Electronic Records and Digital Preservation Objects

A comparison

Margarita Korenkova
ACKNOWLEDGMENTS

I would like to thank all people that have helped me through this work and in that way made this thesis possible. I want to thank my supervisor Ph. Licentiate Hugo Quisbert at Luleå University of Technology for good advice, viewpoints and valuable support in bringing out this thesis. I want to thank Jan Aspenfjäll, Development Manager at the LDP Centre for showed interest, engagement and valuable contribution in this investigation. I also want to thank my colleagues, students of Information Systems Sciences at Luleå University of Technology for viewpoints and feedback to my work and finally, I want to thank my colleagues the Research Trainees at Luleå University of Technology for support, understanding and encouragement that helped me forward in that demanding and exciting period.
ABSTRACT

Today the rapid development of technology leads to information created in electronic form is growing exponentially at the same time, as hardware and software become obsolete fast. Information and long-term preservation is one of the most acute questions for all kind of preservation institutions because of large problems to deal with digital information. The solutions for long-term digital preservation and access that exist today are insufficient. One of the biggest challenges is to maintain high level of authenticity and reliability of electronic records. To find solutions for the long-term preservation problems, the LDP Project was started. The LDP Project gave rise to the LDP Centre; and its activities rely upon the outcomes of research and development from the LDP Project.

Today the LDP Centre has developed a model for structure of a digital preservation object, which is called Information Delivery Package. The aim is to deliver the archival material from an organization to an archive for long-term preservation. In the beginning of the development the Open Archival Information System (OAIS) was utilised, which later was relinquished in order to use some solutions of best practices. It means that its theoretical fundament is rather poor.

The purpose of the investigation was to point out possible deficiencies in the structure of digital preservation object, if any, developed by LDP Centre. This would be carried out using the Records Continuum Model as a “template” for records towards the developed structure of preservation object, analysing likenesses and differences. My efforts were made in order to show that capturing essential metadata, would make it easier to understand and preserve records. In the investigation, consideration was taken into electronic records at a conceptual level, regardless of format, medium, technical or legal aspects.

The analysis of the empirical data shows that the structure of preservation object developed by LDP Centre contains all concepts within the Records Continuum Model at three levels, namely Create, Capture and Organize. The concepts embracing the fourth level, namely Pluralise in the Records Continuum Model are not presented in the developed structure of preservation object. When concerning the specific case of the Swedish National Archives, the managerial group, organizational structure and legislation put a barrier to apply the idea about alliance building and consequently the idea of time/space distanciation. Hence, the comment on whether the structure of preservation object has deficiencies is unsuitable so far. This because, the LDP Centre followed Swedish legislation and norms of the Swedish National Archives, when developing the structure of preservation object.
### TABLE OF CONTENTS

*Introduction* .......................................................................................................................... 1
  Archive ....................................................................................................................................... 1
  The LDP Project ......................................................................................................................... 2

*Purpose* .................................................................................................................................... 4

*Scope* ........................................................................................................................................ 5

*Theory* ....................................................................................................................................... 6
  The Concept of Records ............................................................................................................. 6
  The Records Continuum Model .................................................................................................... 9
  The OAIS Reference Model ......................................................................................................... 12
    The OAIS Concept of Information ............................................................................................. 14
    Information Package .................................................................................................................. 15
  The Preservation Object ............................................................................................................. 16

*Method* ...................................................................................................................................... 17
  Research Approach ................................................................................................................... 17
  Research Purpose ...................................................................................................................... 17
  Research Strategy ..................................................................................................................... 18
    Qualitative-focused Research .................................................................................................. 18

*Information Collection* .......................................................................................................... 19
  Analysis Approach .................................................................................................................... 21
    Credibility .............................................................................................................................. 21
    Transferability ........................................................................................................................ 22
    Dependability ........................................................................................................................ 22
    Confirmability ........................................................................................................................ 22

*Empirical Data* ......................................................................................................................... 23
  The Preservation Object ............................................................................................................. 23
    Creating document .................................................................................................................. 25
    Capturing records ................................................................................................................... 26
    Organizing Records in the Archive .......................................................................................... 27
    Archives ................................................................................................................................... 28

*Data Analysis* ...................................................................................................................... 29
    Document ............................................................................................................................... 29
    Actor(s) .................................................................................................................................. 29
    Transaction ............................................................................................................................. 29
    Trace ...................................................................................................................................... 29
    Record .................................................................................................................................... 30
    Activity .................................................................................................................................. 30
    Unit(s) .................................................................................................................................... 30
Conclusions and Discussion .......................................................... 34
Method Discussion ........................................................................... 36
Proposal to Further Research ......................................................... 38
References ....................................................................................... 39
Interview guide ................................................................................ 41
CHAPTER 1

Introduction

During the last years, the development of digital technologies has been very fast and for the most unpredictable. This not only affected the way people are using the new artefacts (computers, cellular phones etc.) to exchange and disseminate information, but also organizations and society as a whole. (Runardotter et al., 2006)

Organizations are using digital technologies to use and manage information in their working processes. As a consequence the amount of information (records) created in electronic form is growing exponentially. At the same time, the development of technologies has created unforeseen problems, such as the obsolescence of hardware and software. There is no information system capable to retain and make the information accessible for more than 30 years. Constant upgrading of operative systems and software applications in one hand and new storage technologies on other hand, are leading to the very difficult access to information. File formats are experiencing the same problem because what today is standard for some application shifts or disappears within a decade. (Lim Siew Lin et al., 2003)

The solutions that are today for preservations and foremost to make digital material accessible over long time are insufficient. Archival inquiry “Archives for Everyone – Now and in the Future” (SOU2002:78) called attention to these great shortcomings and argued that models, methods and platform-independent software for preservation of the digital material are omitted largely. (Marklund, 2002)

Archive

Information and long-term preservation is one of the most acute questions for all kind of preservation institutions like archives, libraries, local authorities and also for society as a whole (Runardotter et al., 2006).

Archives are documenting and preserving legal aspects of human social activities and national memory. According to Swedish law, archives are part of the Swedish national cultural heritage (Marklund, 2002). According to Runardotter et al., (2006) digital information has to be available for the future society by legal, historical and democratic reasons. Archivists have as their mission the preservation of records and its continuing value and the providing access to those records (ICA, 2005).

According to Borglund (2006), preservation of electronic records is problematic because it is difficult to maintain the authenticity and reliability of
electronic records. Further Borglund states that authenticity and reliability are requirements necessary to fulfil, in order to obtain electronic records of high quality, which can be used in daily work processes as evidence over the transaction of its creation.

Electronic records unlike the paper records are sensitive to inattentive and undetected changes within content, context and format if they do not hold in the safety and under defined and reviewed procedures. These procedures must guarantee that the electronic records are authentic and exact same representations of the transactions and ensure that electronic records have been secured from changes. (Lim Siew Lin et al., 2003)

So far, theory is more advanced than practice. In practice, archives have been used methods developed since 1970-1980. These methods cause losing of functionalities of the records when archiving. Archival science has not developed as rapid as technology, which has been used to create electronic records. (Ruusalepp, 2005)

Electronically created records have put archival science on test. Electronic records have being different from traditional records, leading to a paradigm shift in archival science. (Borglund, 2005)

Today it is stated that archival authorities in the world have enormous problems to deal with digital information. Hence, society and authorities must take a challenge of finding some sustainable solutions to ensure that digital records of today become accessible and understandable in the future. Otherwise, digital material and likewise one part of society’s future cultural heritage and history are in danger of disappearing forever. Consequently, legal and democratic demands cannot be fulfilled. (Marklund, 2002)

The LDP Project

Today the problems have been identified and acknowledged. A number of research projects and institutions over the world have responded to the challenge of finding solutions of the described problems. These projects foremost are focusing on research, the development of preservation activities and education. (Runardotter et al., 2006)

According to Borglund (2006), research on digital preservation, concerning the preservation of electronic records can be divided in two different perspective or traditions:

- **Technical approach of electronic records** – that is problems concerning how electronic records should be understandable throughout its existence. Computers, information systems, file
formats, storage mediums, operative systems are all examples of technical artefacts in continuing change and development, making preservation of records difficult.

- **Methodological and theoretical level of electronic records** – such research is the Records Continuum Model, which constitutes the foundation in the Australian school of recordkeeping.

In Sweden the Swedish National Archives, Luleå University of Technology, the Municipality of Boden and the National Social Security Board have started a research and development project, the Long-term Digital Preservation Project aiming to develop methods and models for long-term digital preservation. The LDP Project gave rise to a centre of competence within long-term digital preservation field, named the LDP Centre. This centre was established by the Swedish National Archives, Luleå University of Technology, the Municipality of Boden, the Royal Library and the Swedish Archive of Recorded Sound and Moving Images. The activities of the LDP Centre rely upon the outcomes of research and development from the LDP Project. The centre is continuing the activities of networking and information dissemination initiated by the LDP Project. (LDP, 2006)

Research and development work is directed on how digital material can be preserved and be available over long time with a high degree of reliability and trustworthiness. It means that the material has to be available over long time to citizens, authorities, companies, other organizations and researchers; and information have to be presented for these interested with content, structure and context intact. It is necessary for the interested to rely that information is correct and to be able understand and interpret it from the situation where it was created once. (LDP, 2006)

One of the major activities for the LDP Centre is systems development. The aim for this activity is to develop methods and techniques for digital archiving in the public service archiving sector. Some efforts have been archived in how information is “prepared” and transferred to archive authorities. The LDP Centre is working with the development of context and structures of preservation objects and focusing on development of a detached digital preservation object that is independent of technology. In this way, the Centre expects to reduce the load on technology itself and the preservation object can be preserved without consideration to software or hardware which created or kept the original piece of information. This is viewed as the main key to a possible solution of long-term digital preservation problems.

Today the LDP Centre has developed a model for structure of a preservation object, which is called **Information Delivery Package**. This structure is intended to be used by archive creating authorities in order to create records of organizations containing produced content and context and which later could
be delivered to an archive for long-term preservation. As starting point in the
development, they have used a conceptual model for Archive Information
Systems, Reference Model for an Open Archival Information System (OAINS)
that was accepted as an international standard in 2003: ISO 14721:2003.
Later, during development time the Centre have relinquished from the model
and continued using some solutions of best practices, even though some
concepts and terminology of the OAIS Reference Model is still in use.
According to LDP Centre, the developed structure of preservation object was
not tested in an extensive scale, because of time limit. Hence, the structure
might be not complete. It means that its theoretical fundament is rather poor,
which supports the statement done by Ruusalep (2005) that:

“during the last twenty five years despite that some research was
done within the problem area, the practical solutions where of ad
hoc type, where time for planning has been short”. (Ruusalep,
2005)

The above statement motivates to investigate which concepts are comprised
in the developed structure of preservation object and compare those with
theoretical concepts from archival theory.

My interest and proposal to this investigation grow up during the literature
study. The main inspiration and the idea I have got from the licentiate thesis
“A Framework for the Development of Archival Information Systems” carried out by
Quisbert (2006). Later, the investigation proposal was discussed with LDP
Centre and the supervisor. Both parties approved my proposal interesting on
the basis of the research perspective.

**Purpose**

I have chose to not have any research question, but start the research from a
research purpose, because the research purpose is covering wider scope in the
problem area than en research question and in that way gives an increased
value for research.

The purpose of the investigation is to point out possible deficiencies in the
structure of preservation object, if any, developed by LDP Centre. This will
be carried out using the Records Continuum Model as a “template” for
records towards the developed structure of preservation object, analysing
likenesses and differences. My efforts are made in order to show that
capturing essential metadata, would make it easier to understand and preserve
records.

The thesis is intended foremost to staff at the LDP Centre working with
development of the digital preservation object and to other interested within
the framework for the research area. The findings from this thesis can be
taken into consideration when developing information system that manages electronic records.

**Scope**

In this investigation, I only take into consideration preservation objects in digital form and electronic records at a conceptual level, regardless of format, medium, technical or legal aspects. I limit my work to illustrate the concepts that embrace existing elements within the structure of developed preservation object. The concept of structure in this research is referring to a general structure concept that describes the context of records in contrast to a general structure concerning visual elements like layout, links etc. Since I use the Records Continuum Model as “template” for comparison, I do not take into consideration other concepts that are not handled by this model, but that can be part of the structure of the preservation object.
CHAPTER 2

Theory

This chapter is a presentation of a theoretical base to the investigation. This begins with explanation of the notion “record”. Further, the Records Continuum Model and the OAIS Reference Model explains. The Preservation Object ends the chapter.

The Concept of Records

In the research two definitions of records has been adopted. The first one was defined by the Australian Standard (ISO 15489), which is one of the components in the Australian school of recordkeeping and was based on the Australian Standard AS 4390-1996: Records Management.

"recorded information , in any form, including data in computer systems, created or received and maintained by an organization or person in the transaction of business or the conduct of affairs and kept as evidence of such activity" (AS 4390.1 1996:6)

The second was defined by International Council on Archives Committee (ICA). This definition is different to the definition of the Australian Standard. This definition gives a more detailed explanation about comprising fundamental characteristics of a record.

"recorded information produced or received in the initiation, conduct or completion of an institutional or individual activity and that comprises content, context and structure sufficient to provide evidence of the activity regardless of the form or medium" (ICA, 2005)

ICA Committee was argued, “…a record has to be related to an activity carried out by an organization or an individual, and this activity and the function it supports determine the provenance of the record, and the record is evidence of that activity”. Further, they insist that all organizations need to capture and maintain records of their business functions to meet business needs and legal requirements. (ICA, 2005)

To support business functions and to provide evidence, a record most possess certain characteristics (ibid.):

- **Authenticity** is defined as “persistence over time of the original characteristics of the record with respect to context, structure and content”. It means that a record is what it purports to be; and
- **Reliability** is defined as the ability of a record “to serve as reliable evidence”, referring to the authority and trustworthiness of records as evidence.

A record can only be evidence of a transaction if a record is reliable and authentic. Authenticity and reliability follow each other. If one of them is low, the other is often low as well. (Borglund, 2006)

The International Standard on Records Management ISO 15489-1 have also ascertained the significance of these characteristics and added two closely related characteristics:

- **Integrity**, referring to the fact that a record is complete and unaltered; and

- **Usability**, defining the ability to locate, retrieve, present, and interpret a record.

Records that possess these characteristics will have sufficient content, structure and context to provide a complete account of the activities and transactions to which the records relate, and they will reflect decisions, actions, and responsibilities. If such records are maintained in an accessible, understandable and usable manner, they will be able to support business needs and to be used for accountability purposes over time. (ICA, 2005)

Records are always transactional bounded and should be able to be used as evidence. The evidential value of a record can only exist if the content, structure, and context are preserved. The context is the link between different records that belong together and to the process where the record was created. The relation of records to transactions is both what makes records different from information and enables the evidential functionality of records. (Borglund, 2006)

**Content**
Content is the subject matter of the document (Public Record Office, 1999) and referred to intellectual substance contained in information object (Lazinger, 2001).

**Structure**
The concept of structure “is related to how the record is recorded, which includes the use of symbols, layout, format, medium, etc.” There is a distinction between physical and logical structure: whereas the physical structure of a traditional record is apparent to user, it is not in the case of electronic records. The physical structure of an electronic record is variable and dependent on
hardware and software; its logical structure (i.e. the relationship between its component parts) renders it intelligible. (ICA, 2005)

This supports a statement by Borglund (2006) that “there is no difference in expected functions between records and electronic records, but there are differences in structure and form”. The majority of traditional records are paper-based. A traditional record is a physical entity, often a document. The electronic record is more of a logical entity, of which integral parts can be managed at different places within an information system, or even in different information systems.

**Context and Metadata**

There are three aspects of the context of a record, but these are not necessarily complete:

- Contextual information – contained in the record (for instance, the signature).
- Relationship between a record and other records in the fond.
- Activity in which the record was created.

Contextual information links records to the administrative and functional environment (activities, processes) from which they were created, and to other records. (ICA, 2005)

The contextual information also allows for the authenticity, reliability and integrity of the records to be proved. This is of particular importance for electronic records. The business processes and functions and the recordkeeping system of a record-creating organisation are part of context of its records. Context can be preserved by internal elements of records (e.g. attached documents, annexed information, links, numbers and reference codes) or by external elements (e.g. metadata). (ICA, 2005)

Metadata are a crucial part of contextual information. Metadata are defined as “data about data”, whereas the International Standard on Records Management explains them specifically in the records management perspective as “data describing context, content and structure of records and their management through time.” (ICA, 2005)

For electronic records, the concept embraces all kinds of information, which are needed to make the record understandable and usable (e.g. system documentation). Metadata can serve different purposes, such as retrieval, usability, authenticity, reliability, maintenance, preservation and appraisal. For electronic records, metadata are particularly important, because they establish “the relationship between a record and its functional and administrative context. Thus, electronic records are heavily dependent not only on a well-documented
administrative context, but on metadata describing how the information is recorded.” (ICA, 2005)

The Provenance Principle
The provenance principle states that the original order or context should be preserved and explained. The provenance stands for the origin of records and implies that these matters should be clear when records are being archived. According to authors the provenance principle state that it is information that is to be preserved since the context matters because data without context is not possible to interpret and can never transform into information. (Runardotter et al., 2006)

The Records Continuum Model
The Records Continuum Model theory was conceived by the Records Continuum Research Group, at Monash University in Australia. The Records Continuum Model has as its fundament the recordkeeping theory and the structuration theory of Anthony Giddens. The model is object oriented. (Upward, 1996, 1997, 2000)

The Record Continuum Model is a theoretical instrument for analysis of records in recordkeeping systems. The model explains different elements and the phenomena that embrace records. The model handles records from the continuum perspective, which means that (the model expresses the continuing value of records) records never die if are in use and useful. The model conceives records as logical entities rather than physical entities and takes into consideration that records can be both on paper and in digital form. The model is based on the Gidden’s Structuration Theory that for many is philosophical and can be considered for a new thinking. (Upward, 1996, 1997)

The Records Continuum Model is based on four axes, what is called Continuum: Identity, Evidence, Transaction and Recordkeeping Containers and four dimensions: Create, Capture, Organize and Pluralise (cf. Figure 1). The axes take its foundation thinking in archival science. Every axis presents four coordinates that can be linked dimensionally to each other. (Upward, 1996, 1997, 2000)

The model highlights the idea about establishing the custody of records. This idea, according to Upward (1997) is expressed in space/time distanciation. Since the record or the archival object expresses its continuing value (by means of information exchangeability) and is seen from the continuum perspective, the record not necessarily are tied to archival institutions as places (Quisbert, 2006).
The model also highlights the role of recordkeeping in long-term preservation in relation to organizational, legal and social responsibility. It provides a possibility of building alliances with different disciplines, concerning organizational, legal, social, cultural and historical responsibility. These alliances with other disciplines should automatically identify multiple uses for records over time. (Quisbert, 2006)

Records continuum is continuous and is a time/space construct not a life cycle model. There are not separated parts of a continuum and its elements pass into each other. Upward tried to emphasize in the model the future independence of some period of time or work overload. The model is self-referencing. Within any implementation environment, the terms will be given specific interpretations and meanings and the way the elements will join with each other will vary. (Upward, 1996, 1997, 2000)

The recordkeeping axis – this deals with the vehicles for the storage of recorded information about human activities. Its coordinates are those of the document, the record, the archive and the archives (Upward, 1996, 1997, 2000).

- The document – within the model, is based in an act and is a pseudo representation of that act. In this form, it has content, structure and context of creation. The document has not yet been communicated,
or if it has been then there is no way of establishing the context of that communication, other than by the content, context and structure of the document itself. (Upward, 1996)

- **The Record** – is a memorialized (disembedded) form of the document usually linked with other documents. It should have additional layers of context to those present in a document, and may be a surrogate record of that document. It is this additional information about context, which is the key to “disembedding” the document from its narrower contexts of creation and carrying it through time and space as a record. (Upward, 1996)

- **The Archive** – is the aggregated record viewed as all the archival documents of an organization (Upward, 1996).

- **The Archives** – is the archive in plural form. It contains the records of a number of organisations because of either spatial spread or temporal transmission to another organization or institution. (Upward, 1996)

**The evidence axis** – this axis consists of the trace of actions, the evidence, which records can provide, and their role in corporate and collective memory (Upward, 1996, 1997, 2000).

**The transactional axis** – this axis presents the act, activities, functions and purposes as co-ordinates. In the model, the terms are simply labels. This axis reflects an emphasis upon records as records of activities undertaken in the conduct of affairs, and upon the way, these activities create links between documents. It reflects, in a basic manner, functions of organizations, and the way these are broken down according to subcategories of activity, or built up from the acts themselves. The purpose is the function viewed from a broader societal perspective. (Upward, 1996, 1997, 2000)

**The identity axis** – this axis represents the actor, the work unit with which the actor is associated (which may be the actor alone), the organization with which the unit is associated (which may be the actor or the unit) and the manner in which the identity of these elements are institutionalized by broader social recognition. Identity relates to the authorities by which records are made and kept, including their authorship, establishing particularities of the actors involved in the acts of records creation, the empowerment of the actors and their identity viewed from broader social and cultural perspectives. (Upward, 1996, 1997, 2000)
Dimensions

- **Create** – in this dimension, records of business activities are created as part of business communication processes within the organizations (e.g. through e-mail, document management software, or other software applications) (Upward, 1996, 1997, 2000).

- **Capture** – recordkeeping systems manage similar kinds of transactions and records series documenting processes at the work unit or single function scope of complexity. This involves the addition metadata, which disembeds information from the immediate contexts of its creation. (Upward, 1996, 1997, 2000)

- **Organise** – in this dimension, a recordkeeping system embraces multiple collections of records that service the entire documentary needs (business, cultural, educational, and historical) of a single juridical entity. Records become part of a formal system of storage and retrieval that constitutes the corporate memory of the organization. (Upward, 1996, 1997, 2000)

- **Pluralise** – in this dimension records that are required for purposes of societal accountability (e.g., by corporate law) or other forms of collective memory become part of wider archival systems, which comprise records from a range organizations. It involves the use of information in ways, which are less predictable or controllable. (Upward, 1996, 1997, 2000)

The Records Continuum Model opens a new perspective to see records. According to the model, the record has its “own life”. It means that nobody needs to take consideration to the preservation place. The records exist irrespective of space or time. The records can be everywhere within the model (cf. Figure 1), these can be in different dimensions, but some concepts must be filled before the record can move further to the next following dimension. (Upward, 1996, 1997, 2000)

**The OAIS Reference Model**

OAIS Reference Model (Open Archival Information System) is a conceptual framework for an archival system dedicated to preserve and maintain access to digital information that was developed by US Consultative Committee for Space Data Systems (CCSDS). The OAIS information model defines the broad types of information that would be required and is to promote a broader understanding about these needs. The model constitutes the functions for long-term preservation and access to digitally stored information. The model also is aimed to guide system developers to create
software for records management. The model can be applicable to all approaches or organizations that deal with preservation and provide digital information. The OAIS model does not provide a detailed specification for an archive implementation. (CCSDS, 2002)

The model is compound by four major foundation components (ibid.):

• A **producer**, (person or machine) which produces the information to be preserved.

• A **manager**, which sets up a policy or framework for preservation.

• An **archive** (an implementation of the OAIS model).

• A **consumer**, (person or machine) which interacts with the archive in order to retrieve information.

The figure below (Figure 2) shows the conceptual model of the OAIS Reference model. A system build according to this model should contain six level processes (ibid.):

- **Ingest** – receives the information and prepares it for preservation.

- **Archival Storage** – storages, maintains and retrieves the preserved information.
- **Data Management** – coordinates description information from an archival object and system information, which is used to support archival operations.

- **Access** – helps the consumer to identify and retrieve information.

- **Administration** – plans the ubiquitous archival activities. Monitors the environment and assess the development of new standards and policies.

- **Preservation Planning** – provides recommendations for conversion, migration, monitoring changes in technology.

**The OAIS Concept of Information**

Information itself is defined as any type of knowledge that can be exchanged, that is independent of the forms used to represent it (i.e., physical or digital), and this information is always expressed by some type of data. Both persons and systems have some Knowledge Base, which allows them to understand the received information. If the person/system does not have sufficient Knowledge Base they need Representation Information to understand that information. It means that data is interpreted using its Representation Information, which produces information. In order to successfully preserve such a generic Information Object, it is critical for an OAIS to clearly identify and understand the Data Object and its associated Representation Information, which together forms what is called an Information Package. (CCSDS, 2002)

![Figure 3 – The OAIS Information Model.](Lavoie, 2000)
**Information Package**

The Information Packages contain both the archived file as well as its descriptive metadata. The description is divided into four categories (cf. Figure 3): Content Information, Preservation Description Information, Packaging Information, Descriptive Information. The Content Information and Preservation Description Information are viewed as being encapsulated and identifiable by the Packaging Information (cf. Figure 4). The resulting package is viewed as being discoverable by virtue of Descriptive Information. (CCSDS, 2002)

![Diagram of Information Package Concepts and Relationships](image)

According to OAIS there are three types of packages (cf. Figure 2 and Figure 3), which move around in the system and which change their status depending on where they are in the archive. There are three types of packages (ibid.):

- **Submission Information Package (SIP)** – this is the digital object package, which is sent to the archive by the information producer. In some cases, it may well not arrive with adequate metadata.

- **Archival Information Package (AIP)** – this contains all the information needed in order to preserve the digital object. All the relevant metadata will be added and technical specifications to render the object.

- **Dissemination Information Package (DIP)** – the package that the user will receive is the digital object and some of its metadata. There is
little need for the user to have access to all the complex technical metadata stored in the AIP.

**The Preservation Object**

“Preservation Objects are those to be archived. These objects are physically compound by set of files in maybe different formats. The objects logically are compound by data – the content for instance birth date of a person, and metadata – that should describe the context when the birth date of a particular person was used.”  
(Dollar, 2000)

In the scientific material I did not found any uniform definition of the notion Preservation Object, but the notion can be compared to approach of Persistent Object articulated by the US National Archives (NARA) in cooperation with the San Diego Supercomputing Center (SDSC). The persistent object method presumes that all records can be represented as objects with their specific characteristics and behaviour (Ruusalepp, 2005). According to Ruusalepp the persistent object is what is created at the ingest stage from a document deposited in a digital archive for preservation – essentially an AIP (according to the OAIS Reference Model). For creating a persistent object, all significant properties of the object that is to be preserved are identified and expressed in explicit, abstract (formal) models in XML. Such significant properties would, as a minimum, include (ibid.):

- The internal components of the object.
- The sequence of components within the object.
- The attributes of presentation of the preserved object.

Once the properties for a range of typical records have been defined as a model (e.g., typical structure of an e-mail, a letter, a bill, etc.), these records can be encapsulated in metadata as defined in their respective models. After that, other technical characteristics of records that are dependent on specific hardware or software, proprietary, or subject to obsolescence are to be eliminated by converting to a platform-independent format.
CHAPTER 3

Method

This chapter describes how the investigation has been done. The chapter explains research approach, research purpose, research strategy, how information was searched, methods for information collection and analysis approach.

Research Approach

To relate theory and empirical data, I chose a deductive approach. Based on existing theories about preservation object, OAIS Reference Model, concept of records and the Records Continuum Model I drew conclusions about individual phenomenon that embrace the concepts of preservation object. My choice was based on deductive approach because it was more appropriate for my study since I first entered deeply into a theoretical area. This influenced later work that is what information has been gathered, how the information has been interpreted and how the outcome has been related to existing theory. The deductive approach, according to Bryman (2002), is bracing the objectivity in the research just through that starting point takes already in existing theory. The consequence will be that research process become coloured of the subjective understanding of the researcher in the less extent. The disadvantage of deductive approach, according to Patel & Davidson (2003), is a risk that the already existing theories which researcher is starting from will be direct and influence the research in such a way that interesting new findings cannot be discovered. The statement is correct about the influence on research process, but it has not influenced my research process, because I carried out a descriptive research.

Research Purpose

The majority of investigations in any research can be classified on the basis of how much a researcher knows about some problem area before the investigation starts (Patel & Davidson, 2003). In this thesis, knowledge level is descriptive, which means that there is certain amount of knowledge about the investigated problem area that embraces the developed structure for preservation object. I limited my work to illustrate a specific problem area. This area embraces existing concepts within the developed structure for preservation object. The investigation is based on the knowledge that existed within this problem area.

According to Bryman (2002) when a quality-focused researcher reports his/her study, the descriptions often seems to be filled with large quantities of detailed information. Because this investigation is quality-focusing the descriptions are detailed, which Bryman recommends. This is because
descriptions are important to understand the context of the social reality that is been studied and that has to be interpreted within a context. If I did not take into consideration some special environment in which, for example a preservation object structure was created, it means that no one can be able understand the context and the behaviour of it.

**Research Strategy**

This investigation characterized by a hermeneutic approach. According to Patel & Davidson (2003), hermeneutics stands for interpretive science and today has a scientific direction in which a researcher studies, interprets and tries to understand the fundamental conditions of human existence. In the contemporary debate between the positivistic approach and the hermeneutical approach, hermeneutics stands for a system for qualitative interpretation and understanding. In this system, the researcher has an open (subjective) and engaged attitude. The main idea in hermeneutics, according to Bryman (2002) is that the researcher that is analysing a text will reproduce the meaning of that text on the basis of the perspective that the author had. The focus lays on the historical and social context in which the text was produced.

A hermeneutic approach constitutes (in this case) a perspective on the analysis of the data that makes possible a link of understanding the text, between ones own visual angle and the social and historical context in which text was created.

I agree with the hermeneutical standpoint, which states that the human existence can be interpreted through language because humans have intentions, purposes, language and actions that are possible to interpret and understand its meaning. In this investigation, I am approaching the research object subjectively on the basis of my own understanding. My understanding, thoughts, impressions, feelings and the knowledge that I have, I see like a resource but not like an impediment to interpret and understand the research object.

**Qualitative-focused Research**

This work is a qualitative-focused case study, which means that the theoretical knowledge standpoint is interpretative and it is very important the understanding of social reality. This work is based on a research strategy that stresses text rather than quantifies collections and analysis of data. According to Bryman (2002), qualitative research distances from especially positivistic research. Qualitative research instead focuses on how individuals understand and interpret their social realities. Bryman further argues that qualitative investigation can be used not only for generating new theory, but for examining a theory as well.
The fundamental form for a case study, according to Bryman, consists of a detailed study of one case and examines the specific complexity and nature of the case. Further, the author means that a case study often choose qualitative methods because the methods works well when a researcher wants to make an intensive and detailed review of one case. A case study distinguishes from the other approaches in that the researcher often is interested in illustrate unique features of a specific case.

The case study, described in this thesis, was carried out at the LDP Centre. This Centre, amongst other tasks, works with the development of structure for so-called Preservation Object. This structure for preservation object is for delivering archive records from an authority to an archival organisation (in this case the Swedish National Archives).

According to some critics, it is difficult to generalize qualitative research results beyond the situation in which such results were achieved because one case cannot be representative for many other cases (Patel & Davidson, 2003). The goal with this investigation was to carry out a detailed study of only one case that embraces a structure for preservation object developed by LDP Centre. Hence, the intention was not to generalize the research results. The results can however be taken into consideration when researching other similar cases.

**Information Collection**

To gather information I used documentation and conducted an interview, which according to Patel & Davidson (2003) are significant techniques for qualitative focused research.

The literature that constitutes the base for this work, I looked for at library of Luleå University of Technology. I used the national catalogue Libris, the electronic database Ebsco, the electronic database Elsevir and the electronic portal Samsök. I also used the search engine Google to get access to relevant material and for expanding my knowledge in the subject. To find suitable material treating problem area I used key words, which according my opinion captured the core of the investigated subject.

Besides from using search words and synonyms for identifying key sources, I followed up references in the books and articles to other works that appeared to be important and relevant. To locate them I used referencing list of articles or books. I searched also through author’s name that often showed up in the literature or research within a certain part area. To choose key sources, I used some appraisal criteria, for example if some article had scientific characteristic and was relevant to the investigated area.
The documentation that I used was both in printed and electronic form. It embraced conference papers, reports, and specialist literature, scientific articles, books, electronic books and licentiate thesis. To appraise about relevance of literature I looked on the title, the author, the publisher, the publication year and reading summary and conclusions.

Qualitative interview was another technique to gather information. The reason to this choice was partly that the developed structure of preservation object was not documented and partly that interviewing was a god technique when concerning to get so detailed picture of the investigation area as possible. According to Patel & Davidson, qualitative interviews have a low degree of standardization that is the questions gives scope for respondent to answer with his/her own words. Bryman (2002) means that the interviewing process is flexible and that the stress must be on how the respondent understands questions and the course of events during the interview. This understanding might concern events, patterns and behaviours; from this understanding he/she answers questions embracing what he/she believes is important.

To address specific interview questions I used a semi-structured interview because my intentions were to begin the investigation with a proportionately straight focus on the concepts in the developed structure of preservation object. Hence, I drew up a list, an Interview Guide, over the questions and subjects that would be covered in the interview. The advantage with such guide is that questions can be asked in different order than the order the questions were written in the interview guide. The questions, which not were included in it, could be asked if I saw a connection to something that the respondent has said. In this kind of interview, the respondent has a great freedom to articulate answers of her/his own way.

When drawing up the interview guide I followed up fundamental advices described by Bryman. First, I created order in current themes so these should be followed successively in a god way. The first theme was embracing a comprehensive describing of the developed structure of preservation object that later was followed by the specific concepts of a record, which the Records Continuum Model handle. Later, I tried to formulate questions in a way that should make it easier to answer the questions and also tried to avoid leading questions. I began the interview guide with questions about background facts, for example about the respondent’s position in the organization and his/her employment.

Bryman, however, advises that a researcher will be aware that an approach can be very time-consuming and that hi/she will need good equipment (i.e., the tape recorder, capacity of the batteries and the microphone). Transcribing also creates a frightening large heap of paper.
The interview I recorded on the tape and later wrote it down. This way, according to Bryman is important to perform a detailed analysis in qualitative investigations and to capture answers of the respondent in his/her own terms. This approach was suitable because I interviewed only one person and in this way, I could follow my planning. The investigation object in the case study was the manager of development at the LDP Centre. He is responsible for coaching the developers, for all existing technology, which is used in the development work, for monitoring technology and other aspects of the digital preservation field. From the beginning, he was involved in the LDP Project with the purpose of carry forward the results of the project and to further develop these results.

**Analysis Approach**

To work up the collected information I selected a qualitative method, which can be applied on working up text material. According to Patel & Davidson (2003), there are not unambiguous rules that are designed for how a qualitative analysis would be carried out. Hence, I selected information analysis based on criterions for qualitative focused research.

To get a picture of quality in an investigation for some researcher, according to Bryman (2002), reliability and validity have no relevance. The author means that the researcher will assume a number of standpoints when concerning with this question. I chose one of fundamental criterions to appraise the qualitative investigation, **Trustworthiness**. This criterion consists of four part criterions which have correspondence in the quantitative research.

- **Credibility**, which is a correspondence to internal validity.
- **Transferability**, which corresponds to external validity.
- **Dependability**, which can be compared with reliability.
- **Confirmability**, which corresponds to objectivity.

To apply reliability and validity criterions on qualitative investigations these criterions presuppose that it is possible to come only to one and absolute picture of the studied social reality. It means that there is an absolute truth about the social reality and this is a task for social researcher to reveal. (Bryman, 2002)

**Credibility**

According to Bryman (2002), credibility in the results includes that the researcher has secured that the research was carried out according to existing rules. Furthermore, the researcher reports the results to the people who are part of the social reality that has been studied. This because these people will confirm that the researcher has understood this reality in a correct way. It is often called **Respondent Validity** or **Participant Validity**.
To achieve credibility in the results I have tried to secure that the research has been carried out according to existing advices and instructions, which was described by Patel & Davidson (2003) and Bryman (2002). I also took into consideration advices and comments of my supervisor. Furthermore, I carried out respondent validity by submitting a report to the respondent in order to get a confirmation that the describing I have intermediated was correct.

**Transferability**

Since, qualitative research in usual cases includes an intensive study of a little group or individuals of some common qualities (because it deals with deep and not with breadth), the qualitative researcher, according to Bryman (2002), is challenged to produce what is called *Thick Descriptions*. These thick descriptions are part of a culture. In the investigation, I was striving to reproduce a detailed report (a thick description) in order to provide other people a “database” that could help to appreciate how much the result is transferable to another environment.

**Dependability**

To appraise an investigation when dealing with dependability I adopted a reviewing approach, so-called *Auditing*. It means, according to Bryman (2002), that the researcher has to ensure that a complete and available report was created of all phases of the research process. These phases are: problem-formulation, choice of investigated persons, interview printout, decision concerning analysis of data and so on. To guarantee dependability in the investigation my colleagues and my supervisor have been reviewers or “auditors” during whole research process and latest when the investigation began to be finished. In this way, they could appraise the quality of the chosen approaches and how these approaches were applied. I took into consideration all comments that I received during the research process.

**Confirmability**

To confirm, according to Bryman (2002) means that the researcher, on the basis of the insight that it is not possible to get any complete objectivity in social research, is trying to ensure that hi/she has acted in good faith. It can be with other words obvious that the researcher unconsciously has let personal values or his/her theoretical focus to influence the accomplishment and the conclusions from an investigation. One of reviewers’ tasks, the author means, is to establish in what extent it is possible to confirm the results. During the whole work time I was aware about this criterion, hence I strived to be objective in order to avoid influence in carrying out the interview, analysis and conclusions by my personal values.
Empirical Data

The chapter describes empirical data received from the interview. The chapter begins with a describing of main notions constituting the structure of preservation object developed by LDP Centre. Later, the chapter presents the answers on the questions about concepts concerning digital preservation object based on concept of records in the Records Continuum Model.

The Preservation Object

The LDP Centre is striving to develop a preservation object that should be independent of technology. It means that the information that is the preservation object will be entirely independent of any programming language, operating system or platform. In this way, such preservation object can be transferred between different technologies. The LDP Centre says have good tools for this matter. There are some formats, which settle to be running independently of operating system or software. There are structures in XML, which also are independent of any platform or operating system, but it has not been tested yet because the test platform has not been built yet.

According to LDP Centre, a preservation object will include the object itself, preservation information about the object and technical information about the object.

Preservation information creates a semantic picture of the object to put it in the context of its creation, which is a part of preservation. Preservation information consists of the data object itself and representation information. The representation information describes the purpose of the object, a preservation perspective and guarantees authenticity of the preservation object.

At the LDP Centre, the structure of a preservation object, is called Information Delivery Package has been developed (cf. Figure 5). The package consists of a base structure and the specific information preservation object. The base structure in the delivery package consists of different sets of metadata according to international metadata standards. These standards are: Metadata Encoding & Transmission Standard (METS), Encoded Archival Context (EAC), Encoded Archival Description (EAD) and PREservation Metadata Implementation Strategies (PREMIS).
The start point for this general base structure has been Open Archival Information System Reference Model (OAIS). The different concepts together in the base structure is corresponding the whole information model in the OAIS (cf. Figure 3).

- **METS** – is the structure to identify the package. METS mainly corresponds to Packaging Information, according to information model in the OAIS standard. METS contains a catalogue of files and classify these files in the package. METS also will contain technical metadata linked to the individual files and contain a link to the describing package content in the EAD, which corresponds to Descriptive Information about Package in the information model in the OAIS standard.

- **EAC** – is the structure to describe archival creator that is persons, family or organization that delivers archival material contained in the delivery package.

- **EAD** - is the structure to describe the archive that is how the delivered archival material is structured. EAD corresponds to Descriptive Information about Package, which according to information model in the OAIS standard describes content of package.
• **ERMS** – is the structure to deliver material from records and document management systems. ERMS contains metadata about the object and about structure of the system. The ERMS structure is divided in different levels, which means that one can choose what levels that would be delivered. These different levels are Class and Subclass, which describe the process in the delivering system. Folder, Record and Component belong to subclasses in the ERMS structure. Folder level describes the different classes of the errand, which is kept in the delivered material. Record describes included records in some errand. Component keeps the technical information about the file itself; however, the LDP Centre has chosen to migrate technical information to the PREMIS structure instead of using the Component level. This way gives faster access to technical information when converting and makes easier changes of the file formats in the archive.

• **PREMIS** – is a method to document changes that has occurred for example in a record. It means that the PREMIS will carry up all technical information as well as give, together with a safe process for management of changes in the record, a guarantee for authenticity of the record to and within the archive. Technical information in this respect is not general information about file format but specific information concerning a specific file. There is no requirements that the PREMIS would be included in the system but according to the LDP Centre to use the PREMIS or any other structure with the same aim creates a higher reliability of the files, which later be delivered to the archive.

• **Content Information** – is specific information, which would be delivered to the archive. Content Information contains, apart from the digital object, also representation information for this object. Representation Information is information that render possible to put the object in its context. This information makes possible to understand the object both out of technical and semantic perspectives. In the Information Delivery Package structure of the LDP Centre, the unique technical information is kept in the PREMIS structure, not in Content Information.

**Creating document**

When a document is created (i.e., a file), content, structure and context are included in the file. Information about how the file is created and why it has

---

1 The specification of ERMS of this level is Folder and corresponds to ”ärende” in Swedish, which is a bit different than its conceptual peer in English.
been created may be retrieved from the ERMS structure. However, the structural information, which is created besides is keeping in the format itself, independent of its type.

To identify who created the document or its owner, the structure at Record level of the ERMS is used. There is information about who is responsible for a document or an errand, who has made changes in a document and whom the document concerns.

Information describing in what transaction the document was created and also information describing why the document was created, that is in what purpose the document exists can be retrieved from the ERMS structure.

**Capturing records**

Record is a level in the ERMS structure that itself is a record. Records management systems keep records that keep documents together. There is also information about the document(s) and a link to some other documents, even though it is not so usual in the case of the LDP Centre to keep the elements of records together in the ERMS structure. However, there is a possibility to keep the concepts of records together in the ERMS structure on the Records level.

Within the structure for Information Delivery Package, a part of the content could be seen at the Record level in the ERMS, which describes what information the file is keeping. The description of structural part (structural context) is not visible but it is kept in the file itself because the format keeps its own structure and metadata.

Within the EAC structure, it could be seen information about in what activity(s) the record was created. When talking about the processes in an organization the EAC describes the structure of this organization. The record is linked to that part of the organization, which has created this record. One part of the information is kept even in the EAD and the ERMS structures at the Folder and the Record levels. There is information that describes the processes of the organization in which the record was created, persons that have created or changed the document and also the roll of these persons in the organization.

The structure of the Information Delivery Package makes it possible to receive information about where in the organization, in what unit(s) the record was created. However, it is optional for organizations to have such description. Even though the structural possibilities exist, the organizations prefer to avoid it; because it deals with a constantly organizational change, where a department that is today can be a unit tomorrow, and so on. This leads to constant changes in the description, which is problematic for the
organisation. On the other hand, for the LDP Centre and on the basis of archival perspectives, it is desirable that organizations have such description in the delivery.

In the Information Delivery Package is included a description of evidence which the record obtains and in what purpose the record was created and is kept. The ERMS structure describes the processes that are following up persons that were involved in the creation of the record. The processes in turn generate some type of information, which is a part of the record(s). In this way, the records are linked to specific processes. The above-mentioned makes possible that the evidence is visible that the record is authentic.

**Organizing Records in the Archive**

The basic idea of organizing the records in an archive, when concerning the Swedish National Archives (Riksarkivet in Swedish) was to build it within the EAC structure. When searching after an archive belonging to an organization, information about all deliveries that has been submitted to the National Archives can be retrieved. Every delivery can be divided by series at the Class level in the ERMS through the EAD structure and later can be divided in sub-series, in other words, which describes every delivery. However, it is problematic when concerning to structure up the large amount of records. During one year, an organization can produce up to 20 millions records, for example the National Social Security Board (Försäkringskassa in Swedish). It means that in practice, it is difficult to organize and put all together into one single archive. One can just imagine the amounts of records created by this organisation after ten years if every year it produces about 20 millions records. Theoretically, it can be possible, but practically very large amounts of records have to be unused. It was not real addressed how multiple records would be structured because some material can be classified as secret. The material is not available apart from some researchers, in other words it is still an uninvestigated area.

Within the Information Delivery Package can be identified that the archive belongs to a certain organization, for example by process searching or by person searching.

To retrieve information about what function within the organization that the archive supports, the ERMS structure is used. Within the ERMS structure, information about main process, which has used within the organization, information about sub-processes and so on, can be received. The purpose with the incorporation of process thinking in the structure was to get this description information more independent of organizational changes.

The purpose with the archive is regulated by Swedish legislation. The organizations have demands in archiving its information. The EAC structure
describes why the archive exists and the EAD structure describes the purpose with every level and every sublevel in the archive describing. In Sweden the EAD structure may be used, nevertheless it will be described according to the Swedish scheme.

**Archives**

The National Archives collects archives from several Swedish authorities. According to the LDP Centre, the following information is absent in the structure of digital preservation object:

- The concept embracing content, structure and context of the archive collection.
- Information describing role of the archive in collective memory or purpose of the archive in the society.
- Information describing purpose of the functions of organizations in the society.
- Institutionalizing of the archive and how these archives is identified in the structure of Information Delivery Package.

However, these concepts will not be contained either. Archives from different authorities are separated in separate parts within the Swedish National Archives. To search information across all archives is very restricted. For example to search information across several archives by personal number is illegal in Sweden. The managerial group has got a commission by the Swedish Ministry of Culture to collect the archives and the managerial group in turn takes the decision about how the archives will be structured, with other words this structure is too broad to be achieved of the Swedish National Archives. The archive itself cannot keep so much information about the organization because it does not concerns the digital archive, nevertheless it concerns semantic information about the organization, which keeps information and hence is ending up at higher outside level.
CHAPTER 5

Data Analysis

The chapter presents the analysis of empirically received information in relation to the theoretical foundation that lies in the background for the investigation. This I carried out using the Records Continuum Model as a “template” for records towards the developed structure of preservation object, analysing likenesses and differences.

Document

According to the Records Continuum Model, a document is created as a result of business actions. Hence, it constitutes part of the business communication processes within the organization. A document is a result of human action that has content, structure and context. The document of archival value should be identified by creator, owner or other involved actors. The document should: a) be linked to the transaction in which it was created, b) have a purpose and c) be traced. (Upward, 1996) These above-mentioned concepts are presented in the structure of preservation object. A digital document with its content, structure and context creates automatically in a file and in that way the elements of the document keeps together. Contextual information, for example about how the file was created and why the document was created can be retrieved from the ERMS structure.

Actor(s)

Contextual information about creator, owner and other actors that contributed in the creation of a document in some way or other manipulative acts with the document is presented in the ERMS structure at the Record level. That, according to the Record Continuum Model (Upward, 1996, 1997, 2000) constitutes the identity of the created document. It means that the information about who is responsible for the document, who has made changes and whom the document concerns can be received.

Transaction

The concept of transaction is representing act that is a part of a business activity within an organization (Upward, 1996, 1997, 2000). It means that there should be information describing the transaction within the organization in which the document was created. Within the preservation object in the ERMS structure can be retrieved information about the transaction in which the document was created.

Trace

According to the Records Continuum Model (Upward, 1996, 1997, 2000) the acts in which documents were created, should be traced. The created
document will serve like evidence of the transaction. To trace acts the information about in what transaction the document was created and in what purpose the document was created should be possible to receive. In the ERMS structure within the preservation object, information describing in what transaction the document was created and also information describing why the document was created, in other words in what purpose the document exists can be retrieved.

**Record**

Records are captured as result of that recordkeeping systems manage similar types of transactions. When a record is produced, it may consist of several documents, which are linked to each other. A record, according to Upward (1996) should have content, structure and context. It means that the additional layer of context is necessary to that presented in the document that describes the record. In the structure of preservation object, information about such elements can be retrieved through the ERMS structure at the *Record* level. Record management systems keep records that keep documents together. There is also information about the document(s) and a link to some other documents, even though it is uncommon in the case of the LDP Centre. One part of content keeps at the *Record* level in the ERMS structure, which describes what kind of information the file contains. On the other hand, structural context is not visible because it is kept in the file itself because a format keeps its own structure and context. There is however a possibility to keep together these elements of record in the ERMS structure at the *Record* level.

**Activity**

Activity consists of number transactions within organization. A record, according to Upward (1996, 1997) is result of an activity and hence the information about this activity and information about to what activity the record belongs should be in the recordkeeping systems. There should also be information about to what transaction the document belongs and to what activity the transactions belong. Within the EAC structure in the structure of preservation object, information about in what activity the record was created is visible. The EAC structure describes the processes in an organization. The record is linked to that activity in what it was produced. Some pieces of information are kept also in the EAD and the ERMS structures at the *Folder* and the *Record* levels.

**Unit(s)**

Unit is a smaller part of the organization. The actors belonging to a unit can be linked to some activities. Information about where in the organization, in other words, in what unit(s) the record was produced and how these units and the actors are linked to each other should be described. (Upward, 1996, 1997,
The structure for the *Information Delivery Package* makes it possible to retrieve information, which identifies in what unit(s) within the organization the record was produced. Since it is optional for organizations to have such description, they drop it because of constant changes in the organization. Reorganizing demands that the organization changes the context and that is something they want to avoid. From the archival perspective, it however is desirable that organization has such description in the *Information Delivery Package*.

**Evidence**

Every record, according to Upward (1996, 1997) should have a specific purpose in business activities, which will be described in the context. It embraces the information about why a record was produced and kept, in what purpose and what evidence the record can provide, in other words, the role of the record in the organization and its authenticity. In the *Information Delivery Package* is included a description about evidence that the record provides and in what purpose the records were produced and kept. The ERMS structure describes the processes, which follow up persons that were involved in producing the record. The processes in turn generate some type of information that is part of record(s). In this way, records are linked to specific processes and by that give visible evidence that records are authentic.

**Archive**

It is about multiple collections of records that are organised by means of recordkeeping systems, which belong to the organization. According to Upward (1996, 1997) it needs an additional layer of context. When the collections of records are organized in an archive, the archive should have content, structure and context. In the structure of preservation object when addressing a specific case of the Swedish National Archives, the idea was to organize that archive within the EAC structure there information about all deliveries from the specific organization to the National Archives could be retrieved. Information about how the delivery was divided in series and later in the sub-series can be provided through a search path from the EAC through EAD to ERMS structure. In this way, the information describing every delivery can be provided. The structuring of a set of records within an archive however, is problematical because large amount of records. For example, the National Social Security Board during ten years has produced up to 200 millions records; hence, there are no structure solutions for such archive. Theoretically, the set of records can be organized within an archive but practically the organization will not be effective because a large amount of records will stay unused. This is an uninvestigated area because in the specific case of the National Social Security Board, all material is classified as secret, which means that the information is inaccessible apart from some researchers.
**Organization**

Every organization has its own archive. It means that identifying some archive belonging to its creator is necessary. According to Upward (1996, 1997) there should be information describing the authority (creator) itself, how it is organized in units, actors in these units, their responsibilities and authorship, which supports the records. In the *Information Delivery Package* can be identified that the archive belongs to some organization and even other describing information. It can be done through process searching in the EAC structure, and even through person searching.

**Function**

Every organization has a function in business making. In the recordkeeping systems, according to Upward (1996, 1997) this function will describe the manner the function breaks down into activities and further to sub-activities. There will also be information about what function in the organization the archive is supporting. In the structure of preservation object to retrieve information about what function within the organization that the archive is supporting, the ERMS structure is used. The ERMS structure keeps information about the main process of an organization, information about sub-processes and so on. The purpose with the structure of the organization was to keep the context more independent of organizational changes.

**Corporate/Individual Memory**

An archive constitutes the memory of an organization. The context should describe the purpose of the archive. This, according to Upward (1996, 1997, 2000) includes information about why the archive exists and what evidence the archive provides for the organization. According to LDP Centre, the purpose with the archive is regulated in Sweden by Swedish legislation. The legislation demands that organizations should create an archive over its activities. In the structure of preservation object, in the EAC structure is described why the archive exists and the EAD structure makes possible the description about the purpose of every level and sublevel of the archive. In Sweden however, the organizations may work with the EAD structure only according to Swedish archival schema.

**Archives, Purpose, Institution, Collective Memory**

This is about why records are needed in society, for example according to collective law or some other form of collective memory. This embraces a number of archives from a number of organizations. According to Upward (1996, 1997, 2000) an additional layer of context is needed. He means that records become a part of further number of archival systems that contain records from a number of organizations. The archives demands content, structure and context. Information about how the archives are institutionalized and how the archives are identified in this process is also
needed. Furthermore, information describing what purpose the functions of a number of organizations have for the society and what role the archives have in collective memory, in other words what purpose the archives have for the society. When addressing the specific case of the Swedish National Archives, the National Archives collects and keeps archives from a number of authorities. The above-describing concepts, which are included in the Archives dimension in the Records Continuum Model, do not exist in the structure of preservation object. According to LDP Centre, the concepts will not be in the preservation object. The explanation for this is that, within the National Archives, all archives are divided in separate parts because information search across all archives is very restricted; information can be searched only in one archive at a time. For example to search information by personal number across several archives is illegal in Sweden. The Ministry of Culture gave the mission to the managerial group in the National Archives to collect archives. The managerial group in turn, makes decisions about how the archives should be structured, in other words the structure is too comprehensive (conceptually and physically) for the archive of the National Archives. The archive itself cannot keep so much information about the organization because it does not concerns the digital archive, nevertheless it concerns semantic information about the organization, which keeps information and hence is ending up at higher outside level.
Conclusions and Discussion

The chapter presents conclusions that I came to in the research and discussions around the conclusions.

The purpose of the investigation was to point out possible deficiencies in the structure of preservation object, if any, developed by LDP Centre. This would be carried out using the Records Continuum Model as a “template” for records towards the developed structure of preservation object, analysing likenesses and differences. Based on the analysis, I came to following conclusions:

- The structure for preservation object developed by LDP Centre is containing all concepts corresponding concepts of records in the Record Continuum Model at three levels, namely Create, Capture and Organize. When concerning the concept Unit(s) at the capture level, information that identifies in what unit(s) within organization the record was produced, the organizations sort out such description. This because of constant changes in the organizations and the option for organizations makes possible to avoid some changes in the context of a record. However, such description is desirable from archival perspective, but this is a question for Archival Science.

- At the Organize level, there is no solution to how large amount of records should be organized by the authority, in other words there is not structure to organize the archive in the organisation. Theoretically, it is possible, but practically a very large amount of records can be unused. On one hand, to organize a small amount of records is easier, but on the other hand, if a large amount of records is removed, then valuable information might be lost. In the specific case of the National Social Security Board, the material is classified as secret. The implications of that need to be investigated by mainly Archival Science but also by other disciplines.

- The concepts embracing the fourth level, namely Pluralize, in the Records Continuum Model are not presented in the developed structure of preservation object. The first reason is that in the National Archives the managerial group takes all decisions. To introduce these concepts in the structure, a re-institutionalizing in the archival authority is compulsory. The second reason is that one cannot search information across all archives, because of legal aspects. For example to search information by personal number across all archives
is illegal in Sweden. Hence, it is a question of legislation. On that way the legislation constitutes a barrier to introducing concepts within the *Pluralise* dimension in the structure of the preservation object. This restriction applies to all Swedish archives that contains sensitive information such as diseases or personal number (i.e., archive from the National Security Board). However, these restrictions do not affect archives from, for instance the Swedish National Archive of Recorded Sound and Moving Images (SLBA - *Statens ljud- och bildarkiv* in Swedish).

- When concerning the specific case of the National Archives, the managerial group, the organizational structure and legislation put a barrier to apply the idea about alliance building and with it the space/time distanciation. Hence, I cannot comment on whether the structure of preservation object has deficiencies. This because the LDP Centrum followed up Swedish laws and norms of the Swedish National Archives when the structure for preservation object was developed.
Method Discussion

The chapter addresses the discussion about the methods I used in the investigation.

During investigation time, I realized that detailed description, which Bryman (2002) has recommended for qualitative-focused research, has significant importance. The detailed description increased considerably the possibility of understanding the context of the social reality that has been studied and that has to be interpreted within a context.

I have carried out the case study at the LDP Centre. The reason for this choice was that the LDP Centre is the only institution working with the development of a structure of preservation object. With other words, I had only one case at my disposal.

To gather information I chose to carry out a semi-structured interview with one respondent that I considered was the most representative for the case study. I did this choice because the respondent is the manager of development at the Centre. The respondent possesses necessary knowledge to describe the whole picture of my investigation area (it concerns the structure with a fixed construction of concepts). Furthermore, the respondent knows about the Records Continuum Model; this made easier the communication and understanding during the interview. Hence, I believe that if I increased the number of respondents, working within same development process, it have not affected the results in the investigation.

In the investigation, I used the Records Continuum Model as theoretical foundation. During the work time, I discovered followed weaknesses of this model:

- The model is very abstract. It describes the concepts that are too abstract to be applied at a more concrete level, in this case, it is the practical solution of structure of preservation object. To apply the model in practice it is needed a more concrete level of context to overcome between the logical level and the physical level.

- It is difficult to hold balance between the contextual levels of the concepts. That is because the concepts have different levels of abstraction. For example, the concepts of Recordkeeping axis, namely Document, Record, Archive and Archives, present a more concrete level of abstraction than other concepts are.
• Interpretation of the concepts can create problems for any interested in archival and records research or anywhere the model is used; for example, a researcher and a respondent might interpret a variety of things in different ways.

To get a picture of quality in the investigation, I have chose **Trustworthiness** that is one of fundamental criterions to appraise qualitative investigation, according to Bryman (2002). I concluded that this criterion was relevant to my investigation and helped me to increase the quality in my work. **Credibility** contributed to increasing the quality, there I as researcher could confirm that I understood the studied reality in a correct way, by carrying out **Respondent Validity** and taking into account all advices and comments from the supervisor. **Transferability** increased the quality by means of detailed report. In this way, other people might have the possibility to appraise how much the result is transferable to another environment. **Dependability** has increased the quality by means of my colleagues and supervisor, who was “auditors” and reviewed my research process and methods during several seminar occasions. **Confirmability** increased the quality through I tried to be objective and in this way I could avoid to influence interview, analysis and conclusions with my personal values.
Proposal to Further Research

The chapter presents my ideas as proposal to further research.

- During my work, I discovered that there are many definitions and opinions about the Digital Preservation Object. However, there is no general and uniform definition of it. Hence, it is valuable from a research perspective to investigate about the notion of Preservation Object, its nature and characteristics that can facilitate to give a general definition of it.

- Investigate how the Records Continuum Model can be applied in the practice. Presumably, concerning to find some suitable or appropriate models that can overcome between the logical or abstract level towards the physical level, which later can be tested.

- Because I limited the investigation only to concepts of the record in the Records Continuum Model, I propose that further investigation covering the remaining concepts in the structure of preservation object, and how those might relate or affect (the theory of) the Records Continuum Model.
References


LDP Centre homepage. Available on: http://ldb.project.ltu.se/projectweb/portalproject/Index.html [cited 12 October 2006]


APPENDIX

Interview guide

Introductory questions

1. What is your position at the LDP Centre?
2. What are your responsibilities within the development of preservation object (PO)?
3. Is there somebody else that works with the development of PO?

1 theme: ”Describing Preservation Object”

1. What is a PO?
2. Describe, what the PO consists of? Clarification: What elements are included in the PO?
3. If you see from the OAIS model perspective, what correspondence every element in the PO has in the OAIS model?
4. What purpose has each element in the PO?

2 theme: ”Records concept on the basis of the Records Continuum Model”

My intention is to find the concepts that the Records Continuum Model describes in the developed PO. My expectations are that you will help me to point out these concepts in the PO.

Create document (1:1 dimension)

1. When a document is created, it has content, structure and context. Can one see the elements in the PO?
2. Can you point out where in the PO are the elements?
3. Can one identify who created the document, i.e., who is the actor, who is owner of the document?
4. Where in the PO is this information?
5. Can one see in what transaction(s) the document was created?
6. Where in the PO can one see this information?
7. Can one trace actions in what the document was created?
8. Where in the PO is this information?
Capture record (2:d dimension)

When a record is created, it can consist of several documents, which are linked to each other, i.e., a record has content, structure and context. It means that the additional layer of context is needed.

1. Can one see these elements in the PO?
2. Where in the PO are the elements?

3. Can one see in what activity (within organization) the record was created?
4. Where in the PO is this information?

5. Can one see where in the organization, i.e., in what unit or units the record was created?
6. Where in the PO is this information?

7. Can one see in what purpose (evidence), i.e., why the record has produced and keeps?
8. Where in the PO can one see this information?

Organize records in an archive (3:d dimension)

It is concerning multiple collections of records that are organised by means of recordkeeping systems, (systems to produce, receive and disseminate memory) in an archive within an organization, i.e., organize records as memory. An additional layer of context is needed.

1. The archive should have content, structure and context. Are these elements in the PO?
2. Where in the PO are these elements?

3. Can one identify that the archive belongs to a special organization?
   *Clarification: how authorities, responsibility and authorship that a record supports are mirrored in the PO?
4. Where in the PO is this information?

5. Can one see what function in the organization the archive supports?
   *Clarification: an organization has a function to make business. The function can be broken down to the sub-functions that are activities.
6. Where in the PO is this information?

7. Can one see the purpose with the archive, why the archive exists, what evidence the archive provides for the organization?
8. Where in the PO can one see it?
Archives (4:th dimension)

It is concerning records that are needed in society, for example according to collective law or some other form of collective memory. It embraces a set of archives from several organizations. An additional layer of context is needed.

1. These archives have content, structure and context. Are these elements in the PO?
2. If not, where are these elements?
3. Can one see what roll the archives have in collective memory, i.e., what purpose the archives have in the society?
4. Where is this information?
5. Can one see what purpose functions of a number of organizations have for society?
6. Where is this information?
7. Can one see how the archives are institutionalized, i.e., how the archives identifies?
8. Where is the information about it?