comparative case study on metadata schemes at Swedish governmental agencies,* the authors Nils Troselius and Anneli Sundqvist examined two agencies' recordkeeping practices and their implementations of metadata schemes. The two agencies were the Swedish Tax Agency and the Swedish Transport Administration. The study showed that the increased digitalisation requires more complex structures when capturing data to guarantee the authenticity, reliability and to guarantee access to captured data over time. The theoretical standpoint in the study was the records continuum model and the results show that using the model as a guideline in recordkeeping gives a holistic approach and can be used to

[50] Duranti (2007) p. 453-454
[51] Hänström (2007) p. 84-85
[52] Bearman (1996)
[53] Eastwood et al. (2010) p. 152-153

design tools for metadata schemes. In their study, they presented the purpose of metadata within recordkeeping management viewed from three different perspectives: a business perspective, a records management perspective and a use perspective. The records management perspective serves the purpose to capture essential attributes of records from the creation through time. This perspective reflects all four dimensions in the records continuum model.[54]

In the article *Recordkeeping and Information Architecture – A study of the Swedish Financial Sector,* the authors Sundberg and Wallin argue that in modern recordkeeping theory it is of great importance to be able to trace a course of event in a record, e.g. for organisations which have many customers to trace certain transactions or activities between the organisation and the customers. The study shows that the traceability could only be made internal in the different systems leaving the context unseen. The lack of a simple metadata model made it difficult to gain understanding of records from different sources and systems. To have a satisfying recordkeeping system, that gives authenticity, reliability and traceability to records over time, demands a pro-active holistic approach in accordance with the records continuum model. In a continuum perspective, recordkeeping goes beyond time with a continuity which is a necessity in current e-business processes over time.[55]

4.4 OAIS

The OAIS model, *Open Archival Information System*, is a reference model and was initially created by NASA to secure and store digital information for a designated community. It is an international standard, ISO 14721, and the word *Open* shows its openness regarding implementations and new development. It is a model that is often used within digital long term preservation. The model contains six entities *Ingest, Archival storage, Access, Data Management, Administration* and *Preservation planning*.[56] (See figure 2)

[54] Troselius, Sundqvist (2012) p. 10-18
[55] Sundberg, Wallin (2007) p. 37-41
[56] CCSDS 650.0-M-2 2012

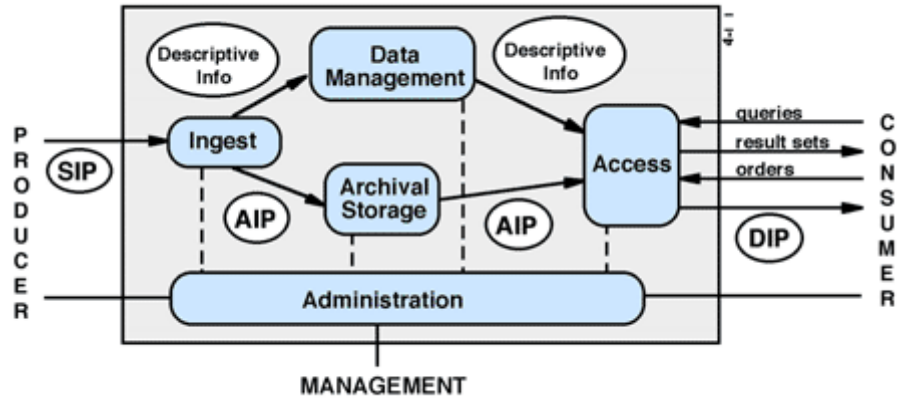

Figure 2. OAIS Functional Entities.

Source: http://nssdc.gsfc.nasa.gov/nssdc_news/dec00/oais.html

These entities cooperate with *Producer, Consumer* and *Management* which is outside of the archive. Producer provides the information to be preserved and delivers a *SIP (Submission Information Package)* to the archive. Management can be an archive in which one or many *AIP (Archival Information Package)* are created through the SIP that e.g. a governmental authority delivers. Consumer can be a governmental authority which requires to order an AIP that will be delivered as a *DIP (Dissemination Information Package). Ingest* receives the SIP into the archive in which it is checked and managed for storage. The storage of the package (AIP) occurs within *Archival storage*. *Access* is the entity where users, through services and functions, search and order a DIP. The data that is received into the archive is identified and described in *Data Management*. The entity *Administration* is maintenance of the archive and coordinates the activities of the other five entities. *Preservation planning* monitors e.g. new technology in order to maintain usability and readability of document and records for long term preservation.[57]

To secure and protect sensitive information within an archive the OAIS provides security services e.g. the *Data confidentiality service*. This ensures that no unauthorized individual can disclose information or make information available for others. The *Non-repudiation service* can be used to indicate different entities' participation in an information exchange. By using this service a participant cannot deny involvement and works in two ways. First, the recipient of data is provided with proof which makes the sender not able to deny sending. This proof contains the origin of the data so the content cannot be denied either.

[57] Brissman, Carlzon (2006) p. 20

Second, the data being sent is also provided with proof so that the recipient cannot deny receiving it.[58]

The OAIS model is used by many electronic archives to manage digital information. Annika Sjöberg describes in her master thesis, *E-arkivering hos stadsarkiv – Mellanarkivslösning i sitt sammanhang,* the close relationship between OAIS and, the records continuum model. Sjöberg presents Anneli Sundqvists' interpretation, in the dissertation *Search Processes, Use Behaviour and Archival Representational Systems,* of Sue McKemmishs' idea that records are "always in a process of becoming"[59] as:

> *"During a continuous process records could be amended, completed, reorganized, and used in various ways. According to the continuum theory there is no fixed point, where the records are finished and where the contextual relations are established once and for all. This means that not only the origin of creation, but also further use can be incorporated in the concept of records."*[60]

According to Sjöberg, this idea connects with the way that information is managed within the OAIS model i.e. information is organised in different ways depending on if it is received, preserved or extracted. This could be, according to Sjöberg, when a consumer receives a package (DIP) from the archive it will create new information, in this case, new usage will be a part of the information.[61]

An electronic archive that uses the OAIS model as a reference may need to implement other programs and systems, e.g. metadata schemes. Even though the model can be customized after the need of the users, other implementation might be of importance. The model is useful in managing information once it is received into the archive.

However, a drawback that has been pointed out regarding the model is the entity *Ingest*. In the article *OAIS as a reference model for repositories*, Julie Allinson argues that the model lacks a *pre-ingest* function since the model is perceived to provide insufficient guidance. Within this function, there would be a closer relationship and cooperation between the *Producer* and the archive regarding the package to be delivered to the archive, since this is the first step into the archive and to the remaining functions. According to Allinson, the model also lacks descriptions of how metadata schemes have been applied in each information delivery *SIP*, *AIP*

[58] CCSDS 650.0-M-2 2012
[59] McKemmish (2001) p. 334
[60] Sundqvist (2009) p. 46
[61] Sjöberg (2014) p. 24

and *DIP*. A solution to this could be to implement a metadata standard to visualise the metadata in order to facilitate the usability and to make it searchable for a specific user.[62]

4.5 Time Stamping

There has been a rapid development in digital technology which has led to increased business transactions in different domains, e.g. the financial sector in online banking services and security trading services. These business transactions are getting more intricate and numerous which is making it more difficult to uncover alterations and secure documents in these digital documents compared to paper-based documents. This increased digital technology developed a need to secure and prove digital data, and one way of doing that is by using the technique of *time stamping*. Time stamping has the capability of ensuring the integrity and authenticity of digital information for a long time period and is a method to guarantee that a certain information was created at a certain point in time.[63]

4.6 Time:stamp

A time stamping scheme is quite difficult to comprehend and research shows that the users (predominantly archival staff) need to have a good understanding of the subject and also accurately evaluate their security levels.[64]

A time stamp can be used in different ways. One way to secure the evidential value of a paper document is to mail the letter to oneself to get the evidential value from the time postmarked. However, this may not be the most reliable way to secure the evidential value when it can still be manipulated and is not usable on digital documents. Digital documents on the other hand are even easier to manipulate and that is why digital documents, in particular, need a time stamp. First, it is the data of the document that needs a time stamp not the medium. Second, it should not be possible to manipulate the actual time and date on the time stamp.[65]

Enigio Time AB is an innovation company based in Stockholm and has developed the service time:stamp in order to secure documents, records or images in time.[66] Many time stamping techniques is based on a trusted third party which is the temporal guarantee of the specific data.[67] The time:stamp service, which is the time stamping method this thesis focuses

[62] Allinson (2006) p.7-8
[63] Une (2001) p. 1
[64] Une (2001) p. 2
[65] Haber et al. (1999) p. 1
[66] https://enigio.com/#timestamp
[67] Une (2001) p. 4

on, is built on a linked scheme but does not require an identifier *ID,* or a trusted third party to validate the time stamp. This means that in a long term perspective the time:stamp would still be valid since the actual validation of the time stamp is a mathematical algorithm. The scheme is explained as below:

1. The time stamp requester is an entity which requires a time stamp for certain data *M.* This request is sent with a hash-value *H* to the data *M* to a time stamp issuer.

2. The issuer then makes a digital signature *S* which includes *M* and creates a time parameter which was the given time when the request was received *T.* The created time stamp *TS* corresponding to *M* includes *H, T,* and *S.*

3. When the time stamp *TS* is created it is sent to the time stamp requester.[68]

To provide evidence to the time stamp, Enigio is using a reverse publishing method. Instead of publishing in a newspaper or on Twitter – an event from a newspapers front page – such as a result from a sport event – is used as a reference in the created linked scheme. This will secure the document in time. The service is based on mathematics, giving the time stamped documents unique digital fingerprints.[69]

This time:stamp by Enigio, follows the International standard ISO/IEC 18014 series *Information technology – Security techniques – Time-stamping services*, which presents requirements for digital time stamping:

> *"- A time variant parameter shall be bound to the data in a non-forgeable way to provide evidence that the data existed prior to a certain point in time.- Data shall be provided in a way that it is not disclosed."*[70]

[68] Une (2001) p. 3
[69] Correspondence with Enigio
[70] ISO/IEC 18014-1 p. 4

5. Investigation

5.1 About the archives examined

The organisations in focus for this study are the city archive of Västerås, the county council of Sörmland, the Swedish Transport Administration, the Swedish Tax Agency and the National Library of Sweden. This chapter contains the background of each organisation examined based on policies and guidelines regarding information security within the electronic archive domain, as an introduction to the following interviews.

5.1.1 The city archive of Västerås

The city archive of Västerås (CAV) did a pilot study regarding an introduction of an electronic archive in 2014. This pilot study is a review of the vision of how to introduce the electronic archive with a current situation analysis. The coming electronic archive of Västerås will follow the line of the National Archives' regulations (RA-FS) concerning digital recordkeeping and the common specifications for government authorities (*förvaltningsgemensamma specifikationer, FGS*). CAV do not preserve in a long term perspective in the way the archive is constructed today but, this will be introduced with the electronic archive. The purpose of the electronic archive is secure digital storage, to increase the accessibility and to reduce the costs of the management of the archival process. The information that is created within the organisation will be delivered to the electronic archive and be used within the electronic administration of the city. The Swedish Association of Local Authorities and Regions (Swedish SKL) has procured a framework agreement with different suppliers within the electronic archive domain and CAV will employ a service from this framework agreement. It is not yet decided which service it will be.[71]

CAV has policies and requirements regarding information security which informs how the archive may store or send information in-house or externally within the local authority. In 2011 they established a policy concerning information security according to the standard ISO/IEC 27000 as a support to the continuous process of information security. In January 2012 the municipality of Västerås decided on requirements for information security where the criteria of information was to guarantee *accuracy, accessibility, confidentiality* and *Traceability.*[72]

[71] Förstudie e-arkiv Västerås Stad (2014) p. 1-11
[72] Riktlinje för informationssäkerhet (2012) Västerås stad p. 1-4

5.1.2 The county council of Sörmland

CCS constructed an electronic archive, R7 electronic archive, together with seven county councils in 2009 which has developed to nine county councils today. It is an archive with joint operation and management with the purpose to preserve electronic information in a long term perspective. The archive is built in-house and is independent of other technical solutions. It has both a transfer archive solution and a long term solution. In the transfer archive solution the different organisations still have access to the archived information and information that is no longer needed is preserved in their long term solution. The R7-archive receives electronic information which must be searchable in a long term perspective. The requirements for delivery packages describe what kind of data shall be delivered, how this data will be structured and which formats it will be delivered in. All information that will be preserved in a long term solution must be converted to a format suitable for long term preservation and the R7-archive has stated certain XML-requirements.[73]

5.1.3 The Swedish Transport Administration

The vision of the electronic archive of STA is to secure information over time, to increase the accessibility within the organisation and to be geographically independent. The archive, first introduced in 2012, is based on the OAIS model (ISO 14721:2003) and both the service and the product are procured. Legal security is guaranteed through origin information, integrity and presentation of the original content information.[74]

5.1.4 The Swedish Tax Agency

The electronic archive of the Swedish Tax Agency was introduced in 2006.

The Tax Agency has internal policies and documents limiting the dissemination of information concerning the infrastructure of its electronic archive and as such is not available to the public.

5.1.5 The National Library of Sweden

The electronic archive of NLS is based on the OAIS model (ISO 14721:2003) and uses the recommendations according to the Trusted Digital Repository which is based on the standard ISO 16363. The purpose of digital preservation within NLS is to secure access to the collections over time and to verity the authenticity of the information in the archive and the received information. In order to verify the authenticity of information, the metadata connected to the

[73] R7 e-arkiv (2013) p. 1-3

[74] Beskrivning av bakgrund och behov avseende e-arkiv. Bilaga 3.(2010) p. 1-5

information is preserved; and all archived information within NLS is part of the preservation plan. This requires that the ongoing systems follow new technical developments.[75]

[75] Policy för digitalt bevarande (2012) p. 1-5

6. Results

The interview questions were divided into three categories: *Electronic Archive, Recordkeeping model* and *Authenticity/ Traceability.* The presentation and analyses of the interviews are presented using these categories. The organisations will be mentioned with the following abbreviations: The city archive of Västerås as CAV, the county council of Sörmland as CCS, the Swedish Transport Administration as STA, the Swedish Tax Agency as Tax Agency and the National Library of Sweden as NLS.

6.1 Electronic Archive

- *What is your definition of an electronic archive?*

CAV defines an electronic archive as a business system that handles electronic records for a long term preservation or handles records as a transfer archive solution where the owner of the information is still the origin administrator. The forthcoming electronic archive will be introduced in two steps. In the first step it is the city archive itself who is the owner of the information; and in step two, the plan is that the origin administrator will be the information owner.

CCS defined an electronic archive as a traditional paper archive with the aim to preserve information in a secure long term perspective.

STA defined an electronic archive as a secure place which is fixed, rather than revisable, consistent and has a long term perspective in a system independent solution. From a technical point-of-view, the electronic archive should not depend on other solutions.

The definition of the electronic archive at the Tax Agency is that it is an archive for electronic records in an organised system for preservation, appraisal and long term information management.

NLS describes their electronic archive as an equivalent to the underground repositories that has been used for the physical documents preserved in an environment for long term preservation. All digital material in the electronic archive are to be preserved in the same long term way for all time. The archives need a well-functioning preservation plan with descriptions of how documents or records can survive over time, through migration and conversion, since archived documents today are preserved in formats that may not be readable in the future. The preservation plan is to secure this chain to confirm the authenticity in the received documents and present a good representation from the time it was first received. An

electronic archive should contribute with information to give authenticity and integrity to the archived documents.

- *In your electronic archive, do you follow the guidelines and regulations decided by the National Archives of Sweden?*

CAV and CCS follow regulations regarding format (e.g., PDF/A, XML and tiff). STA and the Tax Agency follow the regulations RA-FS 2009:1 and RA-FS 2009:2, additionally, the Tax Agency follows the Common Specifications for Government Agencies (förvaltningsgemensamma specifikationer; FGS).

The NLS works in cooperation with the National Archives regarding structure, metadata content and standards, to maintain structurally identical archives. NLS's electronic archive is controlled by the Legal Deposits Act (*e-pliktlagen; 2012:492*). This means that the NLS stores all electronic documents and records that are published under the Freedom of the Press Act, public authorities, and others who distribute or produce electronic material. These have to, according to this law, deliver material within three months from the first time it was published. Moreover, the material must be delivered in the same format as it was published in, even if another format may suit the library better in their long term preservation (e.g. PDF/A-1). The rationale behind this is to secure the representation of the information, at the cost of making the information less searchable.

- *Do you follow any ISO-standard in your recordkeeping?*

NLS, CCS, and CAV are not certified to any such standards as it is not required.

STA, although not certified, follow ISO 15489, ISO 23081, and to some extent, ISO 30300. They also look at other standards regarding construct information.

The Tax Agency follows ISO 14721:2003 (OAIS), SS-ISO/IEC 27001:2006, SS-ISO 15489, SS-ISO 30300:2011, SS-ISO 30301:2011 and metadata standard SS-ISO 23081.

Every organisation stated that an electronic archive is a system which aims to store records in a long term perspective. In addition, STA emphasises an intention to be independent from other solutions and NLS stresses the importance of maintaining the records' authenticity.

Every organisation follows the regulations of the National Archives of Sweden in some way. CAV and CCS follow the regulations regarding format, STA, NLS and the Tax Agency follow the regulations RA-FS 2009:1 and 2009:2, additionally the Tax Agency follows the FGS.

The standards that were used, explicitly or implicitly, were ISO 15489, ISO 23081, ISO 30300, ISO 14721 and ISO 27000.

6.2 Recordkeeping Model

- *Which recordkeeping model do you use/plan to use in your electronic archive?*

All of the questioned organisations use the OAIS in their electronic archives, although CAV does not follow it explicitly. In addition, STA uses pm3, a maintenance management model, regarding administration and development. The OAIS model has focused on the archive and the pm3 has focused on governance in the organisation concerning e.g. budget. NLS considers the model to be a fundamental model in order to make the archive self-sustaining.

- *Are there any advantages using the OAIS model?*

CAV stated that the OAIS is a good descriptive model, and STA emphasised benefits from a long term preservation perspective. CCS considers the model to be "forgiving" in the sense that it, e.g., allows implementations. The Tax Agency finds it beneficial that the model has distinct separations between its parts and may thus be used by other professionals than archivists.

NLS stated that the largest advantage is the idea of a self-sustaining archive, with a long term perspective, and that it is not dependent on other systems. Since 2012, when the legislation *e-pliktlagen* first was introduced, NLS had to handle much more metadata. To migrate such quantities of metadata would take years and then have to depend on other systems would be complicated since many systems today may work, at best, for ten years. After that new systems will be needed.

- *Are there any disadvantages using the OAIS model?*

CAV thinks of the model as a graphic descriptive model which gives meaning to the electronic archive and they could not see any disadvantages with the model today.

According to STA, the weakness with the model has been that it focuses on already captured information and the problem has primarily been to capture the information. This has been easier after introducing ISO 30300. To capture the information at all and to introduce it in an archival process or into a preservation process has been a problem, but once the information is captured the OAIS is a useful model. One of the advantages with the model is that it allows other systems or standards to cooperate. With other standards, connected to the model, it gives a holistic perspective.

CCS could not see any disadvantages with the model today but discussed the usability of the model in the future. Since the model allows other systems to integrate it can be a problem of compatibility between different systems.

Seen from the management perspective for the Tax Agency the model has been adjusted to fit the use within the organisation. Not every part of the model is useful for the organisation e.g. the Dissemination Information Package (DIP). When certain requests reach the archive, it is the connecting governmental authorities that are responsible for the delivery and the security around it. This function might be easier in a smaller organisation as a part of the electronic archive.

NLS described one weakness with the model concerning the interpretation of the implementation into the archive. Since the model can be interpreted in many different ways with different solutions, it is not always clear how to use the model in an optimal way. Moreover, NLS discussed that the problem of preserved packages has been underestimated e.g. if the information has been structured according to the packages in the storage solution it might be difficult to extract it in the future. The information might need to be structured in a more efficient way to facilitate future extraction. According to NLS, the packages needs to be self-sustaining, i.e. not depending on other external systems. The NLS have therefore chosen to represent the information in regular file systems with a hierarchical catalogue structure. The problem with this, when looking at external storage solution, is that even if external storage solutions may handle numerous amount of data, they are not often based on file systems, but instead based on object-based storage systems. When sending a file to an object-based storage system a key is sent back. This key has to exist somewhere and there needs to be information on what the key is for. There is a risk that the key could be connected to the storage solution e.g. connected to something specific. This is a disadvantage that NLS has experienced using the OAIS-model. According to the model, all specifications to every file format should be included in the packages. This has not yet been accomplished since the NLS is discussing the time aspect. What is self-sustaining over a long period of time? To solve this, NLS will continue to provide information to the archive and the OAIS- model might be the best model to achieve the goal in being self-sustaining.

- *Regarding your electronic archive, will you or have you procured a product or service?*

CAV will procure a service in their electronic archive but it is not yet decided which service it will be. It will be decided through a suborder.

CCS has developed their electronic archive in-house but has stipulated certain requirements to the suppliers who have done the work for them. CCS owns the product, R7 electronic archive.

Both the electronic archive and the product in STA is procured which they conduct and set requirements for. The control that the STA has over the archive is by using audit control over all archived information regulated by certain regulations of security.

The Tax Agency has procured a product for their electronic archive. In order to suit the archive, new functions have been developed in-house and have been added to supplement this product. All technical competence is managed within the Tax Agency.

The electronic archive of NLS has been built in-house even though they have been using open source products. They have procured a native XML-database, but if an equivalent open source product would be available NLS would shift to that. The product is used to handle metadata that is needed due to the fact that NLS does not convert all the files into a certain format. Instead, they write numerous preservation metadata and there are not many solutions that can handle this amount of metadata and still make is searchable, which this product is able to handle.

- If you have procured a service, how would that work in a long term perspective or if the company you bought the product from would disappear? How much control do they /will they have?

For CAV, which has not yet procured a service, this will be decided through the forthcoming deal, but they will always be the information owner. There is a rigorous specification regarding what the service provider may take part of, what information they may handle since the information will be preserved in their servers and in a cloud solution.

CCS, which has developed their own product, has both a transfer archive solution and a long term solution. It is the customers, the councils that presents the specification of requirements on the services CCS should provide.

STA have technical control of their electronic archive. STA is the information owner with total control and control over the administration of the archive. The technical access, which the procured service provides, is regulated with business agreements that limit what the service-provider can see, and limits what may or may not be done with the information in the archive. Much of the activities need approval from STA and to make sure that no information has been manipulated, STA do random samples and audits. If the company that they have procured the service from would disappear, STA would still own the information and

everything around it, and can easily move it to another solution. STA argue that in the era of electronic archives there will never be possible to park information in an archive similar to the paper-based archive, locked and sealed. The electronic archives will always need adjustments and the archived information will always be in need of migration of some sort. When the agreement with the procured service expires, STA must either continue with a new agreement with the same company or move to a new solution. To deliver to the National Archives is not an option for STA, since the information needs to be close to STA.

The Tax Agency procured a completely in-house product so the control is still within the organisation.

The license that NLS bought involved signing a two years contract of support. During this time, NLS receives updates on the product; and if the agreement is not renewed, NLS would still be able to use the product but without having new updates. Theoretically, NLS could have the product for some time after the agreement has expired, but in time, the operative systems would change and there would not be any support for the product. The electronic archive in the library would not be affected if the company disappeared since it does not demand any keys nor is the product exposed to a general public.

All of the questioned organisations use the OAIS in their electronic archives, although CAV does not follow it explicitly. In addition, STA uses pm3.

Each organisation stated advantages with the OAIS model and it was considered as a descriptive model, a beneficial model in a long term perspective, "forgiving" regarding implementations and a good model in the idea of a self-sustaining archive independent of other systems.

CAV, CCS, and the Tax Agency see no disadvantages with the OAIS model, although the Tax Agency finds that some of its parts, e.g. the Dissemination Information Package (DIP), is unfit for their organisation.

STA stated that the OAIS model focuses on already captured information, and a problem has primarily been to capture the information, which has been easier after they introduced ISO 30300.

NLS described the interpretation of the implementation into the archive as their weakness, because the model can be interpreted in many ways, with different solutions. It is not always clear how to use the model optimally. Furthermore, NLS stated that problems regarding the preserved packages have been underestimated, e.g. if information has been structured according to the packages in the storage solution, it might be difficult to extract it in the future.

Such information may need to be structured in a more efficient way to facilitate future extraction. As an example, NLS stated that object-based storage systems return a key for each file that is received. Such keys, along with information on what it is for, needs to be stored too. To circumvent this, and keep packages self-sustaining, NLS have therefore chosen to represent the information in a regular file system with a hierarchical catalogue structure.

CAV will procure a service in 2015 and STA has procured both service and product but owns the information and can transfer it to another solution. The Tax Agency and NLS has procured auxiliary products and the control is still within the organisation. For NLS, this involves two years of service and updates. Theoretically, after this period, NLS would still be able to use the product indefinitely. Their solution does not rely on keys, nor is it exposed to a public access, which means that their archive solution can function independently. CCS has developed their electronic archive, R7.

6.3 Authenticity / Traceability

- *In what way can you guarantee the authenticity in a document in your electronic archive today?*

When a document is delivered to CAV, either as an electronic document or a paper document, it will be checked continuously so the content of the document is preserved from the time it first arrived. It is not a control saying if a document is true or false, it is a matter of trust, and the documents are received to be continuously preserved with the guarantee that it will not disappear or be manipulated. The electronic documents that are received to CAV will be checked and compared to the original system e.g. by studying XML-schemes and comparing it with the original system to see if it is a complete presentation.

The authenticity of the received documents or records in the electronic archive in CCS is guaranteed by checksums. There are often big sets of data so it can only be checked with random samples and these are made by the councils themselves.

Once the information is captured into the archive of STA, the authenticity can be guaranteed by using formats like PDF/A and log files. But if a diarium closes, the authenticity of these records can never be confirmed. The only thing that STA can guarantee is that the records will not change once it is archived, but the authenticity of them before they were captured can always be questioned. Another example STA has experienced was when wrong maps were delivered with the original files from a management system into the electronic archive of STA. This causes problems since STA cannot control if the received information is

authentic or false and if someone looks at this in the future, the information archived will not have changed even though the information itself is incorrect.

The authenticity can be guaranteed with a certain control system within the electronic archive of the Tax Agency, but they cannot publish or share how the solution is constructed.

NLS has total control of their own digitalisation project and this control is guaranteed with a checksum approach. This is made early in the process and follows the digital documents over time. When the material is first received into the archive, it already has a checksum and through any operation these checksums will be verified. By counting the checksums any corruption will be detected, but along with the legislation *e-pliktlagen,* they receive such large amounts of metadata that no checksum is included with the received material. This is why NLS cannot with a hundred percent certainty, guarantee that the content is correct. Once it has been delivered into the archive, the checksum, from when it was received, is preserved and are connected to the material through time. This will guarantee the integrity of the material. To guarantee the authenticity of archived documents is difficult when operating big amounts of metadata e.g. if NLS receives a word document but convert it to a PDF information might be lost through this converting process which will mean that the checksum will not match the initiating checksum. In that case, NLS uses the metadata standard PREMIS. The original file will always be preserved and the changes being made will be presented in new packages including the checksum from the original word file and the converted PDF-file. A researcher can see what software that has been used or when the document was converted. Every event is preserved in this chain of preservation and by writing more preservation metadata the chain gets more trustworthy.

- *How can you see the traceability in a document in your electronic archive?*

There are only a few people who have access to the electronic documents in CAV and can retrieve these records. According to CAV, the original documents cannot be manipulated or changed in any way.

In the electronic archive of CCS, the received document will have the same presentation as it had when it first was received into the archive. If a received document is a PDF-file it will still have that presentation in the archive. A time stamp has not yet been used to secure the time of reception.

When information is captured and is received into the electronic archive of the STA, mapping is used which is equivalent to the FGS.[76] STA has experienced that converted documents may break even if it is only a few percent. When this occurs it is unclear whether the received information is reliable and/or whether the total delivery is damaged. If the reliability and the authenticity cannot be guaranteed, can the archive continue with that small margin of error? These are questions that STA has come across several times when the archive is attempting to guarantee the authenticity, the reliability and the secure traceability of the delivered information.

The Tax Agency uses log files and audit logs in order to secure the traceability of documents.

The traceability of documents are connected with the metadata standard PREMIS in the NLS. The library receives different versions of material e.g. the webpage of a daily newspaper is updated several times during one day and the library preserves every version of every article on that webpage.

- Is it important that the electronic archive has total control or can it be achieved with procured services?

STA which has procured both service and product argues that this high level of security could not have been achieved in-house. One of the reasons is that STA consider the supplier of the service to be experienced and competent and if the archive would have been handled within the organisation it would have just been an administrative interface. It would have been one of many things to handle in the organisation and may have caused negligence in the archival process.

The Tax Agency stores information in-house in two different physical places which have, according to them, a high level of security.

The NLS argued that total control can be achieved with procured services and be handled externally e.g. the part of the system that makes the packages of the material that will be delivered to the archive. The part where it might be more difficult with an external part, according to the NLS, is in the long term perspective and in the preservation planning. A long term license today might be for ten years and the migration process for this big amount of

[76]The FGS is produced by the National Archives and is a way to facilitate the action when information is moved from one system to another and consists of e.g. metadata requirements for information packages, metadata schemes and a specification in a form of a description in text. The National Archive of Sweden - Förvaltningsgemensamma specifikationer för e-arkiv och e-diarium. (FGS)

metadata might take several years to be carried out. This means that a big part of the deal will consist of migration out of one system into another. One way to solve this problem, in NLS' perspective, could be that a special authority could guarantee this type of long term preservation as a service.

- *Can you see any problems or threats to your long term preservation today?*

CAV has several long term storage repositories and the storage in the City Archive lies on small islands which is not sustainable in the long term perspective. The electronic deliveries to these repositories are not searchable and this kind of information is not suitable for access to the general public. Other long term storage options for CAV is the information management systems which are growing to enormous proportions and are really not sustainable in the long run. That is one of the reasons that CAV wishes to introduce an electronic archive. When discussing the forthcoming electronic archive, CAV argued that there are risks to preserve data where the general public is involved e.g. e-services. There has to be a totally different view in securing electronic documents than that for securing paper documents, documents that are born digitally must be preserved digitally.

One threat to the long term preservation process is, according to STA, the ongoing technical development and the digitalisation process. It can also be seen as an opportunity as well, but documents have been replaced with structural information which has to be handled and be represented as reliable evidence. STA argues that the information technology today is spelled with a lowercase i and a capital T since the technical development is an ambitious factor that in a rapid development produces new systems. This, according to STA, causes problems when archiving information in a long term perspective e.g. how to prove certain processes in new systems. The digitalisation era presents opportunities as well as many challenges.

The Tax Agency does not see any threats or problems with their long term preservation.

The NLS is in a situation where they now have to procure a new preservation solution. This is described as a challenge, as it is important to choose a solution that is accessible, that will provide access to future users, to a large amount of metadata that every day is received into the archive. This brings a lot of questions to the table: How will that be in a near future? It is expensive and could the government handle that cost? Do they have to start to make selections of what to save and how do you make these selections? The law *e-pliktslagen* is extensive and produces a substantial amount of data. The NLS focuses on the receiving part, primarily, but this also requires a continuous preservation plan over time. NLS describes that

the future is tomorrow, not in a hundred years, so the preservation plan is something that has to be in consideration in every part of the process. For example, if a migration from one storage solution to another needs to be done, other arrangements should be done at the same time since every part is time consuming. If certain formats are getting obsolete, they will go through a conversion process and the original file will contain descriptions regarding which software has been used and when it was done. It is a pragmatic approach since NLS need to have more control of the content and of formats when there is a risk that documents could be stored in an obscure format that may be difficult to use in the future.

- *Regarding time stamping, is this something that you have been discussing and do you think there is a need for that in your electronic archive now or in the future?*

Time stamping has not been an area of discussion within CAV. One of the reasons for this, was that they still thought of their archive as a paper-based archive where checksums worked as a guarantee of e.g. manipulation. They also use log files that track who has changed something and when. This guarantees the authenticity of the system. The general public will not have access to the electronic archive. They will have to go through certain control questions, a search interface, and the documents will be presented in a read-only presentation. One of the questions CAV had about time stamping was if an archive could time stamp an XML-file since the archive does not only archive and preserve whole documents or pictures but also e.g. charts. CAV also argued that in time, when the electronic archive is introduced and in process, the need for time stamping might be clearer.

To further secure the authenticity in electronic records should not, according to CCS, be an obstacle. The importance, in this issue, might be that the National Archives of Sweden should give clear guidelines which technique to use. Even if CCS has their own electronic archive and does not need to follow the requirements that the National Archives has set out, they play an important part in these matters.

Time stamping has been discussed within STA but has not been introduced into the archive yet. Much of the work within their organisation is administration and time stamping is of current interest. One example, made by STA, is within the construction industry when information about certain time events is of great importance e.g. when damages of buildings or even personal injury has occurred in connection with explosions on construction sites. To prove, the exact time when the explosions took place is of great importance which is why traceability needs to be secure. Another case, argued by STA, is the time for accidents or crises e.g. if an accident has occurred in connection with a railway. A report with a connected time stamp can

contribute to the following report of the course of event. A third case where a time stamp can be needed is when a contact line has collapsed. There are always debates in the media on how much time it took for STA to arrive at the scene and how much time it took to correct the mistake or accident. With a time stamp telling the exact time as evidence, this debate will never take place and it would not have to be corrected afterwards.

With a backup-policy in combination with time stamping in received documents, the Tax Agency considers this sufficient in their electronic archive.

The idea about time stamping was nothing that NLS had been discussing but sees the idea as an important part in securing digital documents. The time stamp NLS uses checksums when material is received into the archive, but this method cannot guarantee that a document existed at a certain point in time. The point in time that is shown to the National Archives is the time when it was received and is according to the legislation of *e-pliklagen*, that authorities have to deliver within three months from the publication date. Further time stamping within the National Archives may not be relevant but if the authorities that deliver to NLS would have this information connected to the material, it would be of interest rather than NLS itself using the technique. It is thus important that the authority or other suppliers add time stamping to the delivered package which would give the exact time of creation and when it was published and not just within three months. This could give e.g. a researcher exact time secured metadata.

- *What kind of documents could need a time stamp in your electronic archive?*

Examples where a time stamp could be of importance, according to the CSS, are different contracts, e.g. contract evidence or in systems of journals. Another area of interest could be financial transactions.

NLS have to preserve every document that has been published even if it was published only for a short time. This could be articles that consists of false information and had been withdrawn from a webpage. Even so, NLS has to, according to the legislation of *e-pliktlagen*, preserve that information. One example being when Statistics Sweden (Statistiska centralbyrån) released a report with false information concerning the Consumer Price Index (CPI) which made the Swedish Central Bank increase interest rates. According to NLS, this report was later replaced and if a time stamp would have been used it had been interesting to see for how long the false information was accessible. Other areas where time stamping could be of great importance could be for researchers who study published documents in the stock market. In that area, the exact point in time can be of great importance and with a time stamp

that time cannot be replaced or changed. Another example, discussed by NLS, when a time stamp could be of great importance is when a catastrophe of some sort has occurred in the headlines. The first information is often just a headline and as further information is being revealed to the journalists the article gets more detailed. At this point, it would be interesting with a time stamp since poor information can lead to a misleading course of events. For NLS, this kind of time stamping is of interest, has good information value, and could be handled in the same way as other metadata is handled in the received packages.

- *Concerning long term preservation, do you think there should be a management requirement regarding secure traceability?*

Both CAV and the Tax Agency consider that it should be a management requirement for secure traceability.

With information security in mind, STA would also like to see a management requirement for e.g. the many information technology systems or solutions since STA produces numerous amounts of information. The more important information regarding the management of the organisation needs more guidance to guarantee authenticity e.g. if this was regulated by the time of the delivery. It is a time consuming and onerous process to check the information when it is received into the archive of STA, but if the traceability was a part of the delivered material it would be less time consuming. If it occurred automatically in the process of e.g. a construction project when information was created and generated it would facilitate the archival process.

NLS and CCS do not see any need for a secure traceability of this kind.

The Tax Agency can guarantee their documents' and records' authenticity and traceability; however, they cannot publish or share how the solution is constructed. CAV can verify all documents and records by comparing the XML-schemata with the original presentations whereas STA does this using formats like PDF/A, log files and audit logs.

CCS can guarantee the authenticity and traceability of their documents and records by checksums and by using PDF/A. CCS's data-sets are often quite large, so the councils themselves verify authenticity by random sampling. For NLS, with the legislation *e-pliktlagen*, they receive large amounts of metadata and the NLS cannot guarantee, with a hundred per cent certainty, that the content is correct. Once the material has been delivered into the archive, the checksum is preserved and connected to the material through time which will guarantee the integrity of the material. To guarantee the authenticity of archived documents,

when operating large amounts of metadata, is difficult. As an example, if NLS receives a Word document and converts it into a PDF, information may be lost through this converting process. This means that the checksum will not match the original checksum. In that case, NLS uses the metadata standard PREMIS where the original file is always preserved, and any changes being made will be presented in new packages, including the checksum from the original Word file and the converted PDF-file. A researcher can see what software that has been used or when the document was converted. Every event is preserved in this chain of preservation and by writing more preservation metadata the chain gets more trustworthy.

Regarding the question if total control over the electronic archive could be achieved with a procured service, CAV would agree. STA argues that a procured product and service has a high level of security and this would not have been achieved in-house. In-house it would be seen as an administrative interface that could cause negligence since the archival process would not have been the only area of responsibility. According to the Tax Agency, a high level of security may be achieved by in-house storage solutions. In their favour, their electronic archive is produced and built in-house and they have gained experience since their archive, R7, has been running since 2009. NLS, which stores in-house, argues that total control can also be achieved with procured services. However, in the long term perspective, it might be more difficult since licences and business deals expire, and if it concerns large amounts of data it would be problematic in migration processes. The solution to this problem could be a special authority (only) dealing with how to guarantee long term preservation as a service.

Several of the interviewed archive representatives had detected threats or problems with their long term preservation. One of the reasons is the ongoing technical development over time. Yet another reason is the digitalisation process itself, which produces an abundance of information. CAV, which has storage repositories on different islands and in different management systems, did not think that their current systems were good solutions in a long term perspective. For this reason, they intend to introduce an electronic archive. The rapid technical development constantly produces new systems. According to STA, this causes problems when archiving information in the long term perspective when certain processes within new systems need to be verified. For NLS this long term preservation is a challenge especially after the introduction of the legislation *e-pliktlagen*. Since NLS receives a lot of data, it requires a continuous preservation plan over time in every part of the process, so that the information will not be stored in obscure formats and can still be readable in the future.

Regarding the need for time stamping, several organisations gave examples of when this could be of interest. One domain, presented by NLS, is in the research perspective e.g. a researcher studying published articles or headlines concerning different catastrophes where the first published headline may contain false or misleading information leading to incorrect interpretations of a current event. There are many cases where the second headline has made a total redirection when more detailed information has been discovered. To be able to see the whole chain of headlines with a time stamp telling when new information has reached the publisher, would be interesting when studying how misleading information can lead to preconceptions or vice versa. What would be interesting in fifty or hundred years is a difficult question to answer and NLS struggles with making information readable in the long term perspective.

Regarding which documents within each archive that might need a time stamp, several archives gave examples seen from their perspective. CCS discussed contract evidence, systems of journals and the area of financial transactions.

Three of the interviewed organisations had discussed time stamping; and according to this study, some could see a need for the technique to guarantee authenticity in documents. For STA, dealing with a lot of administration, time stamping is of current interest e.g. to prove, exactly in time, when explosions took place within the construction industry or to prove an exact point in time when accidents or crises have occurred. Today, this is only written by hand and no special marking of time is done; but with a part of the report in the creation, it would facilitate the process. A third example where a time stamp could be of great importance, according to the STA, is in the ongoing debate regarding when STA arrives at the scene and corrects the mishap when a contact line has collapsed. If time stamping would be used in this matter, it will serve as proof of an exact point in time. Several of the archives examined used checksums to secure digital documents. For NLS, which receives a great amount of information, time stamping is not a pressing need. The problem though, is that according to the legislation *e-pliktlagen*, the authorities have to deliver published information within three months from the publication date. This is not always adhered to; and but if the authority that deliver information to NLS would have a time stamp connected to the information, it would give a future researcher exact time secured metadata.

Three of the interviewed archives thought that there should be a management requirement regarding secure traceability.

6.4 Interview with Enigio Time AB

- *In what way does time:stamp differ from other time stamping techniques e.g. linked time stamping?*

The time:stamp is a variant of a linked scheme but does not handle time stamps individually they are instead linked together in a large tree. No trusted third party is needed and Enigio publishes on independent channels of publications to establish the information. A reversed publication is used as a complement. Their service consists of two different types of time stamping. The first type, positive time stamping, is the evidence of that something existed at a certain point in time. The second type, negative time stamping, is when third party information as a source of reference, is brought into the service to make the linked scheme a more secure and balanced chain in their linked time stamping. The source of reference could be a sport result or another event that will be registered as a hash. The purpose with this reference system is that Enigio connects a publication at a certain point in time with reference information from the same point in time. These kind of connections can be made in every second, once a day or hundred times per second or whatever seems convenient to make a precise decision of the time. The time:stamp is a combination of positive and negative time stamping techniques. Regarding an implementation of the time:stamp into electronic archives, the use of the positive time stamp might be more suitable rather than the negative.

If a document is given a time stamp, this will be sent back to the document so it can be seen, represented as time but it can also be a number which is based on the received information. With positive and negative time stamping there is no need for other partners other than the source of reference and users.

- *Many electronic archives uses the National Archives regulations regarding e.g. technical requirements for electronic records, RA-FS 2009:2. Digital signatures are frequently used and are based on IETF RFC 2315 PKCS #7 which is a public key cryptography internet standard. Since your service does not require any keys, could your service also comprehend digital signatures?*

Enigio argues for a solution without using a public key-based solution. They have not dealt with digital signatures but see possibilities in handling it. The problem with keys is that after a certificate has expired, the keys are not considered safe and the information which was created with the keys is no longer reliable. Another way to handle signatures is if Enigio acts as a third party where information could be presented in protocols to the client, the client identifies themselves through secure authentication of some sort and then makes the signature. This could

be something that Enigio could construct and could be based on their services. Even if Enigio were a third party, the information would be connected to the linked scheme with publication references and that would make the information reliable. It would be based on the same principle, in the same way as the time:stamp.

- Many electronic archives are based on the ISO-standard 14721 i.e. the OAIS-model. Did you have the model in mind when you created your service so that it might be implemented in archives that are using that model?

The model was not in consideration when the service first was created, but connections have been made *a posteriori*. Several employees at Enigio have been dealing with electronic archives in the domains of the county council; and those were based on the model, and they followed the development regarding electronic archives using the model.

- Studies have shown the importance of structured metadata schemes in order to guarantee the authenticity in documents and records. Seen from the records management perspective, on the basis of time stamping, the perspective can be to give records a value of evidence with the information regarding who created the record, when it was created and when it was received i.e. the traceability. Could time:stamp be seen as an auxiliary metadata scheme to guarantee authenticity and reliability in time?

Enigio argues that this might be executable. It may be introduced into an already existing metadata scheme or be an individual scheme; but either way, it would be of a different construction since there has to be a connection with a Time Stamping Authority (TSA) of some sort. It has to have a connection to a source of reference or a connection to something outside of the archive.

- If the time stamp requires a connection to something outside of the archive how does that work with migration or if new soft- or hardware is needed, how will that work in a long term perspective?

Enigio explains that it will be considered as a bookmark and is used to get a confirmation from the TSA. It is a loose connection not depended on certain version of a system. In the tree, which consists of time based information, Enigio publishes to external sources regularly and every time that is done, all information is used to create a publication from this point in time. By the next point in time, all new information in consort with all previous information (i.e. all information that exists at a certain point in time) is collected. This will form a chain of evidence

which the user/client can request from Enigio and can use the same chain. This information could be preserved into an archive, but Enigio claims that it is easier to do this through them since one chain of evidence only leads to one publication, and if that publication disappeared, the chain will lose its value. If this went through Enigio, which has redundancy and has thousands of publications, it would be easier and safer.

- *When it comes to digital recordkeeping archivists are more often dependant on other professional groups, e.g. information technology. What prior knowledge does an archivist need when using your service?*

When using the service, Enigio gives a presentation on how it is constructed and someone from the information technology domain will integrate it. It does not require any certain training or education and Enigio provides continuous feedback regarding the service.

- *Several of the organisations that I have been interviewing use checksums or log files in order to discover errors when receiving data. In what way could time:stamp be seen in this circumstance, can it simplify or replace current systems that intend to secure traceability?*

Enigio argues that time:stamp can intensify the traceability when using checksums. Checksums, in this case, are the keys for communicating and to be able to create secure time stamping. Checksums facilitate and the value it creates can facilitate the integration. There is not any secure time stamping within the system of log files, but there are backups of log files.

- *Electronic archives do not only preserve documents or pictures, it can also consist of XML-files or charts. Can you time stamp an XML-file or a chart?*

Both XML-files and charts can be time stamped according to Enigio. It can be one line in an XML-file or thousands of lines. The volume of it is not an issue; everything is done in the same way.

The time:stamp is a variant of a linked scheme with time stamps linked together in a large tree. This requires no trusted third party since Enigio publishes on independent channels of publications to establish the information. A reversed publication is used as a complement to this. The time:stamp is a combination of negative and positive time stamping techniques. Digital signatures are used frequently within electronic archives, but are often based on public key-based solutions and another way to solve this could be, according to Enigio, by using the same principle as the time:stamp.

Regarding the importance of structured metadata schemes, the time:stamp could be introduced to an already existing scheme or an individual scheme, but it has to be connected to a source of reference, a connection to something outside of the archive. The time:stamp is not dependent on certain versions of systems, and for an archivist, it does not require any certain education. Concerning whether time:stamp could simplify checksums or log files, Enigio argues that time:stamp can identify the traceability when using checksums. Checksums can facilitate the integration with the value it creates.

7. Conclusion

Rapid technological development produces more digital records than ever before, which has complicated the management in securing digital records from manipulation as well as complicated the management of information preservation in an authentic and reliable way in a long term perspective. To capture and store digital information with an imperishable authenticity value has proven to be a difficult task. The need to secure and prove digital information has led to the development of several time stamping techniques, which would be a way to ensure the integrity and authenticity of digital information.

The purpose of this study was to examine the function of time stamping in the domain of electronic archives in order to secure the authenticity of digital records. To do that, five organisations were selected, four of these were interviewed with a semi-structured interview technique and the fifth by correspondence. The interviews were complemented with a case-study of Enigio Time AB and their service time:stamp. The interview questions were divided into three categories; *Electronic archive*, *Recordkeeping model* and *Authenticity / Traceability* with the purpose of seeing differences and similarities of the interviewed organisations on definitions of electronic archives, regulations and standards, recordkeeping models, the guarantee of authenticity and their views of time stamping.

Each organisation stated advantages with the OAIS model; but even so, several organisations had experienced disadvantages using the model. The primary disadvantage is that the model focuses on already captured information, when the complexity lies within the capturing. This disadvantage might be solved with a *pre-ingest function*, described by Allinson, as presented earlier. Using a pre-ingest function could lead to a closer relationship between the *Producer* and the archive since it is of great importance that descriptions of metadata schemes is delivered in a correct way which will in turn reflect on the remaining functions.[77] Once the information is received into the archives the OAIS model is useful and beneficial.

An additional conclusion is that both STA and NLS see a problem with authenticity before the information is captured into the archives. The example STA had with archiving a closed diarium showed that the authenticity of these records from the closed diarium could not be guaranteed. Once it is captured into the electronic archive, authenticity can be guaranteed with checksums or log files but not the time for creation.

Another example of an authenticity problem was presented by STA when false information was delivered. The problem with this, stated by STA, is that if false information is

[77] Allinson (2006) p. 7-8

archived and someone looks at this in the future, the information archived has not changed even if the information itself is incorrect.

The organisations can guarantee that records will not change once they are archived but the authenticity of these records can never be confirmed. This was also experienced by the NLS, who, by *e-pliktlagen*, receives a large amount of metadata continuously and cannot, with a hundred percent certainty, guarantee that the content is correct.

Maybe the view of recordkeeping within archives has to change, being more aware of future use. As Hänström puts it, the responsibility for the archives has long been to preserve the reliability of the archived document but not to guarantee the reliability of it.[78]

This study shows that if information is to be reliable and authentic, in the long term perspective, it needs to be secured in time. It must ensure that no body, not even the creator, can manipulate the content. Since the modern information era produces and preserves enormous amount of data, this will become more important and should start with the creation.

This is discussed by Ruusalepp who argues that to solve the problem with digital preservation, the preservation has to start with the creation in order to facilitate the continuous preservation process, since it effects the quality of records over time.[79] This might be achieved with better structured metadata and when seeking provenance, Bearman argues, that the context of creation and use must be documented to understand records as evidence.[80]

Seeing from the continuum perspective the records are, according to Eastwood, viewed as a process rather than objects since records may be used in different places at the same point in time. The authenticity is guaranteed by visualising relevant information with metadata throughout the existence of the record.[81] This is supported by Sundberg and Wallin who discusses the importance of a pro-active holistic approach in order to be able to trace a course of event in a record e.g. between an organisation and its customers. By using a simple metadata scheme to capture relevant information, understanding of records from different sources and systems would be gained.[82]

For this study I mainly studied the first two continuum model dimensions, create and capture, since it is in those dimensions documents, or traces and evidence of records, are shown. It is also in these dimensions that a time:stamp could fit to additionally secure traceability and authenticity. By focusing on the purpose of time stamping, within records

[78] Hänström (2007) p. 84-85
[79] Ruusalepp (2005) p. 3-5
[80] Bearman (1996)
[81] Eastwood et al. (2010) p. 152-153
[82] Sundberg and Wallin (2007) p. 37-41

management, one perspective is to give a record a value of evidence which tells who has created the record, the time for when it was created and when it was received to the archive and thereby the traceability of it. This perspective reflects the first and second dimension in the records continuum model, create and capture. Time:stamp may be able to be used as an auxiliary metadata scheme which secures a temporal authenticity, reliability and trustworthiness.

7.1 Further research

The problem with guaranteeing authenticity within electronic archives wakens questions regarding preservation of digital signatures. It would be interesting to how see how these are preserved today since signatures are preserved, according to Ruusalepp, without issuing certificates. But are they still considered to be authentic without the certificates?

Digital signatures are often used to demonstrate the authenticity in digital documents. If the digital signature technique is combined with a time stamp technique, one could be able to identify the author and control if the content has been manipulated. According to Ruusalepp, who studied national archives around the world on behalf of the National Archive of Sweden, preserving digital signatures in a long term perspective is problematic. Signatures are based on certificates, (public keys) and proprietary software solutions, and have therefore a short life span, but are needed to authenticate the signature. This means that the digitally signed records may be preserved in the National Archives but without the issuing certificates. The authenticity will be guaranteed through the process that when the record has been received into the archive it will be the archives' responsibility that it will not be manipulated. The archives lack methods to preserve authentic digital signed records e.g. preservation and verification of certificates but use audit trails to seek actions in records. This is, according to Ruusalepp, insufficient and has created several initiatives to guarantee digital certificates in digital signatures e.g. The European Electronic Signature Standardization Initiative (EESSI) and Trusted Archival Services (TAS).[83]

Time:stamp is not based on public keys but has not been used for digital signatures. In this case, in order to provide authenticity to a signature, Enigio would have to act as a third party. The information would be reliable and authentic but would still be in need of a source of reference. It would still need a reference outside of the archive. How can authenticity in preserved digital signatures be guaranteed without public keys or a source of reference?

[83] Ruusalepp (2005) p. 11-12

Acknowledgements

I would like to thank Enigio time AB for giving me valuable feedback and interesting discussions. Thanks to valuable correspondence and meetings with Enigio I have learned the difference between different time stamping techniques and gained understanding about their service time:stamp.

I would also like to thank the City Archive of Västerås, the County Council of Sörmland, the Swedish Transport Administration, the National Library, and the Swedish Tax Agency for participating in this study.

Reference list

Allinson, J. (2006). *OAIS as a reference model for repositories*. Digital repositories Support, UKOLN, University of Bath.

Alvesson, M, Sköldberg, K. (2008). *Tolkning och reflektion – Vetenskapsfilosofi och kvalitativ metod.* Studentlitteratur, Danmark.

Bearman, D. (1993). *Record-Keeping Systems*. Archivaria 36.

Bearman, D. (1996). *Item Level Control and Electronic Recordkeeping*. Archives & Museums Informatics. Vol 10, no. 3.

Brissman, K, Carlzon, D. (2006). *OAIS I praktiken – En studie av OAIS-användning vid skapandet av ett digitalt arkiv*. Master thesis, Lunds Universitet.

CCSDS 650.0-M-2**.** (2012). *Reference Model for an Open Archival Information System (OAIS)*. Magenta Book. June 2012

City Archive of Västerås. (2012). *Riktlinje för informationssäkerhet*. Beslutad av Kommunstyrelsen 18 januari 2012. Västerås Stad. Retrieved from Västerås 2015-04-07

City Archive of Västerås. (2014). *Förstudie e-arkiv,* beslutad av styrgruppen den 8 april 2014. Västerås Stad. Retrieved 2015-03-09

Duranti, L. (2007). *Archives as a Place*. Archives & Social Studies: A Journal of Interdisciplinary Research. Vol. 1, no. 0.

Eastwood, T et al. (2010). *Currents of Archival Thinking*. Libraries Unlimited. Santa Barbara, California

Enigio Time AB - https://enigio.com/

Gilje, N, Grimen, H. (1999). *Samhällsvetenskapernas förutsättningar.* Daidos, Göteborg.

Haber et al. (1999). *How to Time-Stamp a Digital Document*. Journal of Cryptology, Vol. 3, No. 2.

Hartman, J. **(**1998). *Vetenskapligt tänkande – Från kunskapsteori till metodteori*, Studentlitteratur, Lund.

Heslop, H et al. (2002). *An Approach to the Preservation of Digital Records*. National Archives of Australia.

Hänström, K. (2007). *Autenticitet i en digital värld – Långsiktsbevarande av allmänna handlingar*, HUMAN IT 9.1(2007): 67-109.

International Organisation for Standardisation. (2001), *Documentation (Records Management) Part 1: General* (ISO 15489-1) Swedish Standards Institute: Stockholm. Retrieved via e-nav Mid Sweden University Library. 2015-03-10

International Organisation for Standardisation. (2001), *Documentation (Records Management) Part 2: Guidelines* (ISO 15489-2) Swedish Standards Institute: Stockholm. Retrieved via e-nav Mid Sweden University Library. 2015-03-10

International Organisation for Standardisation. (2011), *Information and documentation Management systems for records – Fundamentals and vocabulary* (ISO 30300) Swedish Standards Institute: Stockholm. Retrieved via e-nav Mid Sweden University Library. 2015-05-07

International Organisation for Standardisation. (2006), *Information and documentation – Records management processes*(ISO 23081-1) Swedish Standards Institute: Stockholm. Retrieved via e-nav Mid Sweden University Library. 2015-05-07

International Organisation for Standardisation. (2008), *Information technology – Security techniques - Time-stamping services Part 1: Framework* (ISO 18014-1) Swedish Standards Institute: Stockholm. Retrieved via e-nav Mid Sweden University Library. 2015-04-15

InterPARES trust - https://interparestrust.org/

Lynch, C. (2000).*Authenticity and Integrity in the Digital Environment: An Exploratory Analysis of the Central Role of Trust.* Part of Authenticity in a Digital Environment. Council on Library and Information Resources. Washington, D.C.

McKemmish, S. (2001). *Placing records continuum theory and practice,* Archival Science 1: 333-359.

McKemmish, S et al. (2005). *Archives: Recordkeeping in Society*, Wagga Wagga, N.S.W: Centre for Information Studies, Charles Sturt University, 2005.

National Archive of Sweden. (2008). *Rapport angående elektroniska arkiv (e-arkiv) bevarandeexemplar och system för bevarande – bevarande av elektroniska handlingar hos myndighet*. Retrieved 2015-03-26

National Archive of Sweden – Förvaltningsgemensamma specifikationer för e-arkiv och e-diarium. Retrieved 2015-04-20

National Archive of Sweden. (2009). *Riksarkivets föreskrifter och allmänna råd om elektroniska handlingar (upptagningar för automatiserad behandling) (RA-FS)* 2009:1.

National Archive of Sweden. (2009). *Riksarkivets föreskrifter och allmänna råd om tekniska krav för elektroniska handlingar (upptagningar för automatiserad behandling) (RA-FS)* 2009:2.

National Library of Sweden. (2012). *Policy för digitalt bevarande.* Retrieved from the National Library 2015-04-16

Nilsson, N (1973). *Arkivkunskap.* Studentlitteratur. Lund

Pickard, AJ. (2013). *Research Methods in Information*. Facet publishing. London.

PREMIS. (2012). *Data Dictionary for Preservation Metadata.* Version 2.2

R7. (2013). *R7 e-arkiv Systeminformation*.

Reed, B. (2005). *Reading the Records Continuum,* Archives and Manuscripts, Vol. 33, No 1.

Ruusalepp, R. (2005). *Digital Preservation in Archives: Overview of Current Research and Practices.*

Sjöberg, A. (2014). *E-arkivering – Mellanarkivslösning i sitt sammanhang*, master thesis, Uppsala University.

Sundberg, H, P, Wallin, P. (2007). *Recordkeeping and Information Architecture – A Study of the Swedish Financial Sector,* International Journal of Public Information Systems, Vol. 2007:1.

Sundqvist, A. (2009). *Search Processes, User Behaviour and Archival Representational Systems*. Dissertation. Mid Sweden University.

Swedish Transport Administration (2010), *Beskrivning av Bakgrund och behov avseende e-arkiv, bilaga 3*. 2010. Retrieved from the Swedish Transport Administration 2015-04-21

Troselius, N, Sundqvist, A. (2012). *A comparative case study on metadata schemes at Swedish governmental agencies*, Record Management Journal, Vol. 22 Iss 1 pp. 7-19.

Une, M. (2001). *The Security Evaluation of Time Stamping Schemes: The Present Situation and Studies*. IMES Discussion paper series, paper No. 2001-E-18

Upward, F. (1996). *Structuring the Records Continuum – Part One: Postcustodial principles and properties*. First published in Archives and Manuscripts, 24 (2).

Upward, F. (1997). *Structuring the Records Continuum, Part Two: Structuration Theory and Recordkeeping.* First published in Archives and Manuscripts, 25 (1).

Upward, F. (2000). *Modelling the continuum as paradigm shift in recordkeeping and archiving processes, and beyond – a personal reflection*. Record Management Journal, Vol 10 Iss: 3.

Appendices

Appendix 1. Interview questions

1. Electronic Archives

What is your definition of an electronic archive?

In your electronic archive, do you follow the guidelines and regulations decided by the National Archives of Sweden?

Do you follow any ISO-standard in your recordkeeping?

2. Recordkeeping model

Which recordkeeping model do you use/plan to use in your electronic archive?

Are there any advantages using the OAIS model?

Are there any disadvantages using the OAIS model?

Regarding your electronic archive, will you or have you procured a product or service?

If you have procured a service, how would that work in a long term perspective or if the company you bought the product from would disappear? How much control do they /will they have?

3. Authenticity / Traceability

In what way can you guarantee the authenticity in a document in your electronic archive today?

How can you see the traceability in a document in your electronic archive?

Is it important that the electronic archive has total control or can it be achieved with procured services?

Can you see any problems or threats to your long term preservation today?
Regarding time stamping, is this something that you have been discussing and do you think there is a need for that in your electronic archive now or in the future?

What kind of documents could need a time stamp in your electronic archive?

Concerning long term preservation, do you think there should be a management requirement regarding secure traceability?

Appendix 2. Interview question Enigio

- In what way does time:stamp differ from other time stamping techniques e.g. linked time stamping?

- Many electronic archives uses the National Archives regulations regarding e.g. technical requirements for electronic records, RA-FS 2009:2. Digital signatures are frequently used and are based on IETF RFC 2315 PKCS #7 which is a public key cryptography internet standard. Since your service does not require any keys, could your service also comprehend digital signatures?

- Many electronic archives are based on the ISO-standard 14721 i.e. the OAIS-model. Did you have the model in mind when you created your service so that it might be implemented in archives that are using that model?

- Studies have shown the importance of structured metadata schemes in order to guarantee the authenticity in documents and records. Seen from the records management perspective, on the basis of time stamping, the perspective can be to give records a value of evidence with the information regarding who created the record, when it was created and when it was received i.e. the traceability. Could time:stamp be seen as an auxiliary metadata scheme to guarantee authenticity and reliability in time?

- If the time stamp requires a connection to something outside of the archive how does that work with migration or if new soft- or hardware is needed, how will that work in a long term perspective?

- When it comes to digital recordkeeping archivists are more often dependant on other professional groups, e.g. information technology. What prior knowledge does an archivist need when using your service?

- Several of the organisations that I have been interviewing use checksums or log files in order to discover errors when receiving data. In what way could time:stamp be seen in this circumstance, can it simplify or replace current systems that intend to secure traceability?

- Electronic archives do not only preserve documents or pictures, it can also consist of XML-files or charts. Can you time stamp an XML-file or a chart?