LONG-TERM PRESERVATION OF COMPLEX AND INTEGRATED E-SERVICES

Göran Samuelsson, Lena-Maria Öberg, Erik Borglund[1]

Abstract

Access to public documents is a Swedish citizen's right. The decision processes should be traceable, which requires that information systems must fulfil requirements for traceability between information and processes. This research project is focused on long-term preservation of an integrated and complex e-service named Home Building Guide an e-service which handles the process of planning, building, and living in your home. This research study in two parts focused on the question of how to preserve digitally born information in complex and integrated e-services. It aims too answer the questions **who** will take the responsibility for recordkeeping in this public cross- authority documentation, and **how.**

1. Introduction

In order to be able to reconstruct the traces of our past it is crucial to document the present. Public electronic services (e-services) will form an essential part our current history, i.e. the archives of the future. A pro-active holistic approach is essential if we will be able to preserve trustworthy and accessible electronic information in both textual and non-textual over time as a future resource for business, cultural and scientific needs.

This project will focus on two main questions. The first is concerned with the preserving of the digitally born information in complex and integrated e-services. The second focuses is on the preservation of the e-services itself or (maybe more correctly) on the different processes which together create the e-service. These e-services where local, regional and central governmental authorities, together with enterprises and, of course, the citizens of Europe, co-operate in decision-making to create a large quantity of unique process information.

Public cross-authority documentation has great record keeping functionality hidden in audit trails, as defined in the ISO 15489 Records Management standards. Today few are aware of this and no-one has the tools in place to handle it. If we do not succeed in capturing the complex trace of today, we risk tomorrow's description of this revolutionary period in human technology development becoming like a summer meadow containing only a withered and dried flower from a herbarium. No matter whether the services deal with driving licences, health care or building permits, it is obvious, that we have to document our governmental decision processes even in these new integrated e-services – but **who** will take the responsibility in this public cross-authority documentation, and the next question is **how**? If

---

[1] Mid Sweden University, Gånsviksvägen 2-4, S-887 88 Härnösand, Sweden, goran.samuelsson@miun.se.

we do not have any answers to these questions we will never reach the great efficiency and rational potential of our e-society.

In this study we will produce a checklist for an adequate management of a modern integrated e-service so that it will be possible to retrieve, re-use and research the information and records in the future.

A pro-active holistic approach is essential to be able to preserve trustworthy and accessible electronic records. If the records are not captured and prepared for long-term preservation in current business processes, it will be very difficult and costly to subsequently preserve them for long periods in the requisite trustworthy form. Hence, not only the corporate memory of organisations is endangered, but also the collective memory of human activity is likely to be lost.

1.1 State of the art - long-term preservation and recordkeeping

Research in digital preservation and accessibility of records and archives has been carried out in different projects all over the world, for example the InterPARES project, the SPIRT Recordkeeping Metadata project, among others. Of these, the InterPares (The International Research on Permanent Authentic Records in Electronic Systems) project has become the most well-known. It aims to develop the theoretical and methodological knowledge essential to the long-term preservation of authentic records created and/or maintained in digital form. [1]

A high quality metadata model is essential to facilitate preservation of contextual information, migration of records through future system changes and to ensure retrieval in the future. This SPIRT Recordkeeping Metadata project provided a metadata reference model that will inform the design of the metadata standard for this e-government recordkeeping research project.

The Home Building Project will draw on the findings of these research projects and test their applicability in an actual e-government service.

Further, a number of different initiatives, models and standards have developed during the last 10 years. For example MoReq (Model Requirements for the Management of Electronic Records) is destined to become the guiding principle used by ERM (electronic records management system) system suppliers and developers to provide required functionality, and by outsourced records management service providers to define their service offerings. The standard also promises to revolutionise ERM procurement, auditing and training schemes around the globe. MoReq (i.d. Moreq2) will be used by a broad spectrum of public sector and commercial organisations when preparing requests for proposals, evaluating existing systems and preparing records management training materials. It will also provide a valuable teaching resource for academic institutions. Organisations developing and implementing an ERM system based on the requirements of MoReq2 can be assured that records will be properly managed, accessible at all times, retained for as long as they are required and properly disposed of, once the retention period has ended. MoReq2 is ordered by the European Commission and will be available free of charge early next year (2008). [2]

MoReq2 will be widely used as a Standard throughout Europe; it will be an important reference model for the Home Building Project.

Erpanet is another example; the dominant feature of Erpanet is the provision of a virtual clearinghouse and knowledge-base on state-of-the-art developments in digital preservation and the transfer of that expertise among individuals and institutions. [3]

It will provide an up-to-date source of information and research findings about digital preservation that will inform this project.

The International Records Management standard - ISO 15489 based on the Australian Standard for Records Management (AS4390) is the world's guide to saving, caring for and using the information that every organization, business, urban authority or national government relies on to carry out its functions. The Standard's *raison d'être* is:

"The standardization of records management policies and procedures ensures that appropriate attention and protection is given to all records, and that the evidence and information they contain can be retrieved more efficiently and effectively, using standard practices and procedures" [4]

This Standard has, together with other workbooks and guidelines such as the Australian *DIRKS Manual* (Design and implementation of recordkeeping systems) [5] and ICA:s *Electronic records – A workbook for archivists*, provided the recordkeeping profession with tools to analyse and describe the business process.[6] Business process analysis will be the key to identifying what records are created in the e-service process and which agency is responsible for their capture and preservation.

Other important tools and models for management of current business processes and for long-term preservation is *OAIS reference model* (Open Archival Information System) [7] and the *Records Continuum Model*. [8]

The *OAIS Reference Model* is an abstract model for the way in which an archival information system should be constructed or what functions and components are necessary in order to describe an information object. The *OAIS Model* is a reference model that is intended to serve as a guideline in development of archival information systems. The *OAIS Model* has gained great influence in recent years and is often referred to in projects regarding digital preservation. It is also in the basis of many municipal e-depot projects. [9] The history behind the *OAIS Model* came from the Consultative Committee for Space Data Systems when they realised that they had lost a lot of data from their (NASA's) early space travels. With this in mind they started to develop a model for an Open Archival Information System (OAIS) which was later accepted as an ISO standard (ISO 14721:2003). The OAIS Model arose to address a situation in which digital data was irretrievably lost. This situation is a real possibility in e-services if not long-term digital preservation strategies are researched and implemented. Thus the model is particularly pertinent to this project.

The Records Continuum Model helps us understand the nature and scope of recordkeeping in our organisations and in our society. It presents an overview of a seamless and dynamic recordkeeping regime that transcends time and space to capture and manage records for as long as they are required to satisfy business, regulatory, social and cultural requirements.

Records continuum thinking and practice focuses on logical records and their relationships with other records and their contexts of creation and use. Thus the Continuum is a map of a dynamic, virtual place - a place of 'logical, or virtual or multiple realities' - and it always has been, even in the paper world. The ability of the Records Continuum Model to account for records in a variety of formats in context, encompassing the records themselves, the systems in which they are created and used, as well as their current and possible future uses in society makes this a sound theoretical basis from which to approach the range of issues which must be addressed in this project.

Conclusions

Despite the wealth of these works and research results that have been produced in recent years, there has been non-specific knowledge created about these emerging complex and integrated e-services. We can see an increased awareness about the need for long-term preservation to support re-use and

research in the future.[10] However, no extensive work has focused on the complex nature of electronic records and preservation in modern integrated e-services. This ambition of this research project is therefore to contribute to knowledge essential for tackling a crucial need in society: a practical implementation of strategies to preserve digital records in context for continuing primary accountability requirements as well as for the possibility of future secondary research based on archival records.

2. Research objectives

In Sweden as in most of the European countries big efforts and activities are focused on creating public e-services for their businesses and citizens.[11] With an e-service accessibly, seven days a week and twenty-four hours a day, the intention is that citizens should be able to follow their on-going processes and access public documents for themselves. In Sweden there is an on-going project called the "Home Building Guide" [12] the goal is to build an integrated e-service which will cross agency borders in order to make it easier for a citizen to seek information, for example about a property unit purchase or to follow the different processes that are needed to build a house. One example of such a process is applying for a building permit, during which several authorities could be involved. If those authorities could co-operate in an effective manner the process of getting a building permit would be easier and more efficient.

Today a lot of the public documents are electronic and included in different kinds of computer-based information system. Examples of such systems are: document management systems; record keeping systems and workflow systems or combinations of these. The problem is however, that the requirements for designing such an information system that take account of traceability of processes and general requirements for record keeping over time and cross authority under Swedish conditions are rare. In most cases they are totally missing.

This kind of problem becomes even more complex in an e-service like the Home Building Guide where a lot of different actors are involved, with differing information flows and regulations covering each type of agency.

Even if the following questions have many similarities had we chosen to divide these questions in to two research studies whereas the first has its focus on question dealing with responsibility;

- If we find information which is unique for the e-service, who has the responsibility for taking care

   of this information?

- Who documents and archives the information about the e-service itself?

- Is the "owner" of the e-service and also the agency with formal responsibility for capturing and keeping the archival? How will we deal with this if the e-service is outsourced to a private company?

- If the partners in these e-services create their own e-archive solutions, how will the e-service connect to these multiple e-archives?

- If the e-services offer citizens personal storage and archive possibilities for their enquiries or projects, how and where will this e-archive be created?

- If this e-service creates some kind of "Citizens account" or citizens e-archive who will manage these citizen accounts and at what cost?

In the second study we will focus on questions dealing with which kinds of information and metadata the e-service needs in order to function rationally and to provide a digital archival repository in a proper and satisfactory way. The goal for this part is to map and define:

- Will the e-service create information and records which are unique for this process and can only be understood in relation to other information in the e-service?

- What kind of information and metadata is needed to describe and manage the e-service and its processes?

- Which type of metadata will be necessary to ensure possible future re-use and research. They should also be suitable for the new archival description principles in Sweden which will be implemented in 2008.

- Suitable file formats for document and images for capture, archival storage, retrieval and delivery to future users.

2.1 Results so far

Responsibility

Most of the processes in connection to the e-service will be possibly to manage in according to traditional archive principles, but since our research indicate that the e-service create it's own audit trails which is most obvious in the parts which is marked (dark/red) in figure 1, further research is necessary.

Fig 1
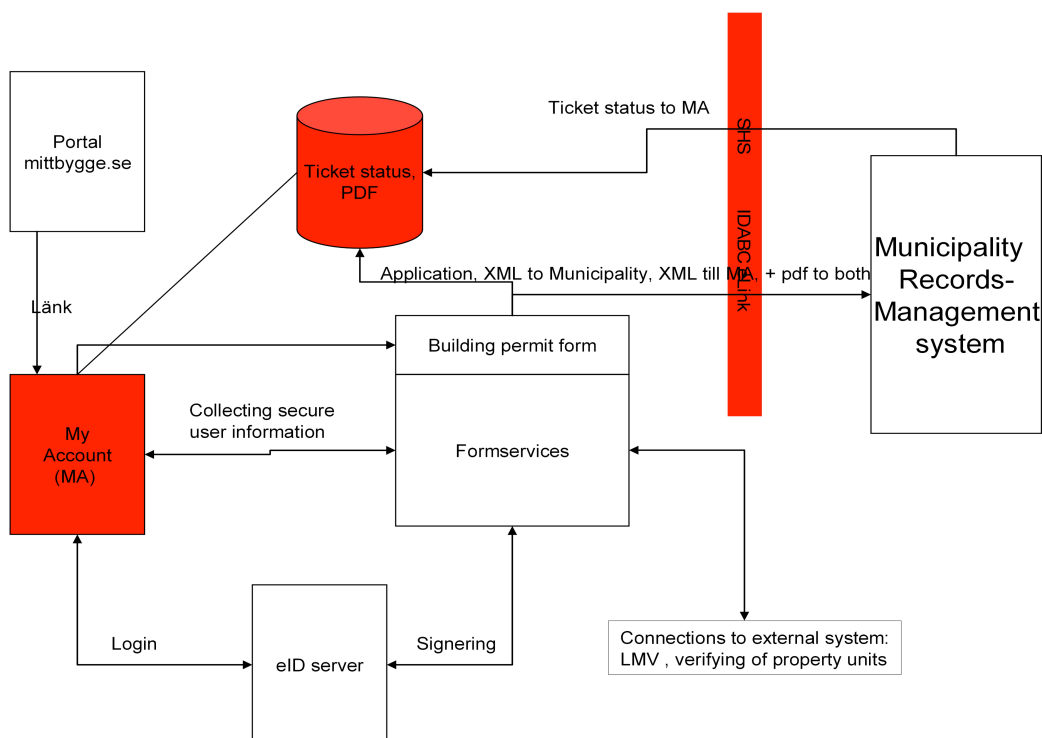
The handling of this information will be even more important to define since the responsibility for the management of the e-service Home building guide probably will be managed by a private company.

Information

In a first study [13] we focus on components that are needed to reconstruct the process for a building permit today. Two Swedish municipalities have been part of this study and it shows that the archived material gives a rather good picture of the internal process and the components that has showed to be important are; contexts, documents and dates. The study has shown that the handling and thereby the archive is limited to a process that include only one actor. The municipalities in this study appear to be in a good position in preserving processes, which is a key factor for openness, traceability and in the future effective and useful e-services, but this study has revealed some problems. The process that is preserved in the archive shows the municipalities' internal processes. The "Home Building Guide" and similar integrated e-services will place the citizen inside the process. This study does not show if a citizen thinks that the preserved process gives a fair picture of what has happened. There are for example few components of dialogue with the applicant during the process of a building permit. During the seminars with the employee it was revealed that dialogue with the applicant takes a lot of time and includes help with drawings and help with filling in the application. The communication will often be recorded by e-mail or by forms. This raises a need for discussion about appraisal of what has to be preserved and not.

There are some audit trails that connect to other agencies' work (for example, consultations with the environmental health committee) but to build integrated, effective e-services the archives must also be more integrated with each other. This will include both co-operators within the Municipalities and external actors if the "Home Building Guide" vision should become reality. How this should be solved is anything but straightforward. Questions that have to be solved are for example; where processes cross agency boundaries, who is responsible for the documents; who captures and preserves the documents and how the context of the documents is maintained and how documents should be exchanged.

Another problem that has been revealed during this study concerns the process borders. When the building permit process is finished maybe a new sub process starts? Or some sub processes may have preceded the building permit process, or be contained within it. How and where should those components of traceability be preserved?

This study has made a contribution about requirements for designing information systems that include public documents. In particular, it has looked at requirements that regard traceability and required components to be able to reconstruct processes.

2.2 Ongoing research

Next step will be an in-depth study of the information flow of the e-service Home Building Guide.

Second tasks to solve are the way the municipalities in the Home Building Guide handle there file formats. Today they do this in a diversified way - frequently creating images (tiff) of the document in the process. This research project will suggest a more object-oriented approach to archiving the information in this e-service and to create a plan for how this will be carried out. This is also according to the underlying incentive for the legal demands in the new regulations for archival description which the National Archive of Sweden will implement in 2008.

This project is significant because it will use the results of the research projects and resources outlined in earlier research, testing them in a long-term preservation system designed to serve a real e-government

service. The two studies will be completed at the end of September 2007 and a report will be produced by the end of 2007

References

[1]   http://www.interpares.org

[2]   http://ec.europa.eu/idabc/en/document/6495

[3]   http://www.erpanet.org/about.php

[4]   ISO TR 15489.2 Guidelines, section 1. Scope

[5]   http://www.naa.gov.au

[6]   http://www.ica.org/biblio.php?pdocid=285

[7]   Reference Model for an Open Archival Information System (OAIS) Blue Book, http://www.ccsds.org/CCSDS/documents/650x0b1.pdf [2004-08-24]

[8]   http://www.sims.monash.edu.au/research/rcrg/publications/Frank%20U%20RMJ%202001.pdf

[9]   See for example the Municipal Archives of Rotterdam http://www.rlg.org/en/page.php?Page_ID=20865&Printable=1&Article_ID=1799 and the City of Stockholm http://www.ssa.stockholm.se/Arkivbeskr/Miniseminarium/Kravspecifikation%20e-Arkiv%20SBK.pdf

[10] See for example Australian Government Information Interoperability Framework 2006: *Sharing information across boundaries*, from The Australian Government Information Management Office (AGIMO), http://www.agimo.gov.au/__data/assets/pdf_file/50725/Information_Interoperability_Framework.pdf

[11] See for example EU IDABC eGovernment Observatory http://europa.eu.int.idabc/en/chapter /140

[12] http://www.mittbygge.se/Pages/Default.aspx

[13] Öberg L-M, *Archiving 2006*, Ottawa, Canada;May 23, 2006; p. 165-168.