Helen Weldearegay Tekulu

Integrating Semantic Web Technologies for Norwegian Records Management Standard Noark5, An exploratory case study

Supervisor: Thomas Sødring

A thesis submitted to the Oslo and Akershus University College of Applied Sciences faculty of Journalism, Library, and Information Science in partial fulfillment of the requirement for the Degree of Master in Digital Library Learning

June 27, 2012

Master thesis
International Master in Digital Library Learning

DECLARATION

"I certify that all material in this dissertation which is not my own work has been identified and that no material is included for which a degree has previously been conferred upon me"

Helen Weldearegay Tekulu

Abstract

Digital archives, Electronic records management (ERM), and digital libraries are organizations that entitle themselves as information providers. Their primary function is often constrained to selecting, organizing, storing, maintaining, and updating information with the support of various application software and systems and often interacting with web-based data. These software applications, systems, and programs are chosen in order to provide better and seamless services to end users as well as to satisfy user needs and expectations for digital archives or digital libraries. For that reason, these information providers are now have to familiarize themselves with various information and communication technology (ICT) mediums in order to distribute and share data over the web. In this century, the information professional, or information supplier, and knowledge society is acknowledging the importance of technology. Therefore, at present everyone is seeking valuable information and knowledge on how to use also the current technological advancements in an environmental that is rapidly and constantly changing. As a result, users' interest, wants, and needs to access, share, retrieve, and in general to manage information and data goes beyond expectation and capacity of the current ERMS, and/or digital archives.

The emerging semantic web technology is in position to help these organizations to catch up with the current revolutionary technological changes through the introduction new ways of organizing, sharing, defining, and providing relevant and meaningful information. On the other hand, the importance of the semantic web is not only limited to defining, organizing, creating, and representing data/ Knowledge, but also able to facilitate, gather, and provide detailed information from various applications and databases on behalf of the human. Doing so makes life easier for humans through saving time, and minimizing efforts when searching and retrieving required information. There are many more benefits of the semantic web and this study takes the first step to introduce and explore the benefits of semantic web for ERM, and archive institutions. The researcher argues that an organization can make cultural and technological, as well as contextual change such as going from a system centric view into more data and user-oriented view by means of semantic web technology. The study adapts technical, logical, and methodological approaches to integrate semantic web technology for the ERM and

uses an appropriate framework and prototype throughout the implementation. Therefore, the overall result of this research work verifies the importance of the semantic web technology for ERMS using the semantic RDF language. In addition, the present results would enable both the ERM and archive organizations to make use of semantic-based data sources, and respectively it opens an opportunity for incorporating external semantic sources into a Noark5 system (A Norwegian Records Management Standard). In addition, semantic web technologies able to expose nested data from a Noark5 system by serializing and parsing the Noark5 extracted data, while at the same time it creates new opportunities of creating semantic relationships in between what we call from an archive perspective, *objects of interest (people), the resources, and places*. Moreover, the current RDF data model results are able to establish a basis for creating semantic Noark5 records and for creating semantic interoperability between various applications and databases without any interruption.

Acknowledgements

Table of Contents

viii

List of Tables

List of Figures

Acronyms and Abbreviations

NOARK   Norwegian Archival Standard

RDF   Resource Description Framework

RDFS   Resource Description Framework Schema

ERM   Electronic Records Management

ERMS   Electronic Records Management System

NARA   National Archives and Records Administration

W3C   World Wide Web Consortium

ICT   Information Communication and Technology

FOAF   Friend Of A Friend

FRL   Frame Representation Language

KRL   Knowledge Representation Language

KLONE   Knowledge Base

OWL   Web Ontology Language

EAC-CPF   Encoded Archival Context –Corporate Bodies, Persons, and Families

DC   Dublin Core

XML   Extensible Markup Language

SGML   Standard Generalized Markup Language

URI   Universal Resource Identifier

HIOA   Oslo and Akershus University of College of Applied Sciences

IsaViz   Visual Authoring tool for RDF

KR   Knowledge Representation

SW   Semantic Web

LOD   Linked open Data

ICA   International Council on Archives

ISO   International Organization for Standardization

EAD   Encoded Archival Description

OAIS   Open Archival Information system

Chapter One

1. Introduction

1.1. The effect of technology

We are living in an information era where information plays significant role in everyday activities. In addition, technology has brought various alternatives ways and opportunities in regards to share or store information on the web. In particular, the existence of information communication and technology (ICT) and the World Wide Web (WWW) provide a new way of handling and disseminating diverse sources. Knowledge societies are now able to use the advantage of technology in their everyday activity to share, transfer, exchange, and access the right information at the right time. In fact, not only the knowledge society but also the information professionals, information providers and other service suppliers are now aware of the technological change. For instance, libraries, archive institutions, data analysts (information analysts) and others are implementing, maintaining, and organizing their resources using user-oriented technologies such as digital archives, digital libraries, content management systems, social networks, virtual community platform, e-learning, record keeping systems and record management systems.

Thus, software applications are widely used to record, maintain, and share heterogonous and massive data across the world. However, it is true that, still those software applications cannot resist the fast growth of technology. Therefore, we should be able to open our eyes every time, every moment and then make use of the new emerging technologies without surrendering to it.

For instance, there are typical new technologies such as semantic web technology that can enable us to maintain interoperability, to establish interaction between machine and human, to increase flexibility, to enhance usability and accessibility of structured and meaningful data on the web. Further, many more technologies are in use for several purposes such as for medical and agricultural purposes.

However, we also understand that the emerging technology have its own significance and influence in all human activities. For instance, lets take from the library or archive perceptive: the challenge for changing the library structure and functions; the need for adapting new skills, such as for digitizing, scanning physical documents, manuscripts, historical records; the challenge for data migration and integration; the need for technical skills to implement, configure, and upgrade software applications etc. These are just an example from the economical, cultural, and human resource perspective that at least library/archive organizations and its staff member could face obstacle to overcome the consequence.

However, since the main objective of these organizations is to increase the use and access of information by their end users, then therefore, it is worth to deal and embrace all the technological factors that are currently emerging, in order to create a better world, a world that can share, link, and access meaningful information on the web.

1.2. An overview of records management and long-term preservation

The "ISO 15489-:2001" standard defined the records management term as a "field of management responsible for the efficient and systematic control of the creation, receipt, maintenance, use and disposal of records, including the processes for capturing and maintaining evidence of and information about business activities and transactions in the form of records" (Australia, 2012).

An organization has a long history and tradition of keeping, organizing, maintaining, recording, and managing their organizational activities including employee records, activities, and business transactions from the first period they are created to the last period of disposal and preservation. At present-day the shift from the traditional records management to the electronic records management has added more value of using and accessing both paper and electronic based documents. Electronic records management system is a system, which is used by records management organization to " manage the creation, use, maintenance, disposal of electronically created records for the purposes of providing evidence of business activities"(Australia, 2012).

In addition, electronic records management system (ERMS) software is widely known for supporting, arranging, organizing, and ensuring captured digital resources or records. For that reason, various electronic records management systems have been developed for the purpose of storing and tracking electronic documents or electronic records of an organization. In addition, several and different type of regulations, requirements, standards, and specifications for electronic records management has been published and practiced by various record keeping and archive organizations. The National Archives and Records Administration (NARA), National archives of UK, National archives of Australia, National archive of the New Zealand, the European Commission, and the National archive of Norway are some of the organizations that have developed different types of standard, requirement, guidance, and regulation to maintain, preserve, and to manage the electronic records management system (ERMS) and records.

Accordingly, the general life cycle and the overall transaction process of the records from the creation period of documents to the transformation and long-term preservation of documents has been able to bridge the difference between the records management and the archive profession.

1.3. Norwegian Archival Standard or Norwegian records management standard (Noark5)

The national archives of Norway use the Norwegian archival standard Noark to maintain and ensure the quality of public administration records. Noark was developed in 1984 as a requirement specification for electronic record systems in government administration, and soon became a de facto standard in Norway. In 1994 Noark-3 and in 1999 Noark-4 was published with the additional specification from Koark (municipal version of Noark). Further, presently the new Noark5 has released with new benefits and functions(Noark5, 2012).

The Noark5 specification and requirement is now widely used in public administration and other sectors in Norway to ensure the quality and authenticity of the record.

Here it is important to note that, in Norway the term archive and records management are often used interchangeably and therefore there is no semantic difference between records management and archives.

However, this can often confuse to non-Norwegian or international readers because internationally the records management functions and activities is distinguished from the archives activities. However, in this context we might use the word records management as to refer to archives and vice versa.

1.4.   Todays Archive Institution

Archival institutions are widely known for providing long-term service through preserving, recording, organizing public and personal documents, cultural and organizational documents, governmental, national, and historical documents for long-term purpose. In addition, archive organizations have a kind of their own culture i.e. protecting, organizing, and preserving sensitive information about public bodies, historical records, governmental documents and other relevant memories. Archive institutions are also known in preserving past and long historical records for instance storing former governmental documents or documents that are not in use at present government, and other related kind of records.

Consequently, they have also a unique structure, context, roles, and regulations in accordance to the right for accessing, sharing, and using of archival records on the web. However, the traditional context of handling, accessing, and sharing resources needs to be changed and currently archival institution are showing the radical shifts in a way of sharing and accessing electronic resources whether by digitizing the physical resources that have been used in the archive or by preserving new digital born documents. Moreover, there was professional argument in relation to the archive organizations and records management, addressing that archives focus more in preserving and protecting the past records than the active records, whilst the records management focus in organizing and providing in hand records or active records of the organizations.

 However, the fact is that both are aiming in providing quality services to their users through organizing, sorting, and maintaining the quality and authenticity of the overall

records. Currently, the use of technology and cultural, contextual change enable both the records management and archives to work together and to deal with their records in accordance to the rule and regulation of records management and archives.

1.5. Motivation

It is true that information and communication technology has been able to change the way society share and access information on the web. The main objective of sharing and accessing huge amounts of information at the right time anywhere by anyone across the world are able to accomplish with the help of various technologies and communication channels such as the World Wide Web (WWW).

The World Wide Web plays significant role for distributing, sharing, linking, and accessing various source of information on the web. Then recently the semantic web technology comes with new promise that is more advanced and flexible than World Wide Web, which enables us to create an environment where machines and humans can both interact and understand the meaning of data.

The WWW inventor Tim Berners- Lee have made his point about the semantic web by addressing that, "I have a dream for the Web [in which computers] become capable of analyzing all the data on the Web-the content, links, and transactions between people and computers. A 'Semantic web', which should make this possible, has yet to emerge, but when it does, the day-to-day mechanisms of trade, bureaucracy and our daily lives will be handled by machines ………" (Badr, Chbeir, Abraham, & Hassanien, 2010, p. v). In reality, his dream is becoming into practice and now several research works and projects about semantic web has shown surprise results and are able to verify the importance of semantic web technology for organizing, sharing, accessing, and retrieving relevant data on the web. In addition, the semantic web technology is also able to address various issues and overcome different problems such as structuring unstructured data, collecting, opening, and linking huge amount of data over the web etc.

On the other hand, the W3C group argues that, "there is lots of data we all use every day, and it is not part of the web. I can see my bank statements in a calendar. But can I see my photos in a calendar to see what I was doing when I took them? Can I see bank statement lines in a calendar? Why not?

Because we do not have a web of data. Because data is controlled by applications, and each application keeps it to itself" (W3C, 2011).

I could not agree more locking and hiding data has been the central problem of several applications. For instance, searching for detail information about well-known artist, his/her biography, works, publications, and his/her current favorite activities etc altogether cannot easily retrieve even if the information is appropriately organized. Of course, the user will be able to retrieve some data of the artist using the traditional key word search engine, which is not satisfactory at all.

The main reason behind the application is that developers developed applications that are less flexible and less accessible by others; this will help them to keep their business competition in the market. However, securing and restricting the application directly affects users and prevents them from using and accessing the actual or right information. Therefore, the semantic web is aiming in overcoming such difficulties, challenges, and solving similar real world problems such as, easy way of knowledge and data representation, improving information use and retrieval, and uncovering or unlocking data from applications.

In addition, the semantic web technology can provide key solutions that help information suppliers as well as users to facilitate and integrate semantic based services, to enhance the usability and accessibility of meaningful data, and to allow easy way of searching and retrieving relevant information on the web. Despite these facts, I become more interested in using the importance of semantic web technologies into records management standard Noark5 and find if this technology will be useable in records management and archives.

1.6. Statement of problem

The Norwegian public administration has used Noark for electronic record keeping system and standard for organizing, managing, and recording their documents. The public administration, or public bodies, such as municipalities of Norway, national archive of Norway, and other private sectors are using the Noark5 based system for ensuring and authenticating their records. The Noark5 based system enables the public administration bodies to structure and stores their digital born documents or digitized documents online.

However, today the problem with the electronic records management system of the Noark5 is that this system has some possible limitation of locking and controlling the municipality data and prevents users from using or accessing the actual data.

This is actually the problem of most application systems, it kept and controlled the data within it self, so that information owners as well as users are not able to get access to the right information at the right time. The public administration could have several solutions for solving the problem for instance using the semantic web technologies such as resource description framework (RDF), Ontology, and linked open data (LOD) and making the data more accessible and usable. Therefore, the researcher believes that introducing the use of semantic web technologies into the records management and archives could have several benefits. Some of the key benefits of semantic web technologies are unlocking the data and information from the system property, allowing data and information understandability by both human-machine, opening an opportunity to shift from the system centric approach to the more information centric approach, allowing easy search and information retrieval, and integrating semantic based service into records management and archives.

1.7. Objectives of the study

The main objective of this research is to explore the importance of semantic web technologies for the records management standard Noark5. As a result, the study will discover and discuss the possible options for using, integrating, modeling, and linking the Noark5 data, metadata, and record structure. The following research questions will be used as a preliminary point for the research.

- How can Noark5 metadata be mapped into resource description framework (RDF)?
- How can RDF be used to represent Noark5 metadata, records, and still retain the traditional archive qualities such as provenance and context?
- How can resource description framework (RDF) be integrated in electronic records management system (ERM)?

1.8. Significance of the study

The researcher believes that introducing semantic web technology into the records management will have its own benefit such as accessing semantic based record management data, improving accessibility, and usability of data over the web. Further, the semantic web is more flexible and interoperable within any system applications that means integrating and exchanging data between several applications can be made easily. Therefore, precisely the study might have the following benefit:-

- It allows the use of semantic web technologies as a technology on the electronic records management.
- It allows easy communication, interaction, and understanding of real objects or actual data by both humans and machines.
- It allows easy way of accessing, using electronic records, and digital archive data
- It introduces the possibility of representing, linking, and interchanging electronic records between various applications and databases.
- It opens an opportunity for using external semantic sources such as Friend Of A Friend (FOAF) and Dublin core (DC).
- It enables the use of semantic search engine technology and semantic web services.

1.9. Organization of the thesis

The first chapter provides an introduction about the effect of technology, brief overview about records management and long-term preservation, background about Noark5 recordkeeping standards, and today's archive institution and their digital shift. In addition, the chapter provides an outline of the research along with the research motivation, statement of problem, objective of the study, and significance of the study.

Furthermore, Chapter 2 contains a state-of-the-art discussion in relation to semantic web technologies, semantic web components, recent literature review of semantic web and the knowledge representation, knowledge management, digital libraries, and digital archives. In addition, an introduction of the Noark5 standard is covered including backgrounds, organizational function of the National archive and other contextual information is taken into consideration.

Moreover, Chapter 3 provides detailed description about the methodological and technical approach of the research. This includes the framework, and how a proposed prototype is adopted to help, to draw the researcher focus and guide the implementation that is undertaken.

Chapter 4 puts the research question and research objective into implementation and the chapter provides a detailed model, representation, mapping procedure, and explanation of the resource description framework (RDF) implementation. Chapter 5 discusses the overall implementation and modeling results, and discusses the research questions, statement problems, research objectives, and chosen methodologies. Finally, Chapter 6 concludes this work and provides some for future work.

Chapter Two

2. Literature Review

2.1. The semantic web and knowledge representation

The idea of representing knowledge and information has been widely used in the field of artificial intelligence, psychology, computer science, and linguistic areas. Researchers and practitioners have been conducting various researches in order to develop further strategy for knowledge representation. In fact, in the last four decades research works about knowledge representation have shown both positive results and discover difficult problems (Patel-Schneider, 2004).

On the other hand, a computer scientist Mylopoulos, from the University of Toronto, provides an overview of knowledge representation scheme that could be use in semantic web data modeling. The logical representation schemes uses the "notions of constant, variable, function, predicate, logical connective and quantifier" in order to represent fact whilst the semantic network representation schemes uses the "term object (nodes) and binary association (labeled edges)" to describe the real world object. Further, semantic network scheme uses various association types such as "classification (Member-Of/Instance-Of), aggregation (Part-Of), generalization (IS-A), and others partitions" in order to relate the subjects or the resource with its object. Then again, there are frame-based knowledge representation schemes, which are used for representing knowledge such as Frame Representation Language (FRL), Knowledge Representation language (KRL), KLONE knowledge base, and web ontology language (OWL) (1980).

In fact, Owl web ontology language is now part of the semantic web language and " is designed for use by applications that need to process the content of information instead of just presenting information to human" (W3C, 2007). Moreover, semantic web languages such as, resource description framework (RDF), resource description framework Scheme (RDFS), and web ontology language (OWL) are used for multiple purpose including modeling, reasoning, knowledge and information over the web.

Accordingly, Patel-Schneider stressed that the semantic web vision opens a new various opportunities for " knowledge representation that have not been exploited in the past as

10

well as some new difficulties that could hinder application of techniques from knowledge representation to the semantic web" (2004, p. 1).

Therefore, the main challenge of semantic web is to provide new possibilities that can solve the complexity of traditional knowledge representation (KR) such as providing a "language that express both data and rules for reasoning about the data and that allows rules from any existing knowledge representation system to be exported on to the web".(Lee, Hendler, & Lassila, 2001, p. 4)

On the other hand, the web along with various search engines has been able to provide and retrieve an organized data and information on line in a way human can read and understand it. For instance, if a person searches for information regarding a former English rugby captain Martin Johnson using the search engine by writing full name of the rugby player "martin Johnson" on the web, then soon after the search engine will retrieve webpage for the rugby players, relating result to his famous artist namesake, and other information.

However, the problem here is that, the result cannot show detail information of Martin for instance, telling the computer that Martin Johnson is a rugby player and that every rugby player is a person. Here is the benefit of semantic adding extra meaning, which is semantic meaning to the web enable computers to retrieve about Martin Johnson as he is a rugby player and this player is a type person. Therefore, semantic web is now beyond the current web and the traditional knowledge representation. It provides rich knowledge representation mechanisms (Breslin, Passant, & Decker, 2009).

2.2.   The semantic web and knowledge management

Knowledge management systems and other related technologies have been used as a means to organize, disseminate, and exchange information and knowledge within the organization context and culture. Intranet is one part of the technology that play significant roles in the development and use of tacit and explicit knowledge. In particular, it opens up an opportunity to interact and connect people with common interests and enable them to share and exchange tacit knowledge, in addition the explicit information is accessible and usable via the intranet(Davies, Fensel, & Harmelen, c2003, p. 2).

11

However, according to Davies et al., the current knowledge management systems have shown some significant weakness in searching information, extracting information, maintaining information, and automatic document generation. Semantic web provide more opportunity to use information in a way both human and machine can benefit for instance, assigning meta-information to the document reduce the retrieval of irrelevant information while also "the explicit representation of meta--information, accompanied by domain theories (i.e. ontologies) will enable a web that provides a qualitatively new level of service" (c2003, p. 4).

Additionally, Antoniou & Harmelen, provides a detiled explanation in relation to the use of knowledge management system and the current web and discuss that knowledge management aims in acquiring, maintaining, and accessing knowledge within an organization. As a result, knowledge management has considered as a main part of large businesses and international organizations activity because they start to look their "internal knowledge as an intellectual asset from which they can draw greater productivity, create new value, and increase their competitiveness" (2008, p. 3).

However, the problem with the current knowledge management system is that it shows some limitation regarding to technology for instance, searching information using traditional keyword-based search, extracting information required human time and effort, lack of maintaining information such as inconsistencies in terminology and failure to remove outdated information, uncovering information, and viewing information. In this case, the semantic web technologies can provide a solution and an opportunity to use more advanced knowledge management systems. For instance, allowing knowledge to be organized in conceptual spaces according to its meaning, to use automated tools to check the inconsistencies and extract new knowledge, to replace keyword-based search with query answering, and to define who may view certain parts of information. (Antoniou & Harmelen, 2008, p. 4)

2.3.   Semantic web technologies

Semantic web aims in making data understandable, readable, useable, and accessible anywhere on the web by both people and machines (Passin, 2004, p. 26). The traditional way of searching, extracting and accessing information using the World Wide Web

(WWW) is the starting point to the current semantic web technology. Therefore, semantic web enables both humans and machines to interact, understand, search, and retrieve information on the web. For that reason Lee et al., address that semantic web is not the substitute of the WWW rather it is "an extension of the current one, in which information is given well-defined meaning better enabling computers and people to work in cooperation" (2001, p. 3).

Furthermore, the World Wide Web Consortium (W3C) group provides a holistic explanation of semantic web as two things i.e. semantic web as a common format that is used for integrating and combining of data from diverse sources and semantic web as a language that allow data to relate with real world objects(W3C, 2011).

In addition, recently several research works have been studied related to the semantic web and most of the researchers produce different results and definitions regarding to the semantic web. Passin, summarize most of the definitions as following (2004, pp. 26–27).

The machine-readable–data view:-"Semantic web is a vision: the idea of having data on the web defined and linked in a way that it can be used by machines not just for display purposes, but for automation, integration and reuse of data across various applications." (w3c, 2003)

The intelligent agents view:-"The aim of the semantic web is to make the present Web more machine-readable, in order to allow intelligent agents to retrieve and manipulate pertinent information." (Cost et al, 2001)

The distributed database view:-"The Semantic Web concept is to do for data what HTML did for textual information systems: to provide sufficient flexibility to be able to represent all databases and logic rules to link them together to great added value." (W3C, 2001)

"A simple description of the Semantic Web is that it is an attempt to do for machine processable data what the World Wide Web did for human readable documents. Namely, to transform information processing by providing a common way that data can be accessed, linked together and understood. To turn the web from a large hyperlinked book into a large interlinked database." (SWAD-E)

The automated infrastructure view:-"in his recent Scientific American article Berners-Lee argues that the Semantic Web is infrastructure and not an application. We couldn't agree more." (Tuttle et al, 2001)

"Therefore, the real problem is the lack of an easy automation framework in the current web…" (Garcia and Delgado, 2001)

The servant-of-humanity view: - "The vision of the Semantic Web is to let computer software relieve us of much of the burden of locating resources on the web that are relevant to our needs and extracting, integrating, and indexing the information contained within." (Carnefild, 2001)

"The Semantic Web is a vision of the next-generation web, which enables web applications to automatically collect web documents from diverse sources, integrate and process information, and interoperate with other applications in order to execute sophisticated tasks for humans." (Anutariya et al, 2001)

The better-annotation view: - "The idea of a 'Semantic Web' [Berners-Lee 2001] supplies the (informal) web as we know it with annotations expressed in a machine-processable form and linked together." (Euzenat, 2001)

The improved-searching view: - "Soon it will be possible to access Web resources by content rather than just by keywords." (Anutariya et al 2001) "The main goal [of the technology described in the paper] is to build a structured index of the Web site" (Desmontils and Jacquin 2001)

The web services view:- "Increasingly, the Semantic Web will be called upon to provide access not just to static documents that collect useful information, but also to services that provide useful behavior." (Klein and Bernstein 2001)

"The Semantic Web promises to expand the services for the existing performed manually and by introducing new applications that are infeasible today" (Tallis, Goldman, and Balzer, 2001).

There are much more working concepts and definitions of semantic web technologies and thus all mentioned definitions indicates that the semantic web (SW) can be suitable and usable within any applications or aspects of an organization.

2.4.   The semantic web and digital libraries

Kruk & McDaniel, on their "semantic digital libraries" book introduces the semantic web technology into digital library and address that adopting semantic web technology into the digital library service will facilitate an easy ways of searching, retrieving, and enhancing accessibility and interoperability. In addition, they also mentioned that digital libraries aim in increasing information usability and accessibility, therefore the semantic web capacity can "offer more efficient solutions for accessing content and metadata." Accordingly, adding the semantic web to the digital library or semantic digital library can allow information integration based on different metadata sources, can provide interoperability with other systems, and can provide more user-friendly search and browsing interfaces(2008,p.73).

However, researchers have been evaluating and examining the applicability of semantic web into digital library using the three digital libraries i.e. BRICKS digital library infrastructure, Greenstone digital library, and JeromeDL semantic digital library. Subsequently, the studies have shown positive outcomes as well as generate challenges into the digital library arena. For instance, the JeromeDL is one of the most surprise results which is able to provide the semantic web and digital library services altogether with its special stress on social semantic whereas semantic web to greenstone digital library has produced challenges.

The authors mentioned that to use semantic web technology to digital library is not free cost and is more complex. This means, to accomplish richer mark-up and to create metadata required more time and effort. In addition, to configure an ontology-enabled Greenstone 3 collection, to prepare FRBR based retrieval, and to achieve internal semantic modeling within documents requires more time, effort, resources intensive and in general, it requires different level of commitment (Kruk & McDaniel, 2008, p. 176).

2.5.   The semantic web and Digital archives

As we explained it in the pervious section, several researches work in relation to the semantic web and digital libraries, semantic web and knowledge management system, and others related factors has been carried out. The role and importance of semantic web technology is continue to be used in various areas such as semantic web with social networks (Peter, 2007), and semantic web and cultural heritage (Benjamins et al., 2004). Moreover, recently Mazzini & Ricci, has conducted a case study of using the semantic web in the historical archives of the Emilia-Romagna region. Their study has been able to introduce the use of semantic web technologies for the archival description of EAC-CPF standard using the ontology language (2011).

Another related study by McDaniel, about semantic web digital libraries for the archives discusses that the semantic web digital libraries are able to provide "faceted search, user tagging, folksonomic search, ontological search, and annotation", therefore the digital archives can make use of those advantages in order to improve the usability and accessibility of archival document (Kruk & McDaniel, 2008).

Moreover, the main capability of semantic web is that it is flexible and interoperable within any applications or disciplines and always has something to offer. At present, several research works and projects are in progress and much more organizations are able to produce surprise results and improvements on their services. In addition, interaction with their customers, employers, and collaboration with several big and small industries are now becoming more easy and possible through sharing and exchanging information and data between them.

2.6.   The semantic web components

2.6.1. Metadata

Nowadays, several sources, documents, web pages, and massive amount of data and database contain detail information about their content, which is metadata. Metadata is structured data that machines can read and understand. It also plays a significant role in providing detailed information when retrieving relevant information on the web.

The main aim of the semantic web is to make data more readable and understandable by machines. Therefore, metadata is one of the main input sources for the semantic web to express and process data easily. Further, metadata is known as *data about data or information about information* (Antoniou & Harmelen, 2008; Yu, c2007).

In addition, " metadata has been with us since the first librarian made a list of the items on a shelf of handwritten scrolls" as a result it was used for multiple purpose from the library cataloguing to the personal indexing of various sources (Breslin et al., 2009, p. 59). According to Breslin et al., metadata can be use to provide a structured description of characteristic, such as the meaning (semantic), content, structure, and purpose of a web resource. Moreover, metadata can be use to enable more sophisticated search engines on the web, to facilitate information sharing, to support intelligent agents and the pushing of data, to minimise data loss or repetition, and to help with the discovery of resources by enabling field-based searcher (2009, p. 59).

There are much more metadata types, standards that are used to describe information about data. Public and academic libraries, archive institutions, record-keeping organizations, private and governmental organizations, large business companies and others are using various metadata elements. Some of the metadata standard and specifications are: Anglo- American Cataloging rules (AACR2), Library of Congress Classification (LCC), Library of Congress Subject Heading (LCSH), Machine Readable Cataloging (MARC), Dublin Core (DC), Dewey Decimal Classification (DDC), Describing archives: A Content Standard (DACS), Encoded Archival Context-Corporate Bodies, Persons, and Families (EAC-CPF, Encoded Archival Description (EAD), and Open Archival Information system (OAIS).

Generally, using metadata help us to ensure document quality and authenticity, to facilitate easy retrieval, to enable linking internal documents within external sources, to identify who create, processes and function the document. In addition, it enable us to recognize *archive responsibility if the system is shared by various organizations* also ensure *access rights and screening of information*, and help archivist to control *preservation and disposal*. (Noark5, 2012, pp. 34–35)

In addition, Standard Generalized markup language (SGML), Extensible Markup Language (XML), and Resource Description Framework (RDF) are meta-languages used

to describe and maintain metadata information. Particularly, Resource Description Framework (RDF) is the common standard that is used to describe metadata.

2.6.2. RDF, RDFs and Ontology

RDF is used to model, describe, and exchange information about *resources*. It also "serves as a base for higher-level language that describes ontlogies." RDF has a capability to model data in simple way using the triples subject-predicate-object and that is why known as a simple data model, which uses things called *resource* and a *statement* that links two resources. Furthermore, it is about modeling specific subject within specific object using specific predicate (Passin, 2004, p. 45).

Moreover, RDF has the power to express meaning in terms of triple format, each triple contains *subject, predicate and object* elements. In addition, using RDF it is possible to make assertions about a document for instance people, webpage, or anything can have properties such as "is a sister" whose value is person, page, or webpage. Further, objects and subjects are identified using uniform resource identifier (URI) (Lee et al., 2001). On the other hand, RDF graphical representation is another means of representing information using the nodes and arcs symbols (W3C, 2004a).

Moreover, Yu, defined RDF as a language recommended by W3C, which is all about metadata that has the capacity to describe any fact (resource) independent of any domain, and it provides a basis for exchanging, coding, reusing structured metadata. In addition, it is machine understandable, and it allows interoperability among applications exchanging machine-understandable information on the web (c2007, p. 40).

Accordingly, RDF has three basic elements i.e. Resource, Property, and Statement. In RDF, resources is known as something and this something can be anything, and Property is used to describe characteristic, attribute, or relation of the resource, whereas *statement* is used to describe properties of resources and has the resource-property and property value format. The property value can be resource or string literal but the subject and predicate can always be only resource. The RDF statements show the relations and links between the two resources (the object and subject) using the predicate. Therefore, this

indicates that knowledge or information can simply represented in terms of triples or statement formats that is subject-predicate-object (Yu, c2007, pp. 40–42).

RDF triples enables us to define resources in similar way as the grammar of the English language and it is known for its simplicity, which offer an opportunity to express statements about any set of resources in an easy way. The RDF triple data model format always contains binaries with the value of object, which contains resources with URI, blank node, or literal and the subject, which contains resource with blank node or URI, and the property name (predicate) with URI, which is used as a means to describe and relate both the subject and object resources (W3C,2004b). Therefore, the RDF triple format looks like this *subject-predicate-object*. For a better understanding of the triple, let us have an example of a statement and the outlook of the statement in the form of triples, RDF/XML serializations, and graphical representations as following.

Example: Ivar Fonnes is the Chief Archivist of National archives of Norway. The National archives of Norway have website http://www.arkivverket.no/ and central office at Oslo.

Based on the above statement in order to model the statement in the form of triples we can use either URI, or Blank node identifier and then make triples data modeling, or RDF graph modeling, and RDF/XML serialization. The (Table2.6.2-1) indicates the *3-tuple* RDF triples, the (Figure2.6.2:1) indicates the RDF graphical visualization of the RDF/XML data model, and (Figure2.6.2:2) shows the RDF/XML serialization.

The easiest way to express the statement is then to start with the triple role i.e. *subject-predicate-object.*

The http://www.arkivverket.no (subject) - http://www.arkivverket.no/isWebsiteOf (predicate), and genid:UNorwayarchivesv (object), and

The http://www.arkivverket.no (subject) - http://www.arkivverket.no/hasChiefArchivist (predicate), and genid:UIvarFonnes (object) etc.

19

Subject	Predicate	Object
http://www.arkivverket.no	http://www.arkivverket.no/hasCenteralOffice	genid:UOslo
http://www.arkivverket.no	http://www.arkivverket.no/hasChiefArchivist	genid:UIvarFonnes
http://www.arkivverket.no	http://www.arkivverket.no/isWebsiteOf	genid:UNorwayarchives

Table 2.6.2-2.6-1 A 3-tuple (triple) representation of the corresponding arkivverket RDF/XML data model

Therefore, an RDF triple can contain the three components i.e. the *subject,* which is an RDF URI reference or a blank node, the *predicate,* which is an RDF URI reference, and the *object,* which is an RDF URI reference, a literal or a blank node.(W3C, 2004b)
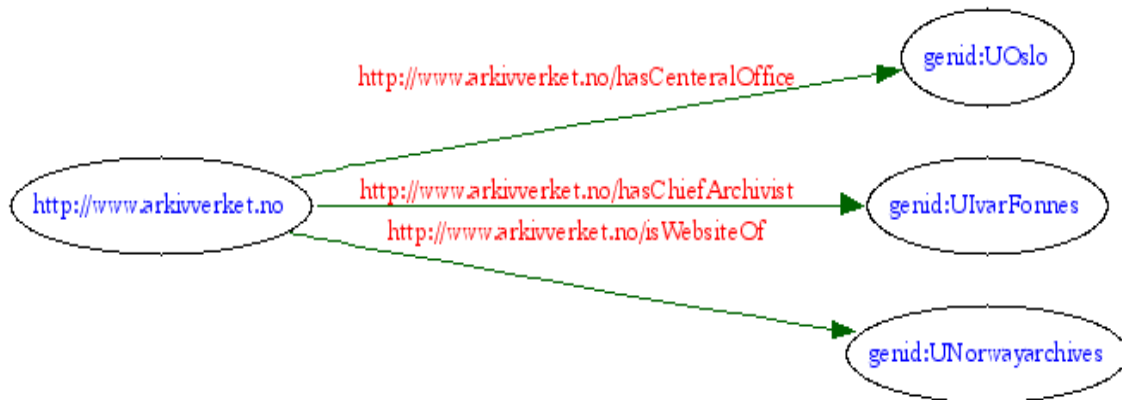
Figure 2.6.2:2.6:1 Graphical visualization of arkivverket RDF/XML serialization

```
1: <?xml version="1.0"?>
2: <rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
3:   xmlns:noark5="http://www.arkivverket.no/">
4:   <rdf:Description rdf:about="http://www.arkivverket.no">
5:     <noark5:hasCenteralOffice rdf:nodeID="Oslo" />
6:     <noark5:hasChiefArchivist rdf:nodeID="IvarFonnes" />
7:     <noark5:isWebsiteOf  rdf:nodeID="Norwayarchives"/>
8:   </rdf:Description>
9: </rdf:RDF>
```

Figure 2.6.2:2.6:2 Arkivverket RDF/XML serializations

Therefore, in general the RDF triple can be used to model and represent anything that is resources and can enable us to relate a given 'thing' to another 'thing', which means RDF is a very interesting and simple way to define and relate 'things' or *resources*.

On the other hand, RDF documents can be used as a stand alone statements and this means machines can still understand, read and can make inference based on the current document, however, RDF is missing vocabularies such as classes, proprieties etc. This indicates that RDF has some limitations, as those kinds of vocabularies never exist in RDF. Therefore, in order to accomplish the semantic web vision of sharing and distributing information on the web in a way more machine processable and machine-friendly, building and adding more vocabularies such as RDFS are necessary. RDF schema is a vocabulary description language *one can use to create a vocabulary for describing classes, properties, and subclasses of RDF resources* (Yu, c2007, p. 74).

Furthermore, ontology is richer language than RDF and RDFS that is why they built ontology language on top of the RDF and RDF schema. This language has been used in philosophy and artificial intelligence as well as in computer science field and each discipline has their own distinct definitions about ontology. Lee et al., provides brief overview of ontology definition based on those area, therefore ontology in the field of philosophy is mean the "nature of existence", of what types of things exist whilst in the filed of artificial intelligence ontology is mean a document or file that formally defines the relations between terms and *an ontology for the semantic web has a taxonomy and a set of inference rules.*(2001, p. 11) Moreover, when it comes to computer science area ontology is "an explicit and formal specification of a conceptualization" (T.R. Gruber's definition, redefined by R.Studer as cited in Antoniou & Harmelen, 2008).

In general,   (Antoniou & Harmelen, 2008, p. 13),

Ontology is a richer vocabulary description languages for describing properties and classes, such as relations between classes (e.g., disjointness), cardinality (e.g., "exactly one"), equality, richer typing of properties, characteristics of properties (e.g., symmetry), and enumerated classes.

2.7. The Semantic Web stack

The semantic web stack provides brief overview of semantic web language development and hierarchical growth. In addition, the figure shows the status of the semantic web technologies. The bottom of the stack, which is the beginning of the layer, contains Unicode the character coding system that allows data encoding and transporting within various systems with no corruption (Unicode, 2012) and the uniform resource identifier (URI) is used to describe or identify resources (W3C, 2004c). Then again, the extensible markup language and the XML schema allow us to write structured web document and publish structured data over the web. Whereas the basic and simple data model RDF allows us to express resources, add meanings, and create relationships using triples. Further, RDF schema is built on top of the RDF language and this language introduce us with new elements that do not exist in RDF such as classes, properties, subclass, subproperty relationships, domain and range restrictions (Antoniou & Harmelen, 2008).

In addition Passin, depicts RDF schema as a "framework for constructing ontologies and is used by many advanced ontology frameworks" (2004, p. 15). Moreover, the ontology vocabulary is a more advanced and richer language than RDF and RDFS, which enables us to represent "more complex relationships between web objects" (Antoniou & Harmelen, 2008, p. 18).

On the other hand, the top layers that are consider as part of logical reasoning (the proof and the logic) are used to ensure the correctness and consistency of data sets. Moreover, the trust layer provides *authentication of identity and evidence of the trustworthiness of data, services, and agents* (Passin, 2004, p. 15).
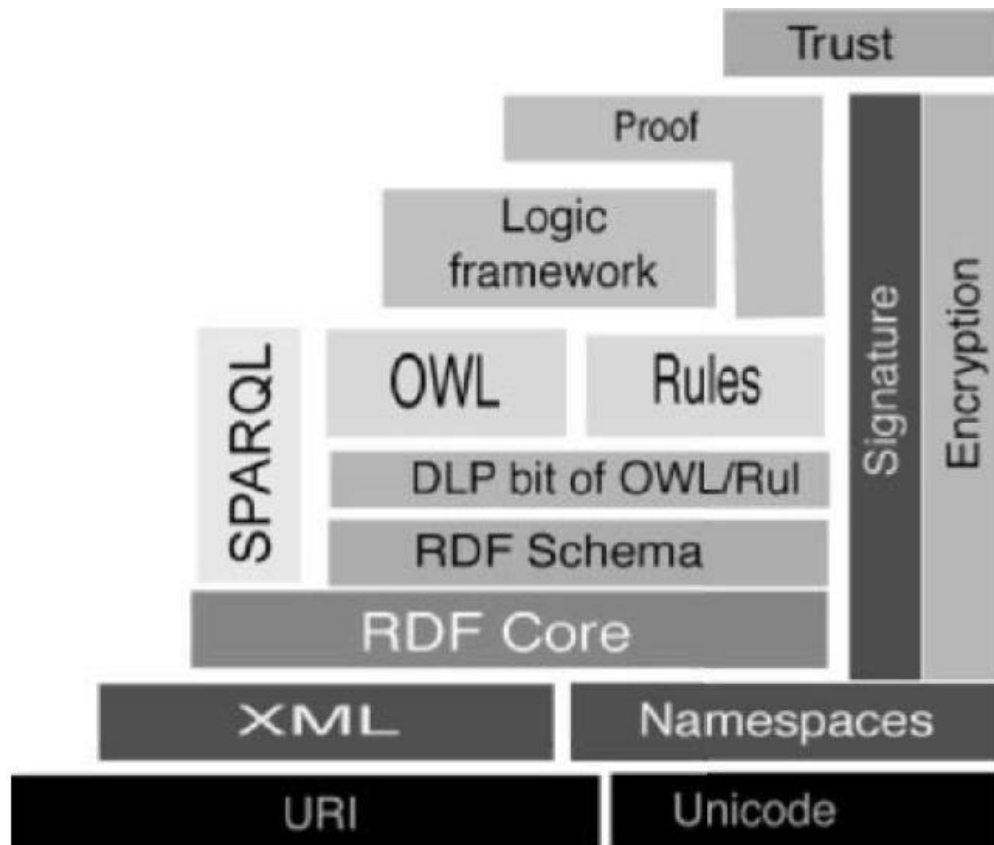
Figure 2.7:1 Semantic Web Stack (Antoniou & Harmelen, 2008, p. 20).

2.8. Recent semantic web (RDF) research works

A number of researchers were conducting several studies about the semantic web RDF data modeling. For instance, Carminati, Ferrari, & Thuraisingham, suggests new approach for using RDF for policy specification and enforcement that "exploit RDF model to allow the specification of high level policy and to automatically entail all the implied authorizations for specific scenario" (2004, p. 5). Whereas, Han, Finin, Parr, Sachs, & Joshi, developed an application called RDF123,which is an open source tools that allows "users to define mappings to arbitrary graphs" using "graphical application and web service" (2008, p. 1).

On the contrary Allemang & Hendler, provides a broad explanation and illustration of RDF data modeling using examples of how data can be distributed and merged using the RDF data modeling and describes the possible options for representing and relating several data from the database. Further, RDF triples and graphical representation is used

to migrated data from the table where the "cell is represented with three values i.e. the identifier for the row is called the subject, for the column is called predicate and for the value is called object"(2008, p. 50).

The study by Styles, Ayers, & Shabir, describes the use of RDF semantic web in order to represent and map the MARC, MARC 21 records and they use an algorithmic approach that allow them to transform the MARC record syntax into the web of data(2008). Their approach for mapping and representing the MARC records enables them to produce RDF based data.

There are many more studies in relation to semantic web technologies and techniques. RDF data modeling is then the foundation of the semantic web technology, which enables us to structured unstructured or semi-structured data in the form of triples and graphical representation in terms of nodes and edges. In addition, RDF is flexible, effective, and preferable language, which has a capacity to organize, gather, combine, and relate huge amount of data from various source of information over the web.

2.9.   The National standard for recordkeeping (Noark5)

Noark5 has been developed for all types of recordkeeping for both public and private bodies who wish to use a new electronic records management system or wish to implement a Noark5 based system. The Noark5 standard is obligatory within the public sector and is developed for electronic records management to ensure the functions and functionalities of the system and the intended electronic documents. In addition, as we earlier try to explain it in the introduction part of this thesis both the records management and the archives have more similarities rather than differences. In particular, in Norwegian both records management and archives do not have semantic difference. The general life cycle of the records are tied both of them, therefore the recordkeeping and archives are connected to each other from the creation of document, storing of records in record keeping until the disposal, preservation, and transformation of the records to the archives. Public administration of Norway uses the Noark5 solution and standard in order to track and ensure their electronic archives system and records.

In general, Noark5 contains various requirements, standards, and specifications and each record keeping system in the public bodies should or must fulfill all the obligatory and

optional requirements in order to be approved as a Noark5 solutions. At present-day, the Noark5 standard for electronic records management system is widely used across the whole Norway country.

2.10. The National archives and the regional state archives

The national archives and the regional state archives (statsarkivene) together form the national archival services of Norway (arkivverket) and are headed by the Director General (Riksarkivaren). The main job of the national and regional state archives is to preserve archive materials from state intuitions and to make this material accessible or available for use to the public or users, and to control the work on records administration within the ministries, regional and country authorities, and to contribute to the preservation of private archives. Further, the national and the regional archive organization are both administration bodies and scientifically based institutions for culture preservation.

In addition, the National Archives of Norway preserve the non-current records of government ministries, directorates and other central offices. Documents from these institutions are brought to the national archives when they are 25 years old and out of administrative use. According to the national archives website, there are more than 125 000 meters of documents, which are kept in the National Archives of Norway and each year the amount of documents increases by approximately 4000 meters. The oldest complete document is from 1189 A.D (The national archives, 2011).

The regional state archives (statsarkivene) aim to preserve documents from the regional and local branches of the state administration in their areas. In addition, the national archive of Norway receives archives from public administration bodies and supervises the archives creation done by public agencies. There are officially eight regional archives located in different part of Norway and a Sámi archives. The eight regional archives are located in Oslo, Hamar, Kongsberg, Kristiansand, Stavanger, Bergen, Trondheim, and Tromsø.

In general, the national archives of Norway aim in recording, preserving, and making data usable and accessible by end users. However, all digital or physical data are not

accessible some records that contains sensitive information are not reachable and accessible. Currently, the public administration and archive organization in Norway are digitizing their physical documents or records and in addition they are also capturing or receiving digital born documents in order to make data easily accessible by everyone (The national archives, 2011).
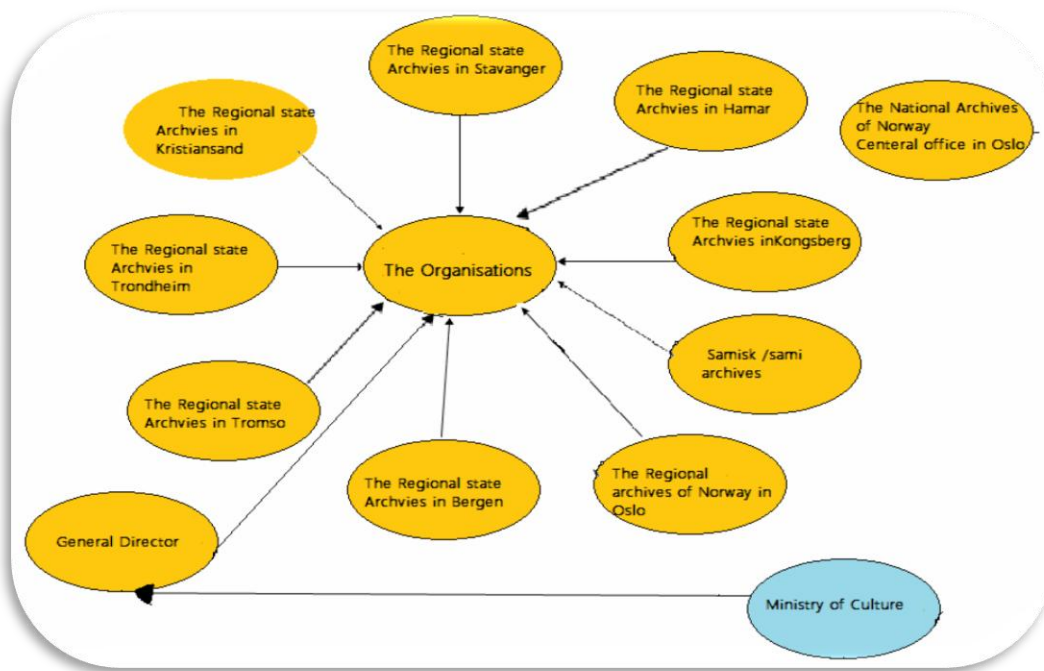
Figure 2.10:1 the eight regional state archives, sami archives, and the National archives of Norway are headed by the General Director. In addition, the Ministry of culture is responsible for the National archive of Norway.

2.11.   Summary

There are plenty of ongoing on projects and studies in relation to the importance of semantic web technology. Several scholars and practitioner within various disciplines are able to address distinct problems and issues using the semantic web technology. For instance, the study by(Antoniou & Harmelen, 2008; Davies et al., c2003) the semantic web and knowledge management and the semantic web and digital libraries by (Barbera, Cortese, Zitarosa, & Groppo, 2009; Kruk & McDaniel, 2008). In addition, semantic web

and social networks by (Peter, 2007) and the semantic web and digital music by(Bonanos, 2009; Engels, ESIS, & Tønnesen, 2007; MusicOntology, 2010; Raimond, Abdallah, Sandler, & Giasson, 2007), and the semantic web and digital archives by (Mazzini & Ricci, 2011) are some of the recent works about semantic web with other arena. In addition, the most research result shows the importance of semantic web technologies as well as issues regarding compatibility, flexibility, and interoperability, which can be used in relation to all data and information aspects.

In general, the literature review section covers studies that are relevant to the research theme and more or less the theoretical framework, the background, and the research that have already been undertaken and all indicates the merit of semantic web technology. Various literature reviews, chapter, books, and other related facts are collected from Oslo and Akershus University College of Applied Sciences library database and learning centre, Google scholar, Google search engine, and other relevant input sources. Furthermore, keywords such as *archives, records management, electronic records management, semantic web, RDF, knowledge representation, Noark, semantic web technologies,* and some others were used when retrieving relevant sources from these databases. Therefore, based on the above studies and evidence-based theories, the new emerging semantic web technology is able to influence and change most of the knowledge society activities in a positive way. In addition, libraries, archives, organizations, companies, non-governmental organizations, and others are making use of such new technologies in order to enhance the accessibility of their resource by their customers. However, there are also researchers who address that implementing, integrating, data in general using semantic web technologies in the existing applications or software that are already in use will require a lot of time, money, and cost.

Finally yet importantly, those who entitled themselves as information providers or information suppliers such as from the library, archive, museum, to the commercial and non-commercial industries and companies, which aim to serve users' should continuously improve and maintain their organizational culture and context in order to provide seamless and endless service to their users. Accordingly, it is true that the fast growth of technologies including the semantic web technology could require a great investment in time and money. However, the main point should not be only limited to the

time, effort, and money that are going to be paid/spend but rather it would also be good to focus the main target of the service, which is users and their needs and expectations.

In addition, at present there are many options of developing and using lightweight open source applications, therefore this in general would be able to save more money and time in particular can enhance and facilitates the use of data. The semantic web is more suitable with any applications and it allows anyone pretty much to say anything about something.

However, when it comes to semantic web and digital archives and or semantic web and ERMS very little work has been done. Therefore it is possible to say that there are no other research works in relation to these fields except the two research results by (Mazzini & Ricci, 2011; Kruk & McDaniel, 2008).

Therefore, the researcher attempts to use and integrate the semantic web technology for the ERMS/archives based on the importance of semantic web, which has been used for facilitating and overcoming several problem of several applications therefore this technology can also have advantage and importance into the area of records management and archives. In the following chapter, the essence of semantic web technology for ERMS and on how the semantic web technology benefits the Noark5 record keeping standard will be covered.

Chapter Three

3. Methodology

3.1. Purpose of the Research

The main objective of this research is to explore and find the possible integration points of the semantic web technologies for records management standard Noark5. In addition, the study involves learning and understanding the Noark5 (Norwegian archive standard) specifications, requirements, and standards. The study will describe, identify, and discuss the means to use and integrate semantic based services with records management. In this section, we will explain the research approach, research design, research strategy, the conceptual model of Noark5, proposed framework, limitation of the study, and ethical considerations.

3.2. Research approach

Research paradigm is considered as an essential part of the research process and serves for the researcher as guidance to his/her research issue. Pickard, in her recent book of research method in information stated that research paradigm "comes before the theoretical perspective of the research, it is 'the world view' that is accepted by members of a particular scientific discipline which guides the subject of the research, the activity of the research, and the nature of the research outputs" (2007, p. xvi).

Interpretive research is one of the research paradigm types, which enable researcher to view the existing meaning, and constructed meaning accordingly. The interpretive research has two distinct groups that are 'empirical interpretivism' and 'critical theory'. The empirical interpretivism "deals with investigation in natural settings of social phenomena" and the critical theory deals with "ideologically oriented investigation, examining current thought and social structures" (Pickard, 2007, p. 11).

The research employs an interpretive research approach in order to get an in-depth understanding of the Norwegian Archival Standard (Noark5). In this context, the use of such paradigm would enables us to have detailed information and knowledge about the

current archive organization system and their contents, as a result the researcher would have possible options to construct and interpret the meaning and reality of the Noark5 system, function, requirement, and standards using the available documents and information. In general, "Interpretive research assumes that access to reality (given or socially constructed) is only through social constructions such as language, consciousness, shared meanings, and instruments" (Myers, 2008). For that reason, the study prefers to use an interpretive approach for better understanding of the Noark5 context and structure.

3.3. Research design

Research design continues to be used as a means for exploring and solving societal problems. Researchers, academic disciplines, and other scholars have been employed multiple methods and methodologies that help them to examine and address their research issues. The most known methodologies are qualitative, quantitative, and combination of both. However, according to Pickard, "research methods cannot be limited only to the qualitative or quantitative methodology because methodology is perspective, the angle the researcher wishes to take on the question being asked" (2007, p. xvi).

This indicates that, there is no limitation to just one design or method, a researcher can imply and use various research approaches based on the issue or problem she/he want to address. However, nowadays the three major research methodologies are widely used.

Qualitative research is "non-numerical terms that aim to describe or interpret social processes or human condition. In addition, qualitative research focus more on how humans understand, experience, and interpret the world." Whereas quantitative research "aims in providing scientific knowledge about various social phenomena which is more data centered"(Keeran & Levine-Clark, c2007, p. 89). Pickard also address in similar way but within the research questions for instance 'how many', 'how often' and 'when' falls in the quantitative angle whereas 'why' and 'how 'kind of questions falls in the qualitative angle (2007, p. xvi).

This can show that both methodologies are far from each other and aims to address different issues from different perspective. In this case, qualitative research approach is preferable because the qualitative approach believes in social constructions of reality whereas the quantitative approach believes in objective reality of social facts (Gorman and Clayton as cited in Pickard, 2007, p. 13).

Moreover, the qualitative research strategies are purposely designed to help the researchers understand the social and cultural context within which people live, what they say and do. This indicates that, the qualitative research enables researchers to understand and see the *context* within which decisions and actions take place (Myers, 2008, p. 89).

The qualitative methodology as research methodology and an exploratory case study as research method has been chosen. In this case, for this research purpose several documents, records, and text for instance the Norwegian archive standard requirement and specification, extracted files and records, the Noark5 metadata, the national archive of Norway website, texts, and other related sources will be use as main input source.

3.4. Research strategy

The research theme by its nature seeks to understand a specific context within a specific purpose, for that reason, the researcher employs the case study as a research strategy to conduct this research. Case studies are the mostly and extensively used strategy particularly in the field of social science for instance in psychology, sociology, anthropology, political science, social science, and employment relations. Basically, the case study concentrate in a single object that can be a group of people, individual, program, an organization or member of the organization for instance departments, types of employee, customer or clients and so on (Cassell & Symon, 2004). In addition, Pickard depicts the case study that allows the researcher to gain "in-depth knowledge of the specific through rich description situated in context. This may lead to an understanding of particular phenomenon" (Pickard, 2007, p. 86).

In addition to that, case studies require "detailed investigation, often with data collected over a period of time of phenomena, within their context. The aim is to provide an

analysis of the context and processes which illuminate the theoretical issues being studied" (Cassell & Symon, 2004, p. 323).

Generally, as we can see the aim of case study strategy is to find, investigate, and understand a particular phenomenon within a particular context, purpose, and time limit. In this case, the Norwegian archive standard (Noark5) was chosen as a case study for two reasons. The first reason is because the researcher is more interested in knowing and understanding the process and function of records management standards and the second reason is that the Noark5 Norwegian archival standard is currently the de facto standard, which is widely used in Norwegian public administration in order to approve and maintain the electronic records management system.

Noark5 was built as specification of requirements for electronic recordkeeping systems for public administration and is used as de facto standard across Norway. The Norwegian archive standard contains multiple functions, recording keeping requirements, specifications, and other obligatory Noark5 components for electronic records management system. In this case, the study attempts to looks at the specific part of the Noark5 that is the Noark5 inner core and the Noark5 outer core. For that reason, a case study is more suitable and preferable to investigate, understand, and conduct specific case, within specific purpose and given time.

3.5.   The conceptual Model of Noark5

The conceptual model in (Figure 3.5:1) shows the present Norwegian electronic records management components. The Norwegian National Archives that developed the Noark5 standard concerns itself both with the classical paper based records as well as electronic records. In this case, the main objective of the research is to explore and understand the general requirement and specification of Noark5 that have been developed since 1984 as a standard to maintain and approve the electronic records management system. Therefore, the (Figure 3.5:1) illustrates the procedure and structure that are used as a means to capture, to describe, and to provide the electronic records.

Consequently, the recent publication of Noark5 noted that as a rule the public bodies must use a Noark-approved system to maintain their registered documents and archives (Noark-Arkivverket, 2011). The Noark5 standard requirement has several elements and

each element has their own obligatory requirements, and specifications that are must be fulfilled in order for a system to be approved as a Noark5 solution. The below (Figure 3.5:1) is then depicts the structure an ERMS as well as system for both electronic and physical records are obliged to follow. The Noark5 standard and requirement in addition to that, metadata is used as a means for describing the captured record. Therefore, generally the Noark5 inner core facilitates the document capture, retrieval, periodistation, preservation/disposal, and transfer. The metadata directory is used to support the data description, search and retrieval, exchange, and migration between different systems. Therefore, the conceptual model of Noark5 shows the overall process and function from capturing to assigning all necessary requirements. The "fonds" is the highest level of Noark5 requirement, which contains captured documents and the "series", which group the fonds with the help of classification system and class. In addition, the *file, record, document description, and document object* are part of the conceptual model and are the basic record structure components of Noark5.
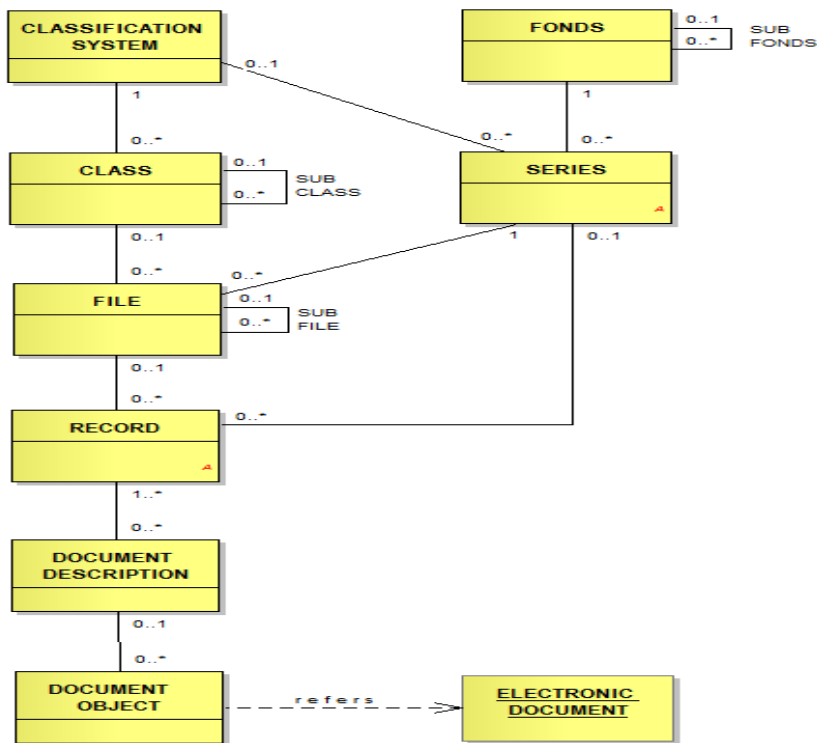
Figure 3.5:1 Conceptual model of Noark5

3.6. Proposed framework

The following proposed model provides an overview of the record keeping standard (Noark5) its structure and function as well as the semantic web component. Consequently, the proposed framework describes the methodological, technical, and logical approach of the study. In addition, this model is used for elaboration and guidance purpose of the research work, which is integrating and modeling Noark5 records management standard and data, using the Resource description framework (RDF) along with RDF visualizations. The Noark5 standard is logically classified into three parts that is Noark5 inner core, Noark5 outer core, and Noark5 complete but technically, they are not classified. Each core has its own functions, requirements, and specifications in regards to the electronic records management system and the actual records. Noark5 system can capture and store metadata relating to both paper based as well as digital resources.

As a result, every document or record that are created during organization activities must be registered or assigned necessary metadata element and in general, public bodies or private sectors within Noark5 system are obliged to fulfill the requirements, specifications, and regulations of Noark5 standard in order to ensure the quality and authenticity of records.

Therefore, during the modeling procedure several primary and secondary sources are used as an input such as the Noark5 metadata elements, Noark5 records structure, Noark5 extracted XML data, websites, and other related factors. In addition, the record structure of Noark5 contains the basic elements such as fonds, series, records, and files.

According to the recent publication of Noark5, *fonds* contain documents that are created, produced, or received by an individual fondscreator and collected as a result of this person's work or activity whilst the *series* is an arbitrarily defined section of the fonds entity where all materials are subdivided and organized according to the single primary classification systems. In addition, the fonds section is often determined to be identically to a record series, but need not necessarily be defined in this way (Noark, 2012, p. 43-44).

Respectively, the *record* type is known as a collective term for the records units' simplified record, including the specializations basic record, registry entry and meeting record. Whereas the *file,* contains a group of document that belongs together in some way, in addition the documents in a file should represent an instance (i.e. an execution) of an activity along with beginning and end (Noark5, 2012, p. 58-66).
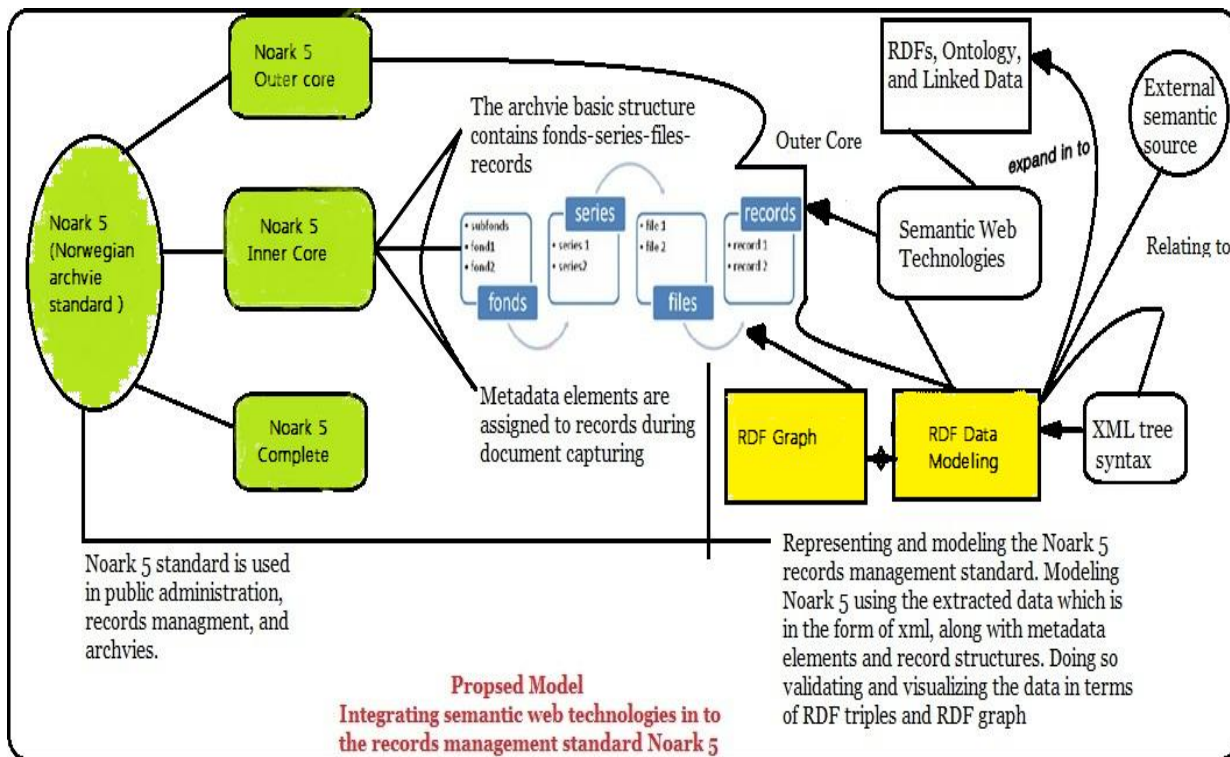
Figure 3.6:1 Proposed Framework of the Noark5 records management (the Noark5 contains the inner core, outer core, and Noark5 Complete), the archive structure, and the semantic web component contains (RDF data modeling, RDF graph modeling, and others).

On the other hand, semantic web technology contains the RDF, RDFS, Ontology, and LOD components and is used for defining, representing, and linking data (see Figure 2.7:1 semantic web stack for further illustration). Based on that, this research work uses the semantic RDF data modeling to represent and model the records management standard as the RDF is the foundation of semantic web technologies, which allows users

35

to model data in terms of triples. The main objective of this study is then to look at the benefit of semantic web for the records management standard within the RDF triples and graphical visualization. The framework will be use to draw the researcher into the right direction of mapping and modeling process. For that reason, Noark5 extracted XML will be parsed into the RDF data model using the RDF/XML serialization and 3-tuple or triple representation of the corresponding data model and the graphical visualization of the data model will be produced accordingly. Therefore, the initial starting for combining the ERMS and semantic web technologies as well as for modeling the Noark5 extracted data is the proposed framework. As a result, the overall creation of RDF data model could then enable researcher or experts to make use of additional and more advanced languages of the semantic web such as ontologies and LOD.

3.7.   Proposed Prototyping

The proposed prototype shows the hierarchical path of the research. The aim is to achieve a semantic based Noark5 standards and records. Therefore, the layer depicts the possibility to represent, map, and link the archive structure, Noark5 inner core, Noark5 outer core, and its intended metadata elements into the form of semantic web. In addition, the proposed prototype will have main significance throughout the modeling process in particular when it comes to maintaining and ensuring the traditional archive quality such as provenance and context. In this case, the bottom to the top hierarchy show that the existing and new layers can be combined altogether without overlapping or converting the existing elements with the primary focus of preserving and or retaining the traditional archive quality such as provenance and context.

Therefore, the layers that are located at the bottom show the overall Noark5 archival functions and structures. Whereas the two top layers are the proposed data model, which indicates the use of RDF modeling, representation, and graphical visualizing for representing data. The research use the pyramid shape, as it helps to show the interconnection and relationships with the Noark5 on the bottom and semantic web (SW). In addition, the pyramid is defined as a shape, which is used in thesis in order to build an argument for example progressively from a basic general premise and is mean to increase rapidly and on a widening base (Thefreedictionary, 2012). For that reason, it is very

important to lay the research work using the proposed framework and proposed prototype.
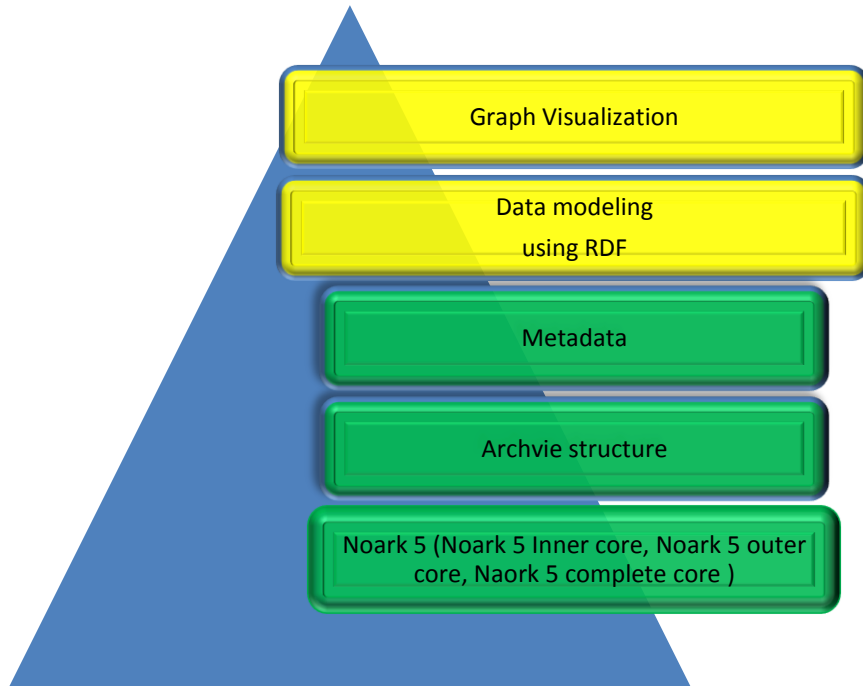
Figure 3.7:1 proposed prototype for semantic web integration into Noark5

3.7.1. *Operational definitions of terms*

Noark5: Noark5 is national standard for record keeping

Noark5 Inner core: contains requirement for basic functionality for recordkeeping and archiving

Noark5 outer core: contains requirement for external system

Metadata: 'data about data'

Resource Description Framework (RDF): is a framework for representing information on the web. (W3C, 2004d)

RDF triples: representing data in terms of (subject-predicate-object)

RDF graph: representing triples in terms of graph (node-arc-node)

Extensible Markup Language (XML): is a simple text-based format for representing structured information

3.8. Limitation of the study

There are some challenges and constraints must first be addressed. First and foremost is that understanding, learning, and implementing the overall process of the study require a lot of time, and effort. In particular, learning, defining, investigating, and exploring two new and distinct fields i.e. Noark5 ERMS and semantic web technologies to a level that gave the required knowhow and knowledge. Moreover, the RDF implementation and mapping of the Noark5 records keeping standard did not make use of neither Noark5 data nor Noark5 system, as there shall be several issues on front for instance asking and requesting permission for accessing and using the Noark5 based Electronic record or system, waiting for approval, and other related issues that waste time. For that reason, the researcher uses the available primary and secondary data source of the Norway public administration, Noark5 extracted data and others for the sake of this research. In addition, the other main reason is that the research has time limitation; there are not enough time to cover all aspects of electronic records management (ERM) and archives standards and systems. For that reason, it is preferable to introduce semantic web applicability and capacity for Noark5 using the Noark5 extracted records and standards. The other main constraint is that, the traditional archive provenance and the context has its own impact for semantic web, as it is not possible to turn and convert the overall archive or ERM traditional structure and context into the semantic RDF. Therefore, if we want to make use of the semantic web technology, then it is very important to assured that the traditional quality of the archives structure is maintained and preserved accordingly. Therefore, this will have its own constraints and may bring new challenge for semantic web.

3.9.   Ethical considerations

The research work did not make use of the actual Noark5 data or the real database instead uses XML records, documents, and other related important notes. In addition, the modeling and representing process has been done using the Noark5 open source software, which is developed by the department of information science and the Noark5 extracted data. However, all the input sources, the open sources software, and related data are only used for the sake of this study.

3.10.   Summary

To conclude, the above research process attempts to provide necessary and detailed explanation about the researcher choice of methodology, research paradigm, research method, research strategy, and research design. In addition, the methodology section introduces new technical and theoretical approaches for the purpose of modeling and representing the Noark5 data. For that reason, framework of the Noark5 records management (Figure 3.6:1) and prototype (Figure 3.7:1) for RDF semantic integration are the two approaches that are proposed for use throughout the thesis implementation. In addition, the study will explore the possibility to use the proposed model and methodological approach for addressing the statement of the problem, and the research questions in parallel. In addition, the interpretive research paradigm and research strategy are considered as an essential part for this research work. Therefore, based on the methodology part the researcher will have better understanding of a given case as well as will be able to achieve the intended objectives and aims of the study.

Chapter Four

4. Data modeling

4.1. Norwegian Archival Standard (Noark5)

Noark5 was developed as a specification and requirement in order to ensure and approve the quality of electronic record keeping system in public and government administration. The Noark5 recordkeeping standard must be use in public administration, archival organizations, records keeping managements, and other private sectors. Most of the public and private records keeping sectors in Norway are currently using the Noark5 standard.

When the record keeping function handles multiple tasks such as recording, storing, and retrieving data in authentic form then the national standard for recordkeeping Noark5 will ensures if the record keeping functions handle its task in appropriate way. The Noark5 specification and standard is classified into three parts, requirements for the inner core, requirements for the outer core, and requirements for a complete Noark5. This specification has been done purposely in order to avoid creation of two sets of Noark5, which is, "one set of requirements for an archive which can be used in any environment and one set of requirements for a complete, independent system" (Noark5, 2012, p. 24).

Therefore, the Noark5 inner core contains requirements for basic functionality of recordkeeping and archiving while the Noark5 outer core requirements contains requirement for the external module/systems. Accordingly, the complete Noark5 contains requirements and recommendations for optional task and administration systems that will form part of a "complete" Noark5 solution. In general, the requirements for the inner core and the requirements for the outer core are must and need to be fulfilled in order for a system to be approved as a Noark5 solution (Noark5, 2012).
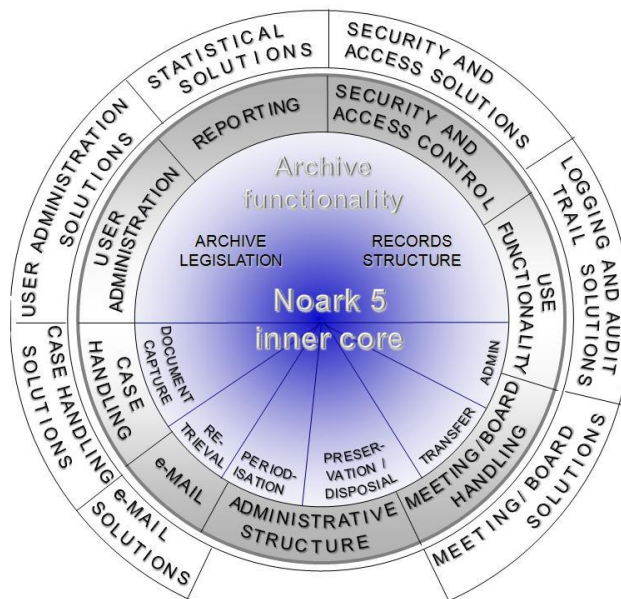
Figure 4.1:1 An overview of the Noark5 Standard for Records Management (Noark5, 2012).

4.2.   Noark5 Inner and Outer core

4.2.1.   Requirements for inner core and outer core in a Noark5

The Noark5 contains requirements for the inner core part of archive functionality and requirements for the laws and regulation for appropriate archiving. Noark5 inner core has record structure and archive legislation as archive functionality and document capture, retrieval, preservation and disposal, periodisation, transfer and administration of the core as modules. On the other hand, the Noark5 outer core contains requirements for the optional or external systems based in the regulations and laws in relation to the archive (Noark5, 2012).

4.2.2. Record structure

It has been tradition of collecting, selecting, sorting, organizing, and classifying records in accordance to their types and their contents. Physical records are organized and arranged in shelves, folders, files, cabinets, or archive boxes for better accessibility and reachability of the records so that users will be able to find the type of physical records easily. The same applies to electronic records i.e. records are organized, selected,

classified, and assigned necessary metadata and classification types in order to retrieve well-organized data. The main similarity between the physical and electronic records is that they organize and sort their records accordingly. However in electronic records, documents or records is assigned metadata, which is data about data that is used to provide detailed information about given records. On the other hand, the other difference between the electronic and physical records is in terms of their accessibility, anyone anywhere and anytime can access the electronic records. Therefore, in any of those cases the record structure is the main inner arrangement of the archives and ensures the archive records accuracy and efficacy. The following XML format shows the overall archive structure in Noark5 based system.

```
<fonds>
        <series>
                <files>
                        <records>
                                <documentdescription>
                                ……………………………………
                                        <documentobject>
                                        ----------------------------
                                        </documentobject>
                                </documentdescription>
                        </records>
                </files>
        </series>
</fonds>
```

Figure 4.2.2:1 the Noark5 Archive structure in XML format

4.3. Archive structure and metadata elements

The following basic archive structures are taken from the Noark5 standard for records management document English version. Each element has their own terminologies and

definitions in regards to the way the document organize, registered, and handled. Some of the basic records structure are mentioned below.

❖ FONDS

Fonds consist of documents that are created, received, or produced by individual fonds creator as a result of the fonds creator work or activity. Fonds are the highest level of the record structure, which contains the captured document by organizations and or individuals. As Noark5 requirement and specification, the fonds creator is responsible for creating their own fonds or fonds entity. In addition, fonds creator can be anyone individual, organization, foundation, and companies. Further, in Noark5 it is possible to create one or more fonds entities by members of an organizations for instance the head office and regional office can also setup their own fonds entity and can have the possibility to share the same fonds.

❖ FONDSCREATOR

Fonds creator is an organizational unit or person who creates fonds as part of his/her activity or work. A fonds creator might be public bodies, a company, an organization, a foundation, etc. or a part of such a unit. A public body may be one fonds creator and therefore have one fonds entity (central registry), or it could have several fonds creators (departments, agencies, etc.), each of which create their own fonds entities (partial fonds).

Series

In addition, series is the next part of fonds entity in which all material, is subdivided and arranged according to a single primary classification system.

A grouping of fonds (entity) constituted by a common classification scheme. Its definition is often, but not always, identical to that of a fonds series

❖ FILE

A level in the fonds structure, a record unit. One or more fonds and associated fonds documents that are linked together under a common identity.

❖ CASE FILES

A specialization of the record unit file in the fonds structure

❖ RECORD

Document created or received by a person or organization as part of the activity which is maintained by the person or organization

❖ REGISTRY ENTRY

An individual record (entry) in a registry, i.e. the information on a case document and any attachments.

(Noark5, 2012, p. 44)

4.4. The overall archive elements and metadata

Fonds elements	systemID, Title, description, fond status, documentmedium, storagelocation, createdDate, createdBy, finalizedDate, finalizedBy
Fondscreator elements	fondscreatorID, fondscreatorName, description
Series elements	systemID, Title, description, recordsectionstatus, documentmedium, storagelocation, createdDate, createdBy, finalisedDate, finalisedBy, recordsperiodStartDate, recordperiodEndDate, referenceParent, referencePrecursor, referenceSuccessor, referenceClassifcationsystem, refernceFile, referenceRecord
File element	systemID, fileID, filetype, title, officalTitle, description, keyword, documentmedium, storagelocation, createdDate, createdBy, finalisedDate, finalisedBy, referenceParent, refernceChild, referenceRecordsection
Case file elements	casedate, administrativeUnit, case-responsible, registrymanagementunit, casestatus, loanedDate, loanedTo,referenceSecondaryClassification,
Simplified record elements	systemID, recordtype, createdDate, createdBy, archivedDate, archivedBy, refernceParent, referenceRecordsection, referenceDocumentdescription, refernceDocumentobject,

Basic record elements	recordID, title, officalTitle, description, keyword, author, documentmedium, storagelocation
Registry entry elements	Serialnumber, registryentrytype, recordstatus, registryDate, documentDate, receivedDate, sentDate, duedate, confidentialityassessedDate, numberofAppendices, loanedDate, loanedTo
Client elements	clienttype, clientName, postaladdress, postcode, postaltown, foreignaddress, emailaddress, telephonenumber, contactperson
Case-responsibility elements	administrativeUnit, executiveofficer, registrymanagmentunit
Documentdescription elements	systemID, documenttype, documentstatus, title, description, author, createdDate, createdBy, documentmedium, storagelocation, refernceDocument
Documentlink element	refernceRecord, linkedRecordAs, documentnumber, linkedDate, linkedBy,
Case party element	casepartyID, casepartyName, casepartyRole, postaladdress, postcode, postaltown, foreignaddress, emailaddress, telephonenumber, contactperson
Documentobject element	systemID, versionnumber, variantformat, format, formatDetails, createdDate, createdBy, refernceDocumentdescription, refernceRecord, refernceDocumentfile, checksum, checksumalogrithm, filesize

Table 4.4-1 the overall archive structure, context, and their metadata elements

4.4.1. Metadata elements

In Norway public administration metadata elements are used as a means to provide detailed information of a given document or resource in similar way as others. In addition, it can be used for the purpose of transferring and exchanging data with in applications and databases. However, there are some exceptional cases when it comes to

the names of metadata for instance the names should be composed of ordinary Norwegian terms, and be as intuitive as possible. In addition, the metadata names must not contain numbers, spaces, or other special characters and always metadata names should start with small letters. Moreover, special characteristics such as Norwegian letters æ, ø, and å should not be used in names. This kind of characteristic is not valid in the XML syntax therefore they should be converted in another format for instance in the following pattern: æ> ae, ø> oe and å> aa. Some operational definitions and terms are translated from the Norwegian language into the English language using the Google translator in order to have brief overview of the metadata elements as following.

4.4.2. Operational definition of the metadata elements

❖ SystemID

SystemID is a unique identification of the archival unit. All archival units except document object contain unique system identification. Document object has no system identification because the document object can be duplicated during data or archive extraction, particularly if the document file is linked to several different registration items. Therefore, each unit records in the archive structure have a systemID, for instance fonds, class, record, document description etc.

❖ ClassID

ClassID is a unique identification of the class within the classification system. Other classification systems within the same file system may contain one or more of the same identifiers. Identification can be purely numerical, but can also be alphanumeric and have a logical meaning. Note that classID is identical to the terms of order value and the file code in Noark4. All classes in a classification usually created when a file system is used. However, some solutions may allow the creation of new classes as needed (most relevant at the object-based classification).

❖ FileID

FileID is a unique identification of the file within the archive file belongs. Other records in the file system may contain one or more of the same codes. The code can be purely

numerical, but can also have a logical structure. Note that, fileID is identical with the combination of case number and sequence number in Noark4.

❖ CreatedDate

CreatedDate within date and time format is assigned to archival unit or records during creation or registration period. All archive units are automatically assigned createdDate.

❖ CreatedBy

CreatedBy is name of person who created / registered the archive unit

finalisedDate

❖ FinalizedDate

 FinalizedDate is assigned when the record unit is completed or closed. FinalisedDate has the data and time format. Therefore, every record unit that is finalized must have createdDate.

❖ FinalisedBy

FinalisedBy contains information about the person name who completed or closed the records unit

❖ ArchivedDate

ArchivedDate has similar format as finalisedDate and createdDate contains date and time format. However, archviedDate is assigned to all the documents that are archived. Archived documents means that documents, which is *frozen* that cannot be alerted, updated, or changed.

❖ ArchivedBy

ArchivedBy is the name of the person who archived/filed the document and froze it for all further editing (The national archives, 2007).

4.5. Noark5 extracted files

The national archivist determines the time and frequency of transfer of archival versions from each system and determines whether the status of the archival version should be transfer or deposit, and may require depositing to take place prior to the time of transfer. Transfer would normally take place when the material is approximately 25-30 years old,

but the National Archivist may also provide for transfer at an earlier or later date, cfr. § 5-2 of the archival Regulation.

In addition, from electronic registry and recordkeeping systems, an archival version is to be deposited as soon as records period (usually a five-year period) is completed, cfr. § 3-17 of the Archival Regulation. Any associated electronic documents must be enclosed on deposit. Associated documents on paper must be enclosed as part of the normal transfer when the material is 25-30 years old. Therefore, during this transaction period electronic records must be extracted and transferred from the Noark5 recordkeeping system into the national archives (The national archives, 2007).

Accordingly, we used the Noark5 electronic document that was extracted for transferring the electronic records into the archives, which is Noark5/arkivstruktur.xml (XML format). The Noark5/arkivstruktur.xml content and structure is written in Norwegian language that is the native language of the country and the XML standard format has been used as a means to extract the Noark5 electronic format. See appendix 1 and 2 for data sample of Noark5 extracted records.

 Further, we used the http://ark1.hioa.no/n5.ui/   project website from hioa (Oslo and Akershus University College of applied sciences). The ark1.hioa.no/n5.ui/ is a project developed by the Oslo University College department of information science with the aim of producing open source Noark5 software. Therefore, the study used this project site and database as a source of information for this research work. The main objective of using this site is that because the current research results could serve as a basis for modeling the actual data of Norak5. The following (Figure 4.5:1) shows an example of the arkive structure and (Figure 4.5:2) show the print screen of the user interface of the Noark5 core.

```
<arkiv
 xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
 xmlns="http://www.arkivverket.no/standarder/noark5/arkivstruktur"
 xmlns:n5mdk="http://www.arkivverket.no/standarder/noark5/metadatakatalog">

 <systemID>403ac99f-1d20-438c-a398-824b32d045ff</systemID>
```

```xml
<tittel>Fonds</tittel>
<beskrivelse>Fonds</beskrivelse>
<arkivstatus>Avsluttet</arkivstatus>
<dokumentmedium>Elektronisk arkiv</dokumentmedium>
<opprettetDato>2011-10-11T16:05:28</opprettetDato>
<opprettetAv>system</opprettetAv>
<avsluttetDato>2011-10-11T19:26:36</avsluttetDato>
<avsluttetAv>system</avsluttetAv>
<arkivskaper>
  <arkivskaperID>964963193</arkivskaperID>
  <arkivskaperNavn>Test Name</arkivskaperNavn>
  <beskrivelse>Test Name Archive</beskrivelse>
</arkivskaper>
<arkiv>
  <systemID>29743769-edf4-46f5-be47-51a0167fbfeb</systemID>
  <tittel>Subfonds5</tittel>
  <beskrivelse>Subfonds5</beskrivelse>
  <arkivstatus>Avsluttet</arkivstatus>
  <dokumentmedium>Elektronisk arkiv</dokumentmedium>
  <opprettetDato>2011-10-11T17:32:44</opprettetDato>
  <opprettetAv>system</opprettetAv>
  <avsluttetDato>2011-10-11T19:26:36</avsluttetDato>
  <avsluttetAv>system</avsluttetAv>
```

Figure 4.5:1Noark5 extracted data in the form of XML see Appendices 1 and 2 for details
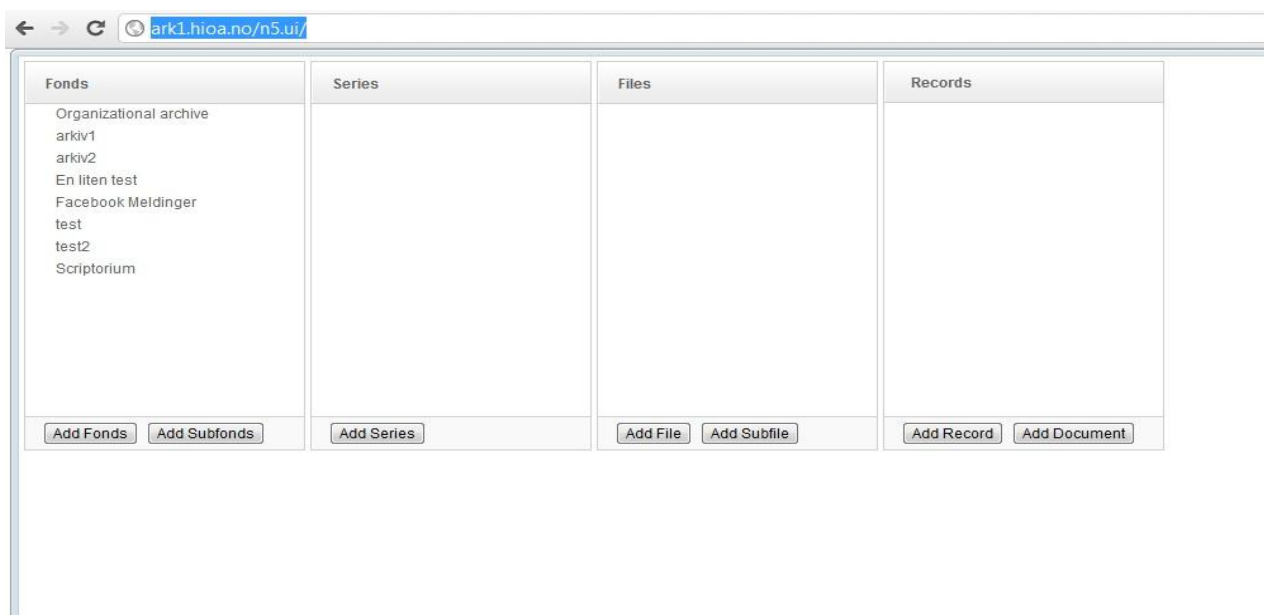
Figure 4.5:2 The Noark5 open source software its user interface

4.6.   Modeling procedure

The first footsteps for modeling Noark5 records using RDF semantic web is to identify what to model, why modeling, and how to model the data. Therefore, for that reason we decided to have a look at the Noark5 inner core and outer core standard and its intended metadata. The Noark5 has three levels and each level has its own functions, requirements, and specifications that must or should be use on the ERM and archival institution. The Noark5 inner core has a capacity to perform and handle organization archival records without depending on any external functions. In addition, the Noark5 inner core archive functionality contains two parts called record structure and archive legislation as well as modules such as document capture, retrieval, preservation and disposal, periodisation, transfer and administration of the core.

In general, "the inner core will handle the organizations archive, i.e. the fonds documents that are received or produced as a result of the activities carried on by the organization." then soon after the fonds document organizes in accordance to the general record structure. The record structure aims to organize, record, and maintain the records accordingly. The record structure is also known as the inner core arrangement of the

50

archive record and can contains large hierarchies and several levels from the top to the bottom. Therefore, the electronic documents or the fonds document must organize and store appropriately in the record structure. In addition, each record structure unit has its own metadata and it is possible to link top record units with its subsidiary record unit using metadata. Moreover, the archive structure contains detailed explanation about the overall conceptual model of the archive structure. The conceptual model shows the linkage between several metadata and their link between physical and electronic documents. Therefore, the main objective of this research is to represent and model the Noark5 inner core and outer core using the available input source from the Noark5 record keeping standard i.e. the Noark5 extracted files, Noark5 metadata elements, and archive structure.

4.7.   The proposed prototype and Framework

The proposed prototype and framework are used as a means for elaboration and guidance throughout the modeling and implementation procedure. In addition, it shows the archive components, semantic web elements and their importances. The framework shows the two distinct areas when they create a relationship and interconnection among them. As a result, the layers in the proposed prototype describe the possibility to represent and link the archive structure and metadata in the form of semantic web. See the proposed framework (Figure 3.6:1) and the proposed prototype in (Figure 3.7:1).

4.8.   XML representation of the archive structure

4.8.1. *Extensible Markup Language (XML)*

The XML language is a language that is designed for both human and computer consumption. The language also allows us to develop well-structured documents on the web and "supports the exchange of structured information across different applications through markup, structure, and transformations (Antoniou & Harmelen, 2008, p. 59).

In addition, the XML syntax enables users to define data in a structured way using different elements and attributes of XML. The following example shows the XML format that contains the fond elements from the Noark5/arkivstruktur file. The fond is the main root element and have child elements (systemID, title, description, finalisedDate, finalisedBy, documentmedium, createdDate, createdBy). In addition,

```
<? Xml version="1.0 encoding="UTF-16"?>
 xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
 xmlns="http://www.arkivverket.no/standarder/noark5/arkivstruktur"
 xmlns:n5mdk="http://www.arkivverket.no/standarder/noark5/metadatakatalog">
```
are an XML declaration which indicates the current document as an XML document and the next two namespace shows the XML based Noark5 recordkeeping standard records and record structures that contains the metadata elements, and the actual data. All the elements have "tag name" for instance   <systemID>403ac99f-1d20-438c-a398-824b32d045ff</systemID> and each elements and its data must always be enclosed with the opening and closing tags.

```
<?xml version="1.0" encoding="UTF-8"?>
<arkiv
 xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
 xmlns="http://www.arkivverket.no/standarder/noark5/arkivstruktur"
 xmlns:n5mdk="http://www.arkivverket.no/standarder/noark5/metadatakatalog">
 <systemID>403ac99f-1d20-438c-a398-824b32d045ff</systemID>
 <tittel>Fonds</tittel>
 <beskrivelse>Fonds</beskrivelse>
 <arkivstatus>Avsluttet</arkivstatus>
 <dokumentmedium>Elektronisk arkiv</dokumentmedium>
 <opprettetDato>2011-10-11T16:05:28</opprettetDato>
 <opprettetAv>system</opprettetAv>
 <avsluttetDato>2011-10-11T19:26:36</avsluttetDato>
 <avsluttetAv>system</avsluttetAv>
</arkiv>
```

Figure 4.8:1 the XML representation format of Noark5, see appendix 1 for details

Therefore, the overall extracted arkivstruktur (archive structure) looks like the above (Figure 4.8:1). Therefore, the research is going to use this XML based structure file for modeling and representing purpose. In order to proceed for data modeling let us first like to see the similarities and differences of XML and RDF as following.

4.9. The difference and relationship of XML and RDF

XML is a language that allows users to define markup for their documents using tags and supports information or document exchanging between distinct applications. However, XML is not concerned with the meaning or the semantics of data and this indicates that the XML data "does not have semantic definition that will enable the machine understand and process the meaning of the data"(Yu, c2007, p. 67).

On the other hand, RDF is a simple data model that is used to represent or model simple statement in terms of triples and has graph-based model. RDF takes advantage of XML syntax in order to support "syntactic interoperability". Therefore, this indicates that, "XML and RDF complement each other because RDF supports semantic interoperability" (Antoniou & Harmelen, 2008, p. 109).

In addition, there are more arguments and discussion concerning XML and RDF for instance, why we wish to use RDF, why not we keep using of the XML language as to write and structure data. The confusion is because both the XML and RDF languages aim to structure and create well-structured data. However, when it comes to understanding the meaning of the structured data XML does not address anything or make sense of what the content is or what type of document it is holding. Therefore, the advantage of RDF is then that it enables us to create structured and meaningful data so that both humans and machines can use and understand it in similar way. Let us assert the statement using an example one how and why RDF modeling differs from XML.

Example 1 XML and RDF/XML representation
<?xml version="1.0" encoding="UTF-8"?>
<arkiv beskriverlse ="arkiv"?>
  <systemID>438c-a398-824b32d045ff</systemID>

```
<tittel>Fonds</tittel>
<beskrivelse>Fonds</beskrivelse>
<arkivstatus>Avsluttet</arkivstatus>
</arkiv>
```

Figure 4.9:1 XML representation format of example 1
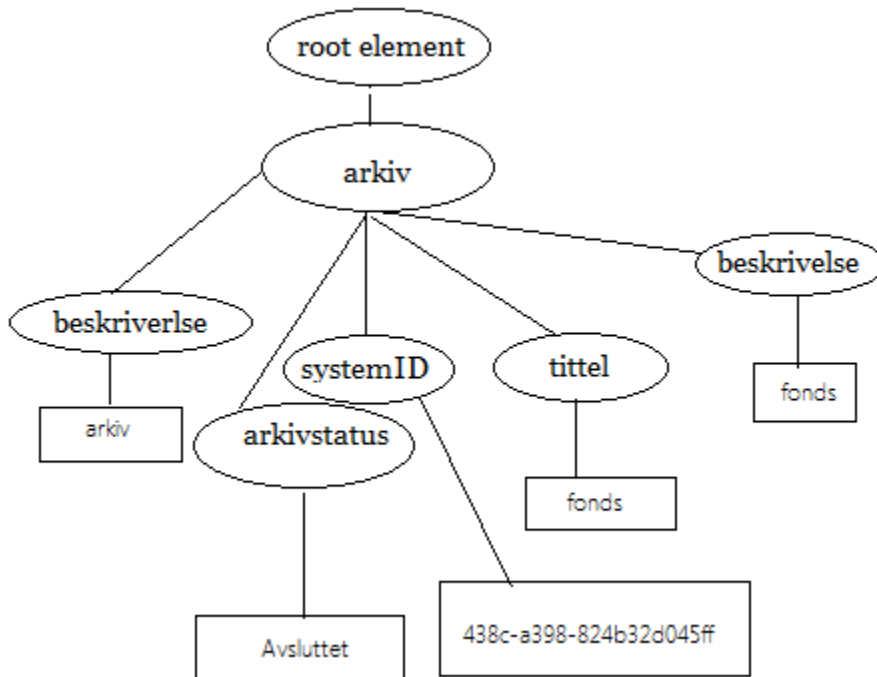
Figure 4.9:2 the XML graphical representation shows the hierarchical tree of the example 1 XML syntax

On the other hand, the RDF data model uses the subject-predicate-object triple format to represent meaning to data. Let us proceed to the next example of RDF.

```
<?xml version="1.0"?>
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
            xmlns:n5ui="http://ark1.hioa.no/n5.ui/" >
                <rdf:Description rdf:nodeID="arkiv">
                        <n5ui:systemID>438c-a398-824b32d045ff</n5ui:systemID>
                        <n5ui:tittel>Fonds</n5ui:tittel>
                        <n5ui:beskrivelse>Fonds</n5ui:beskrivelse>
                        <n5ui:arkivstatus>Avsluttet</n5ui:arkivstatus>
                </rdf:Description>
</rdf:RDF>
```

Figure 4.9:3 RDF representation format of example 1

54

The first line with <?xml version="1.0"?> indicates the XML declaration and the second line shows the RDF namespace, vocabulary, and URI. Further, we have new namespace for Noark5 that is called "n5ui" and http://ark1.hioa.no/n5.ui/ and is used for defining the Noark5 document. In addition, we use the RDF element rdf:Description in order to make statement of the given resource that is rdf:nodeID="arkiv". Furthermore, every RDF document should contain the <rdf:RDF> and </rdf:RDF> valid syntax.

The above (Figure4.9:1) and (Figure4.9:3) are a sample of XML and RDF/XML serialization and the following (Table4.9-1) show the RDF data model in the form of triples and (Figure4.9.4) and (Figure4.9.5) indicates the graphical representation within long namespace URI and the short namespace QNames. We used graphical representation for better understanding of the RDF/XML serialization and it is an optional graphical visualization of the actual data model.

Number	Subject	Predicate	Object
1	genid:Uarkiv	http://ark1.hioa.no/n5.ui/systemID	"438c-a398-824b32d045ff"
2	genid:Uarkiv	http://ark1.hioa.no/n5.ui/tittel	"Fonds"
3	genid:Uarkiv	http://ark1.hioa.no/n5.ui/beskrivelse	"Fonds"
4	genid:Uarkiv	http://ark1.hioa.no/n5.ui/arkivstatus	"Avsluttet"

Table 4.9-1 A 3-tuple (triples) representation of the corresponding example 1 RDF/XML data model

The above (Table4.9-1) illustrates the RDF triples data model of the above archive RDF/XML serialization. The RDF/XML is more convenient and is easy for machine to understand, trace, and read the data whereas for human consumption, the RDF triples and RDF graphic visualization is more suitable. Therefore, the triple always have the *subject-predicate-object* style, where the subject intended to describe or express the resource and the predicate, which is also known as property name used to create semantic relationships, and the object used to identify the value of the resource. In this case, the triples of the data model of an example 1 can be express for instance the first raw, column 1 describes "genid:Uarkiv" as the subject, "http://ark1.hioa.no/n5.ui/systemID" as the predicate within URI, and "438c-a398-824b32d045ff" as object. Most of the time the use

of long URI in predicate is not preferable, therefore, there is another possible option of using short namespace with prefix sometimes known as QNames. Therefore, the following graphic visualization shows the difference between the use of long URI and short namespace.
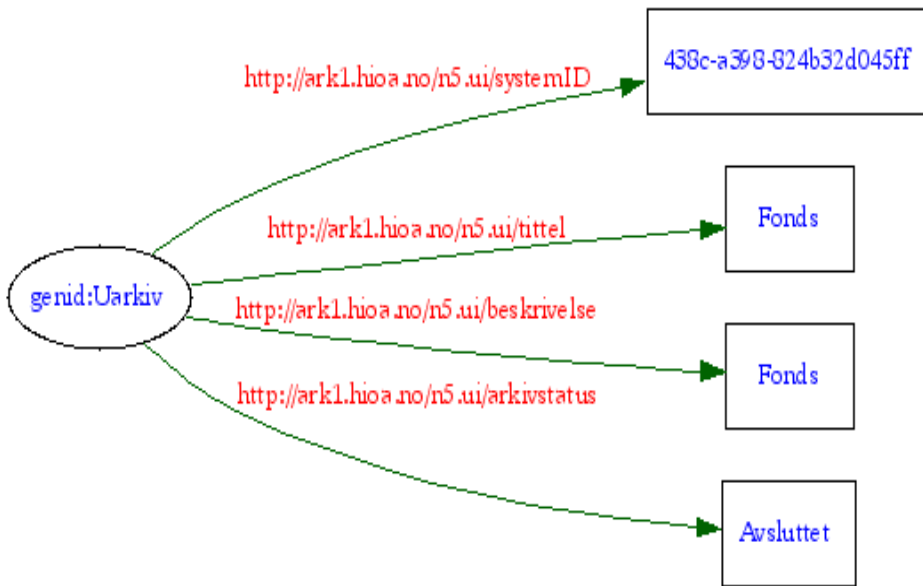
Figure 4.9:4 the graphical visualization of example 1 RDF/XML serialization with the long URI property
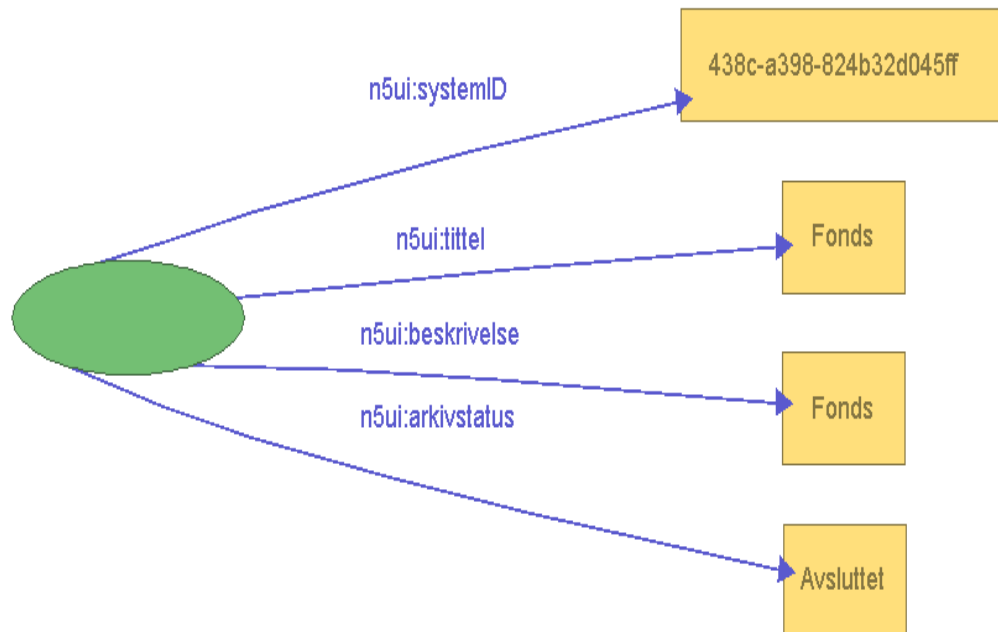
Figure 4.9:5 the graphical visualization of example 1 RDF/XML serialization with the short namespace (QNames)

Furthermore, both examples are used to illustrate and describes the differences between the two languages RDF and XML. In fact, XML is a meta-language is much known for writing and structuring data in a form machine can detect or read it and the RDF is a meta-language as well as a resource description framework for modeling of data.

Therefore, the example in general indicates the RDF and XML ways of representing, and parsing data, which is more interesting for anyone who concern with data, information, and knowledge of structuring, creating meaning, and representing.

In addition, the RDF triples representation and an RDF graphical visualization result varies from the XML way of parsing and XML tree representation. Therefore, in this early stage, in this section at least we are able to address and explain the differences between the XML and RDF and introduce the primary use of RDF and RDF namespaces Such as (xmlns:rdf) and (xmlns:n5ui).

XML	RDF/XML syntax

<?xml version="1.0" encoding="UTF-8"?>
 <!-- archive structure in xml format -->
 <fonds description="fonds">
    <systemID>4584d-6256a-4394704521</SystemID>
    <title>fonds</title>
    <description>new fond document</description>
    <fondsstatus>closed</fondsstatus>
    <documentmedium>electronic format</documentmedium>
    <createdDate> 2011-10-11T16:05:28</createdDate>
    <createdBy>Oslo municipality</createdBy>
    <finalisedDate> 2011-10-11T19:26:36 </finalisedDate>
    <finalisedBy> Oslo municipality </finalisedBy>
 <subfond>
    <systemID>4584d-6256a-4394704521</systemID>
    <description> subfond</description>
 </subfond>
 </fonds>
	<rdf:RDF
	    xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
	    xmlns:n5ui="http://ark1.hioa.no/n5.ui/" >
	    <rdf:Description rdf:nodeID="fonds">
	       <n5ui:systemID>4584d-6256a-4394704521</n5ui:systemID>
	       <n5ui:title>fonds</n5ui:title>
	       <n5ui:description>new fond document</n5ui:description>
	       <n5ui:fondsstatus>closed</n5ui:fondsstatus>
	       <n5ui:documentmedium>electronic format
	    </n5ui:documentmedium>
	       <n5ui:createdDate> 2011-10-11T16:05:28</n5ui:createdDate>
	       <n5ui:createdBy> Oslo municipality </n5ui:createdBy>
	       <n5ui:finalisedDate> 2011-10-11T19:26:36 </n5ui:finalisedDate>
	       <n5ui:finalisedBy> Oslo municipality </n5ui:finalisedBy>
	       <n5ui:hasPart rdf:nodeID="subfond"/>
	    </rdf:Description>
	    <rdf:Description rdf:nodeID="subfond">
	       <n5ui:systemID>4584d-6256a-4394704521</n5ui:systemID>
	       <n5ui:description>subfond</n5ui:description>
	    </rdf:Description>
	 </rdf:RDF>

Table 4.9-2 Example 2 data representation in the form of XML and RDF/XML

In addition to that, each element that are enclosed within open and end tag are now contains additional semantic based information that is <n5ui:systemID> "n5ui"is the added semantic value to the systemID. Therefore, this time both humans and machines are able to read and understand the well-defined RDF data and its meaning, for example the following triples are now understood as the resource *arkiv* have a *systemID* whose value is *438c-a398-824b32d045ff*. Moreover, example 2 expresses and indentifies more data elements and semantic relationships between the elements in the form of XML and RDF/XML serialization. (Figure4.9:6) and (Figure4.9.7), provides graphical visualizations of the XML and RDF/XML serialization.
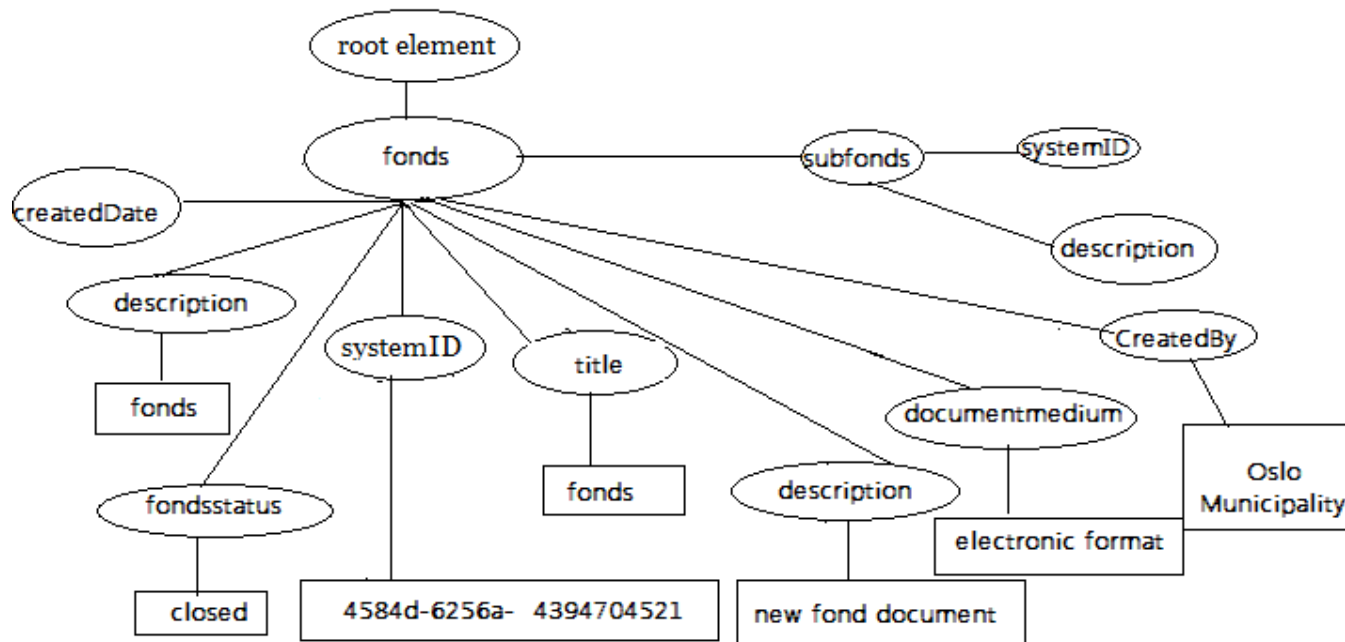
Figure 4.9:6 the XML representation format of example 2

The above (Figure 4.9:6) XML tree representation has one root element and child elements. Users' can write such well structure data using the XML meta-language. It allows users to write, access, and exchange the data within various applications. This has been the main advantage of XML language and is still in use.
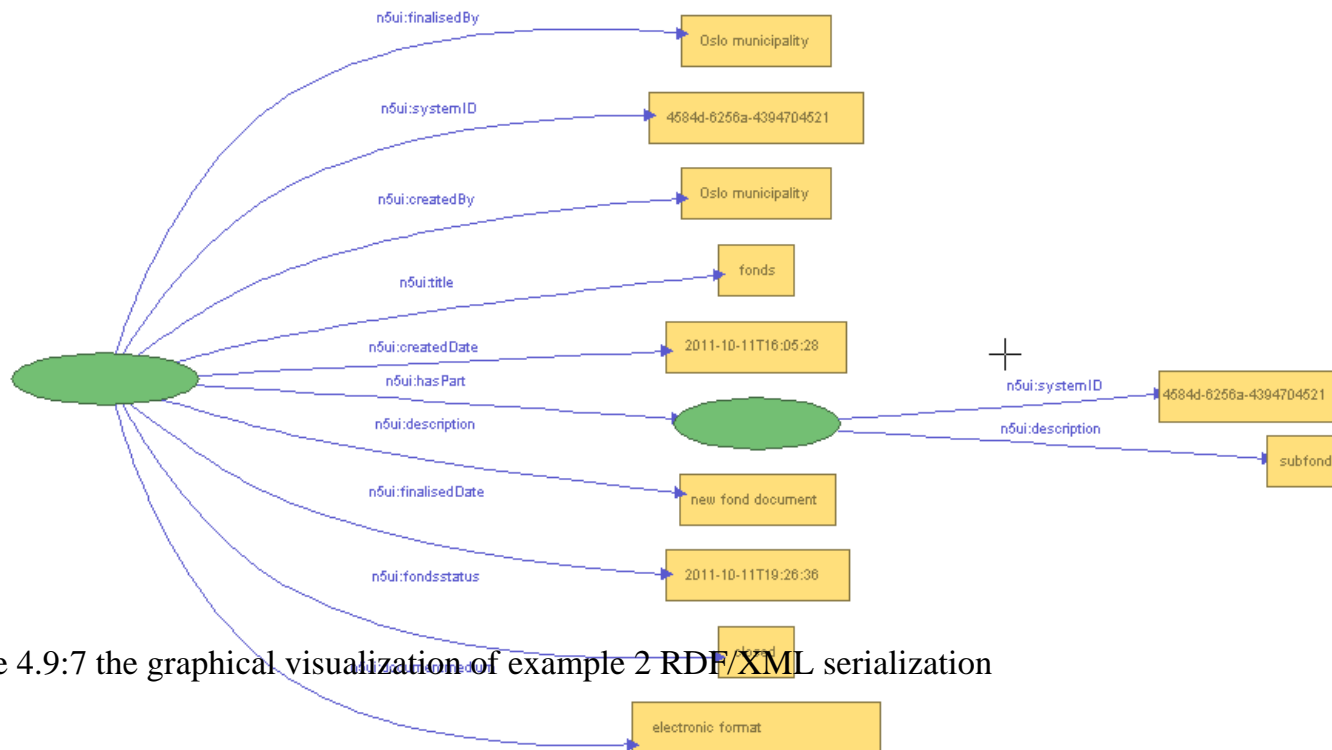
Figure 4.9:7 the graphical visualization of example 2 RDF/XML serialization

In general, to summarize the above example two is similar as example one, but contains more data and elements than the pervious one. As we earlier try to mentioned it the data modeling shows the main differences between the XML and RDF data modeling and cleared the question and confusion of "Why should I use RDF-why not just XML?" (Berners-Lee, 1998). In addition, the above graphical representation shows the added semantic meaning to the data and the semantic based elements are also in the form of triples, for instance all resource (subjects) are connected to its intended resource (object) using the URI predicate relationship. Example of data in triple format can be taken one from the table as following,

The subject (*Genid:Ufonds*) – predicate – (*http://ark1.hioa.no/n5.ui/systemID*) –object – ("*4584d-6256a-4394704521*"). At the time both human and machine can understand that the so-called fonds resource has the property name called systemID whose value is 4584d-6256a-4394704521. As a result, the following section will provide the data modeling of the Noark5 recordkeeping standard in detail using the Noark5 extracted file.

4.10. RDF data modeling

Recently some studies have been conducted in relation to the use of RDF and ontology languages as the researcher described it in the pervious literature review section. The RDF is the basis for the semantic web technology and is the essential part to start modeling using the triple rules and then it will be more convenient to expand the RDF data into other form of semantic such as ontology, which is a richer language than RDF that allows us to create more complex data representation and relationships. Therefore, RDF data modeling is used as a starting point for this study.

The pervious section shows a glimpse of the Noark5 extracted xml data that contains metadata elements and Noark5 data within the archive structure. As a result, in this section the mapping or representing the sample Noark5 record into the RDF data model has been done using the RDF/XML serialization syntax. Further, the overall modeling procedure has been done manually using the free source code editor such as Notepad ++, RDF editor, whereas the process of validating and checking the RDF/XML data is done using the well-known default W3C validation service. Furthermore, RDF visualizing

tools such as RDF service validator and IsaViz a visual authoring tool for RDF is used to generate the RDF/XML data models in terms of graphics.

In RDF, a resource or entity is identified using the global unique identifier uniform resource identifier (URI). A resource is something or anything that can talk about for instance chapter, webpage, people, and other real world objects. Therefore, anyone can make use the existing URI identifier or can create unique URI in order to address the resources. However, URI identifier is uses if and only if the name of the resource is globally known or identified.

In this case, we do have a resource with web identity, which is not yet globally known, or it is not visible to the outside world. Therefore, in order to proceed for data modeling we need to assigned blank nodes. Blank nodes are used to describe anonymous resources or a resource that do not have real resource identifier and can be used as object and as a subject in the RDF triples (Breslin et al., 2009, p. 62).

In addition, the W3C document specification mentioned that blank node identifier "is used on a node element to replace rdf:about="RDF URI reference" or on a property element to replace rdf:resource="RDF URI reference" with rdf:nodeID="blank node identifier" in both cases (W3C, 2004e). Similarly Yu, added that, "a statement using a blank node as its subject" should use an RDF keyword rdf:Description element with an rdf:nodeID attribute as a substitute of rdf:about and "a statement using a blank node as its object then should use a property element with an rdf:nodeID" as a substitute of an rdf:resource attribute. (c2007, p. 54)

Therefore, in this case we used the blank rdf:nodeID as a subject and the rdf:Description element as equivalent to rdf:about. Each blank node has blank node identifier in RDF model and each identifier is used to identify or express the anonymous resources or the resource with no URI reference. Here to note that, the blank node identifier is only valid, accessible, and useable within particular document, therefore this identifier "is completely unknown outside of the scope of the document" (Yu, c2007, p. 53).

In addition, we have defined new namespace prefix or QNames for data modeling purpose i.e. "n5ui" and "content" are used for describing the Noark5 inner core and outer core records and the (http://ark1.hioa.no/n5.ui/) project web identifier, which is used as a means to create open source software of Noark5. The namespace are

4.10.1.Noark5 Inner core: fonds and fondscreator

The Noark conceptual model for archive structure depicts that an organization or person is can possibly create one or more fond entities as part of his/her work. The fond entity then can be shared between several bodies' for instance main office and their branch office can create their own fond entity and share. Therefore, this indicates that one fond creator have the possibility to create one or more fond entity. Using the RDF/XML serialization, RDF vocabularies, and new property name enable us to describe or express the creator and the fond entity as RDF statement.

The parsing result produces the RDF triples, and this triple verifies the author and the object relationship. For instance, (*fondscreator* is *created fond_1, fond_2, and fond_3 entities).* In addition new property name such as *isCreated* and *isCreatedBy*, and RDF keywords, rdf:Description element within the rdf:nodeID are used. Both isCreated and isCreatedBy property name have the capability of linking the subject and the object resource. RDF is the origin for semantic web and it is flexible any one can describe or express something about anything or can make statement of anything about something. In addition to that, RDF allows both human and machine to interact and cooperate. Therefore, currently we are able to relate the subject, which is "fond" with the object, which is "fondscreator" using the Predicate *isCreated* and *isCreatedBy.*

The following RDF/XML document shows the Fonds element with the fondscreator element and their semantic relationships.

Defined properties: - n5ui:isCreated, n5ui:isCreatedBy

```
1: <rdf:RDF
2:          xmlns:n5ui="http://ark1.hioa.no/n5.ui/"
3:          xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
4:          xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#">
5:                  <n5ui:fondscreator rdf:nodeID="fondscreator">
6:                  <n5ui:fondscreatorID>964963193</n5ui:fondscreatorID>
```

```
7:                    <n5ui:fondscreatorName>Oslo municipality</n5ui:fondscreatorName>
8:                         <n5ui:description>creating new fond</n5ui:description>
9:                           <n5ui:isCreated rdf:nodeID="fonds_1"/>
10:                          <n5ui:isCreated rdf:nodeID="fonds_2"/>
11:                          <n5ui:isCreated rdf:nodeID="fonds_3"/>
12:                  </n5ui:fondscreator>
13:                 <n5ui:fonds rdf:nodeID="fonds_1">
14:                          <n5ui:systemID>d7ee5939-4122-4b80-9e1e-057dac4f3d29</n5ui:systemID>
15:                          <n5ui:title>fond1</n5ui:title>
16:                          <n5ui:description>New fond</n5ui:description>
17:                          <n5ui:fondsstatus>closed</n5ui:fondsstatus>
18:                          <n5ui:createdDate>2012-02-21T20:32:39</n5ui:createdDate>
19:                          <n5ui:createdBy>Oslo municipality</n5ui:createdBy>
20:                          <n5ui:isCreatedBy rdf:nodeID="fondscreator"/>
21:                  </n5ui:fonds>
22:                 <n5ui:fonds rdf:nodeID="fonds_2">
23:                          <n5ui:systemID>d7ee5939-4122-4b80-9e1e-057dac4f3d27</n5ui:systemID>
24:                          <n5ui:title>fond2</n5ui:title>
25:                          <n5ui:description>New fond creation</n5ui:description>
26:                          <n5ui:fondsstatus>closed</n5ui:fondsstatus>
27:                          <n5ui:createdDate>2012-02-21T20:32:39</n5ui:createdDate>
28:                          <n5ui:createdBy>Oslo municipality</n5ui:createdBy>
29:                          <n5ui:isCreatedBy rdf:nodeID="fondscreator"/>
30:                  </n5ui:fonds>
31:                 <n5ui:fonds rdf:nodeID="fonds_3">
32:                          <n5ui:systemID>d7ee5939-4122-4b80-9e1e-057dac4f3d28</n5ui:systemID>
33:                          <n5ui:title>fond3</n5ui:title>
34:                          <n5ui:description>New fond3</n5ui:description>
35:                          <n5ui:fondsstatus>closed</n5ui:fondsstatus>
36:                          <n5ui:createdDate>2012-02-21T20:32:39</n5ui:createdDate>
37:                          <n5ui:createdBy>Oslo municipality</n5ui:createdBy>
38:                          <n5ui:isCreatedBy rdf:nodeID="fondscreator"/>
39:                  </n5ui:fonds>
40: </rdf:RDF>
```

Figure 4.10:1 fonds and fondscreator RDF/XML serialization
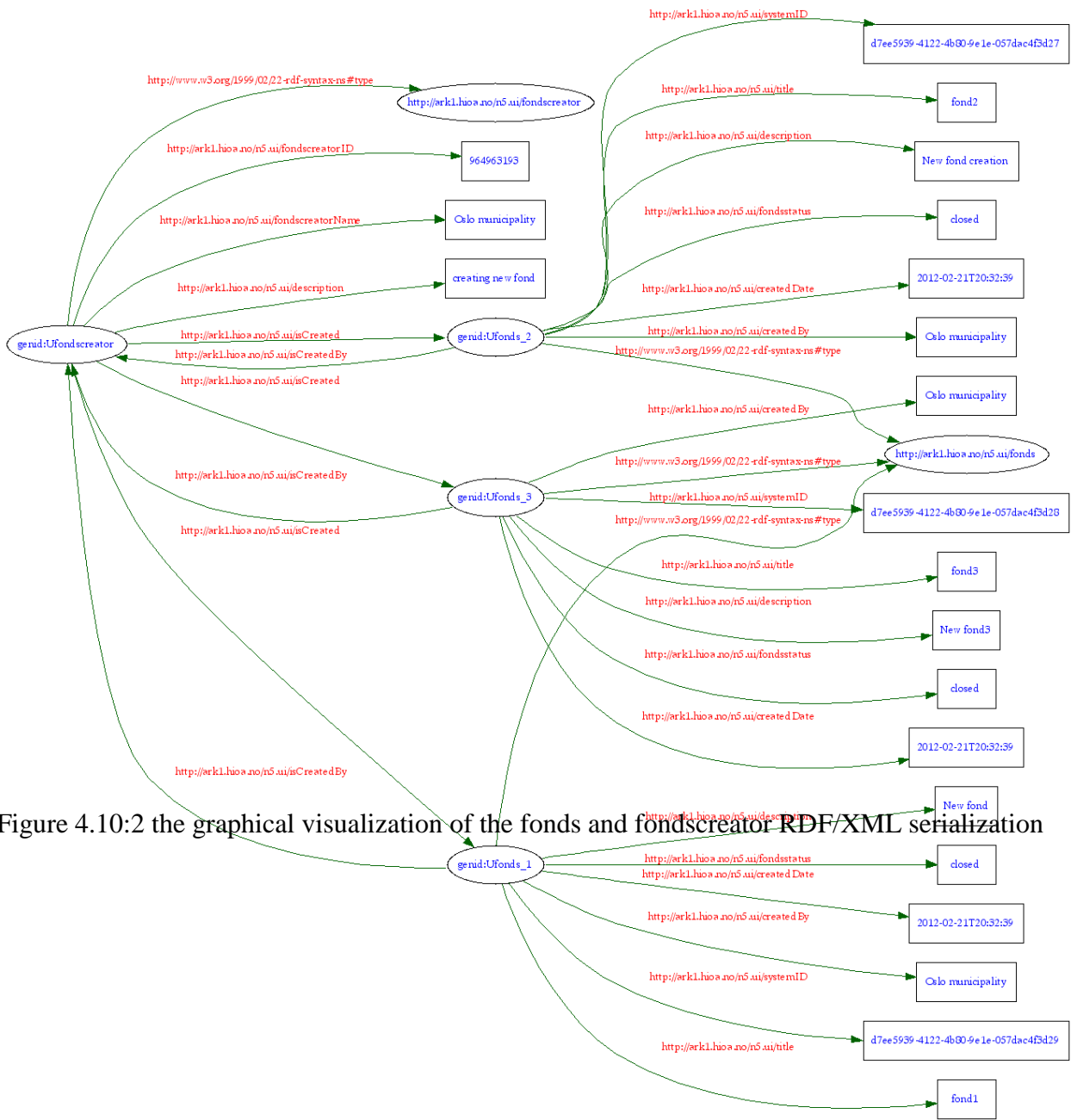
Figure 4.10:2 the graphical visualization of the fonds and fondscreator RDF/XML serialization

Subject	Predicate	Object
genid:Ufondscreator	http://www.w3.org/1999/02/22-rdf-syntax-ns#type	http://ark1.hioa.no/n5.ui/fondscreator
genid:Ufondscreator	http://ark1.hioa.no/n5.ui/fondscreatorID	"964963193"
genid:Ufondscreator	http://ark1.hioa.no/n5.ui/fondscreatorName	"Oslo municipality"
genid:Ufondscreator	http://ark1.hioa.no/n5.ui/description	"creating new fond"
genid:Ufondscreator	http://ark1.hioa.no/n5.ui/isCreated	genid:Ufonds_1

Table 4.10-1 Show the 3-tuple (triple) representation of the corresponding fonds and fondscreator RDF/XML serialization

The above *subject-predicate-object* triples are the result of the fonds and fondscreator RDF/XML serializations, for further detail of the triples see the appendix 2.

4.10.2.Archive structure

The models in Noark5 are *conceptual models,* which are intended to show the link between different metadata and between metadata and physical or electronic documents. The conceptual models in Noark5 state something about how the information should be organised in principle. They will also form the basis for the definition of data structures in connection with electronic communication, integration with other systems, migration from one system to another and for transfer (Noark5, 2012, pp. 37–38).

 The overall records of Noark5 is then organized according to the basic structure which is fond->series->file->record, the representation or mapping of the archive records can be done straightforward using the RDF/XML syntax without influencing the archive structure. For instance, using the *isRelated, issubPartOf, hasPart* property name the data with the "series" can be related with the data in the "files" as well as the "files" with its "records". Therefore, the highest level of the records structure "fonds" is related to "series" and proceeds to the other subsidiary part of the archive structure. The name property here plays significant role in creating semantic meaning between the elements.

Defined property names: n5ui:isRelated, n5ui:issubPartOf, n5ui:hasPart

RDF/XML document

```
1: <rdf:RDF
2:          xmlns:n5ui="http://ark1.hioa.no/n5.ui/"
3:          xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
4:          xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#">
5:                    <n5ui:fonds rdf:nodeID="fonds_1">
6:                              <n5ui:systemID>4e1b554c-dcc7-4b89-820f-7a8e0c44a672</n5ui:systemID>
7:                              <n5ui:description>fond1</n5ui:description>
8:                              <n5ui:hasPart rdf:nodeID="subfonds"/>
9:                              <n5ui:isRelated rdf:nodeID="series_1"/>
10:                   </n5ui:fonds>
11:                   <n5ui:subfonds rdf:nodeID="subfonds">
12:                              <n5ui:systemID>0a7a9c5a-f39a-43f4-a531-8fe6344c2188</n5ui:systemID>
13:                              <n5ui:title>subfonds</n5ui:title>
14:                              <n5ui:issubPartOf rdf:nodeID="fonds_1"/>
15:                   </n5ui:subfonds>
16:                   <n5ui:series rdf:nodeID="series_1">
17:                              <n5ui:systemID>de01097f-03d1-4032-b787-8062264ecf17</n5ui:systemID>
18:                              <n5ui:description>series1</n5ui:description>
19:                              <n5ui:isRelated rdf:nodeID="file_1"/>
20:                   </n5ui:series>
21:                   <n5ui:files rdf:nodeID="file_1">
22:                              <n5ui:systemID>83029702-519a-4933-aeda-89ba6cfcee2b</n5ui:systemID>
23:                              <n5ui:fileID>11/012465</n5ui:fileID>
24:                              <n5ui:description>file1</n5ui:description>
25:                              <n5ui:isRelated rdf:nodeID="record_1"/>
26:                   </n5ui:files>
27:                   <n5ui:records rdf:nodeID="record_1">
28:                              <n5ui:systemID>403ac99f-1d20-438c-a398-824b32d045ff</n5ui:systemID>
29:                              <n5ui:recordtype>document</n5ui:recordtype>
30:                              <n5ui:recordID>11/012466</n5ui:recordID>
31:                              <n5ui:description>record1</n5ui:description>
32:                              <n5ui:hasPart rdf:nodeID="documentdescription"/>
33:                              <n5ui:hasPart rdf:nodeID="documentobject"/>
34:                   </n5ui:records>
35:                   <n5ui:documentdescription rdf:nodeID="documentdescription">
36:                              <n5ui:systemID>29743769-edf4-46f5-be47-51a0167fbfeb</n5ui:systemID>
37:                              <n5ui:documenttype>electronic record</n5ui:documenttype>
38:                              <n5ui:documentstatus>closed</n5ui:documentstatus>
39:                              <n5ui:createdDate>2011-02-21T20:32:56</n5ui:createdDate>
40:                              <n5ui:createdBy>Oslo municipality</n5ui:createdBy>
```

```
41:                    </n5ui:documentdescription>
42:                    <n5ui:documentobject rdf:nodeID="documentobject">
43:                            <n5ui:versionnumber>bf1dfa10-243a-4a7c-9904-89ec9325cca0</n5ui:versionnumber>
44:                            <n5ui:variantformat>production format</n5ui:variantformat>
45:                            <n5ui:format>pdf format</n5ui:format>
46:                            <n5ui:createdDate>2011-02-21T20:32:56</n5ui:createdDate>
47:                            <n5ui:createdBy>Oslo municipality</n5ui:createdBy>
48:                    </n5ui:documentobject>
49: </rdf:RDF>
```

Figure 4.10:3 Noark5 archive structure RDF/XML serialization

Subject	Predicate	Object
genid:Ufonds_1	http://www.w3.org/1999/02/22-rdf-syntax-ns#type	http://ark1.hioa.no/n5.ui/fonds
genid:Ufonds_1	http://ark1.hioa.no/n5.ui/systemID	"4e1b554c-dcc7-4b89-820f-7a8e0c44a672"
genid:Ufonds_1	http://ark1.hioa.no/n5.ui/description	"fond1"
genid:Ufonds_1	http://ark1.hioa.no/n5.ui/hasPart	genid:Usubfonds
genid:Ufonds_1	http://ark1.hioa.no/n5.ui/isRelated	genid:Useries_1
genid:Usubfonds	http://www.w3.org/1999/02/22-rdf-syntax-ns#type	http://ark1.hioa.no/n5.ui/subfonds

Table 4.10.2-1 the 3-tuple (triple) representation of the corresponding Noark5 archive RDF/XML serialization and for detailed triples data model, see appendix 3

The above (Table4.10.2-1) triple data model describes and expresses the Noark5 basic archive structure and it indicates that the archive or the ERM are always followed this basic record structure to maintain and ensure the electronic record quality. Therefore, all the digital data and records are organized and stored with respect of this archive structure. Therefore, the RDF triple shows the detailed data modeling process for instance, the subject called genid:unfonds_1 that have fonds type related to the object called series_1

using the *isRelated* predicate. Like wise, the pervious object binary can also be act and use as subject in another statement for instance the "series_1" can relate to "file_1". Both the subjects and objects resources are then used the property name or predicate isRelated in order to relate and interact between them. See (Figure4.10.2:2) for the graphical visualization.
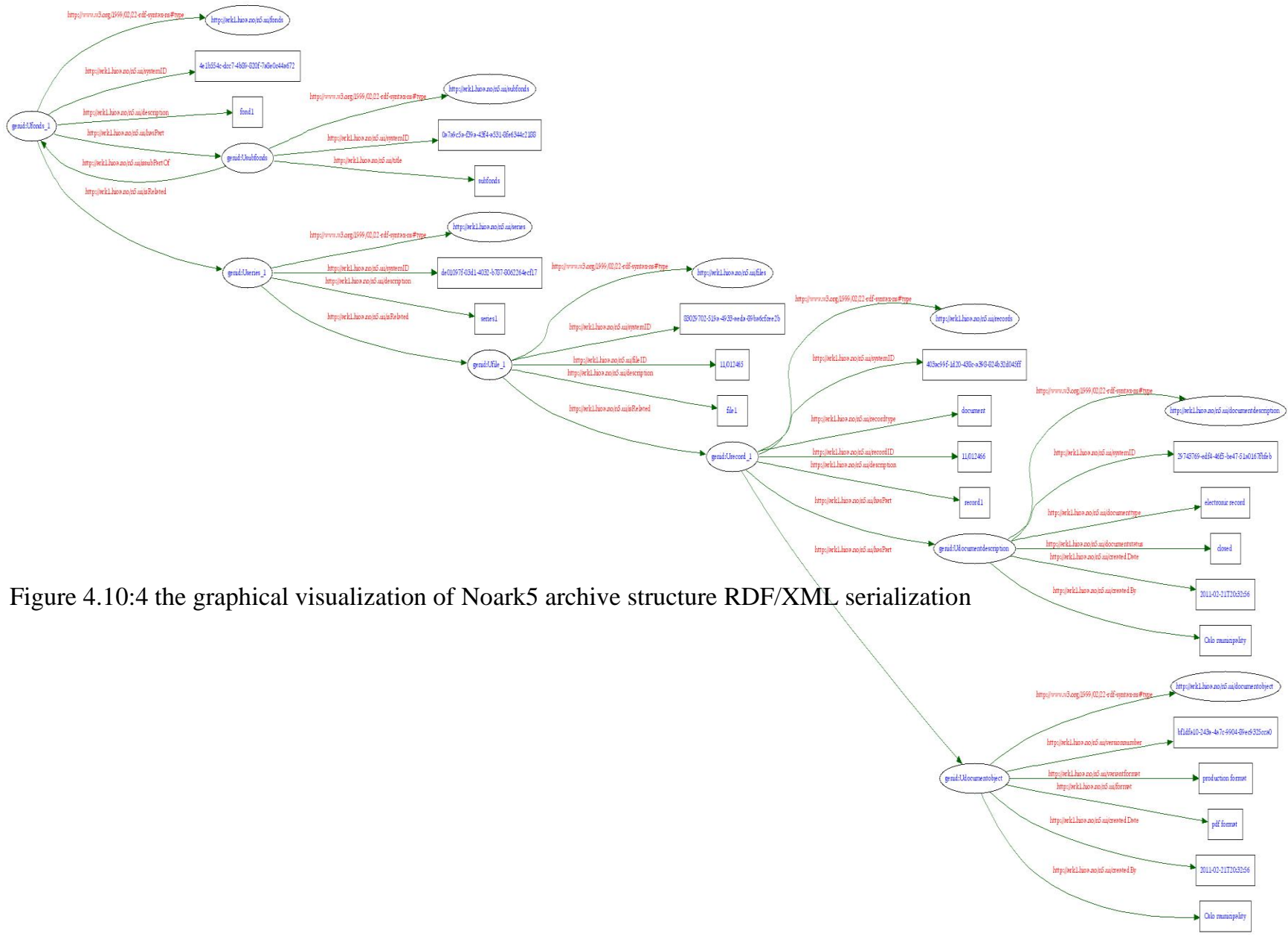
Figure 4.10:4 the graphical visualization of Noark5 archive structure RDF/XML serialization

4.10.3. Fond entity

The creation of several fonds is possible in Noark5 record keeping standard. Therefore, a unit or an organization has the possibility of creating one or several fonds, for instance, a company with several departments or branches offices. Such as "The national archives of Norway" and the eight "Regional State archives" located in different part of Norway could have a common fond entity and/ or can have the possibility to create their own entity and this entity can be linked to the main or common fond entity using semantic relation. In this case, "The national archives of Norway" is the main fond entity. Soon after, every member of the company can then have possibility to use and share the fond entities. Here is the advantage of semantic web it enables us to add extra meaning to data, to represent, and link digital resources from different location within diverse databases. As a result, it also enhances data interchangeability and the usability over the web that is stored, organized, and available for access by the archive organizations. The following RDF/XML syntax shows the benefit of semantics.

Defined RDF properties: n5ui:hasPart, n5ui:BelongTo, n5ui:isPartOf

RDF/XML syntax

```
1: <rdf:RDF
2:          xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
3:          xmlns:n5ui="http://ark1.hioa.no/n5.ui/"
4:    xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#">
5:          <n5ui:fondentity rdf:nodeID="fondentity">
6:                    <n5ui:hasPart rdf:nodeID="fonds1"/>
7:                    <n5ui:hasPart rdf:nodeID="fonds2"/>
8:                    <n5ui:hasPart rdf:nodeID="fonds3"/>
9:          </n5ui:fondentity>
10:         <n5ui:fonds rdf:nodeID="fonds1">
11:                   <n5ui:belongsTo rdf:nodeID="fondentity"/>
12:                   <n5ui:hasPart rdf:nodeID="subfonds"/>
13:         </n5ui:fonds>
14:                   <n5ui:fonds rdf:nodeID="fonds2">
15:         <n5ui:belongsTo rdf:nodeID="fondentity"/>
16:         </n5ui:fonds>
17:         <n5ui:fonds rdf:nodeID="fonds3">
18:                   <n5ui:belongsTo rdf:nodeID="fondentity"/>
```

```
19:        </n5ui:fonds>
20:        <n5ui:fondscreator rdf:nodeID="fondscreator">
21:        <n5ui:isCreated rdf:nodeID="fondentity"/>
22:        </n5ui:fondscreator>
23:        <n5ui:subfonds rdf:nodeID="subfonds">
24:        <n5ui:isPartOf rdf:nodeID="fonds1"/>
25:        </n5ui:subfonds>
26:        <n5ui:fonds rdf:nodeID="fonds1">
27:                <n5ui:systemID>e05862f1-11ac-4b06-9228-a18c0005b205</n5ui:systemID>
28:                <n5ui:isRelated rdf:nodeID="series"/>
29:        </n5ui:fonds>
30:        <n5ui:series rdf:nodeID="series">
31:         <n5ui:systemID>e05862f1-11ac-4b06-9228-a18c0005b205</n5ui:systemID>
32:         <n5ui:isRelated rdf:nodeID="files1"/>
33:         <n5ui:isRelated rdf:nodeID="files2"/>
34:         <n5ui:isRelated rdf:nodeID="files3"/>
35:        </n5ui:series>
36:        <n5ui:files rdf:nodeID="files1">
37:         <n5ui:systemID>e05862f1-11ac-4b06-9228-a18c0005b205</n5ui:systemID>
38:         <n5ui:fileID>12/548752</n5ui:fileID>
39:         <n5ui:isRelated rdf:nodeID="record1"/>
40:         <n5ui:isRelated rdf:nodeID="record2"/>
41:         <n5ui:isRelated rdf:nodeID="record3"/>
42:        </n5ui:files>
43:        <n5ui:records rdf:nodeID="records1">
44:                <n5ui:recordID>12/654132-5874</n5ui:recordID>
45:                <n5ui:hasPart rdf:nodeID="documentdescription"/>
46:                <n5ui:isPartOf rdf:nodeID="files1"/>
47:        </n5ui:records>
48:    <n5ui:records rdf:nodeID="records2">
49:                <n5ui:recordID>12/686522-5412</n5ui:recordID>
50:                <n5ui:hasPart rdf:nodeID="documentdescription"/>
51:                <n5ui:isPartOf rdf:nodeID="files1"/>
52:        </n5ui:records>
53:        <n5ui:documentdescription rdf:nodeID="documentdescription">
54:                <n5ui:hasPart rdf:nodeID="documentobject"/>
55:                <n5ui:isPartOf rdf:nodeID="records1"/>
56:    <n5ui:isPartOf rdf:nodeID="record2"/>
57:        </n5ui:documentdescription>
58:        <n5ui:documentobject rdf:nodeID="documentobject">
59:                <n5ui:isPartOf rdf:nodeID="documentdescription"/>
60:        </n5ui:documentobject>
61: </rdf:RDF>    Figure 4.10:5 Fond entity RDF/XML serialization
```

Figure 4.10:6 RDF graphical visualization of the fond entity RDF/XML serialization

4.10.4. Modeling Noark5 outer core

Noark5 outer core contains requirements and specifications related to external system or optional system. The main aim of Noark5 outer core is to enhance and assure the quality communication and interaction between various pre-systems/optional systems. In similar to Noark5 inner core, a system to be approved as a Noark5 solution the requirement and specification for Noark5 outer core must be fulfilled. Furthermore, "In order for Noark5 to function in an integrated archive system environment, it must be possible to integrate Noark5 with a number of pre-systems/task systems which must be implemented in order to supplement a recordkeeping and archive solution" (Noark5, 2012, p. 112).

The Noark5 outer core has user administration, case handling, e-mail, administrative structure, meeting/boarding handling, use functionality, security and access control, and reporting components. Further, requirements for case handling, case distribution, recording or file moving, splitting, and other related issues are also included in Noark5 outer core. Moreover, the Noark5 outer core contains information related to case parties as well as precedent. Case parties contains information about the object of interest (for instance, the name of the person, the address of the case party, their email, etc), on the other hand, there is called precedent that sometimes act as a guide for case or situation, which required legal decision. Therefore, "a precedent can also be a case that is governing for the processing of other similar cases" (Noark5, 2012, p. 123).

However, Noark5 based system provides records or data that are produced or received from an organizations or individuals. The main objective of Noark5 is then to confirm the record quality, to assign necessary metadata for instance createdDate, archviedDate, author name, and sorting records accordingly, and making the data available for access are some of the functions of Noark.

However, the records that are created do not provide any detailed information in relation to who created the document or the *object of interest*. The main reason is that the object of interest is nested within the system property. The system is able to record and organize both the data as well as the object of interest, however finding the owner identity or anything related to the object of interest is harder to access or retrieve. Therefore, in order

to make both the subject and object accessible and usable we need to have semantic based implementation. This implementation shows the way of exposing the object of interest from the current system using the basic RDF data model and its semantic relationships.

Therefore, what we are attempting to achieve here is to model the object of interest in RDF form, to create semantic relationship between binaries and to allow interaction and data exchange between the subject (resource) and its object (people) in the Noark5 system.

One or more organisations or people can be linked to a case file as parties to a case (*case parties*). A party is the party at which a decision is aimed or which the case otherwise directly concerns. Case party in Noark5 consists of information to describe parties to a case such as (name, address, etc.). Information concerning case parties is not obligatory in Noark5 core. The requirements for case parties are obligatory for solutions that involve parties. Information on case parties is obligatory for transfer (Noark5, 2012, p. 122).

The RDF/XML representation express the case party name, role, address etc, in addition it uses the short prefix "content" in order to differentiate the object from its subject resource. The case party can contain different type such as person, organization, or foundation, and thus different units can create case as parties to a case.

```
1: <rdf:RDF
2:          xmlns:content="http://ark1.hioa.no/n5.ui/content/"
3:          xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
4:          xmlns:n5ui="http://ark1.hioa.n5.ui/"
5:          xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#">
6:
7:                    <content:person rdf:nodeID="case_party1">
8:                          <content:casepartyID>e565939-4187-4b10-9e1e-057dac924518</content:casepartyID>
9:                          <content:casepartyName>Debora</content:casepartyName>
10:                         <content:casepartyRole>Casecreator</content:casepartyRole>
11:                         <content:postaladdress>Toyen gate42</content:postaladdress>
12:                         <content:postcode>0578</content:postcode>
13:                         <content:postaltown>Oslo</content:postaltown>
14:                         <content:foreignaddress>Norway</content:foreignaddress>
15:                         <content:emailaddress>Deb@gmail.com</content:emailaddress>
16:                         <content:telephonenumber>47599874</content:telephonenumber>
17:                    </content:person>
```

```
18:                              <content:person rdf:nodeID="case_party2">
19:                                      <content:casepartyID>d7ee5939-4122-4b80-9e1e-874521e6</content:casepartyID>
20:                                      <content:casepartyName>Helen</content:casepartyName>
21:                                      <content:casepartyRole>Casecreator</content:casepartyRole>
22:                                      <content:postaladdress>Tofet gate 21</content:postaladdress>
23:                                      <content:postcode>0785</content:postcode>
24:                                      <content:postaltown>Oslo</content:postaltown>
25:                                      <content:foreignaddress>Norway</content:foreignaddress>
26:                                      <content:emailaddress>helu@gmail.com</content:emailaddress>
27:                                      <content:telephonenumber>47125474</content:telephonenumber>
28:                              </content:person>
29:                      <content:publicsector rdf:nodeID="case_party3">
30:                                      <content:casepartyID>59d9cff4-32ff-4e3e-a5e7-ac0b6a6edc71</content:casepartyID>
31:                                      <content:casepartyName>Oslo municipality</content:casepartyName>
32:                      </content:publicsector>
33:                      <content:case rdf:nodeID="precedent_case">
34:                                      <content:systemID>e05862f1-11ac-4b06-9228-a18c0005b205</content:systemID>
35:                                      <content:fileID>12/022459</content:fileID>
36:                                      <content:casedate>2012-02-21T20:33:12</content:casedate>
37:                                      <content:casestatus>under process</content:casestatus>
38:                                      <content:case-responsible>oslo municpality</content:case-responsible>
39:                                      <content:administrativeUnit></content:administrativeUnit>
40:                                      <content:precedentstatus>current</content:precedentstatus>
41:                      </content:case>
42:                      <content:person rdf:nodeID="case_handler">
43:                                      <content:firstname>Seliman</content:firstname>
44:                                      <content:lastname>Enger</content:lastname>
45:                                      <content:personRole>executive_officer</content:personRole>
46:                      </content:person>
47:                      <content:case rdf:nodeID="case">
48:                                      <content:caseparty rdf:nodeID="case_party1"/>
49:                                      <content:case_handler rdf:nodeID="case_handler"/>
50:                                      <content:caseparty rdf:nodeID="case_party2"/>
51:                                      <content:caseparty rdf:nodeID="case_party3"/>
52:                                      <content:case rdf:nodeID="precedent_case"/>
53:                      </content:case>
54: </rdf:RDF>
```

Figure 4.10:7Case parties RDF/XML serialization
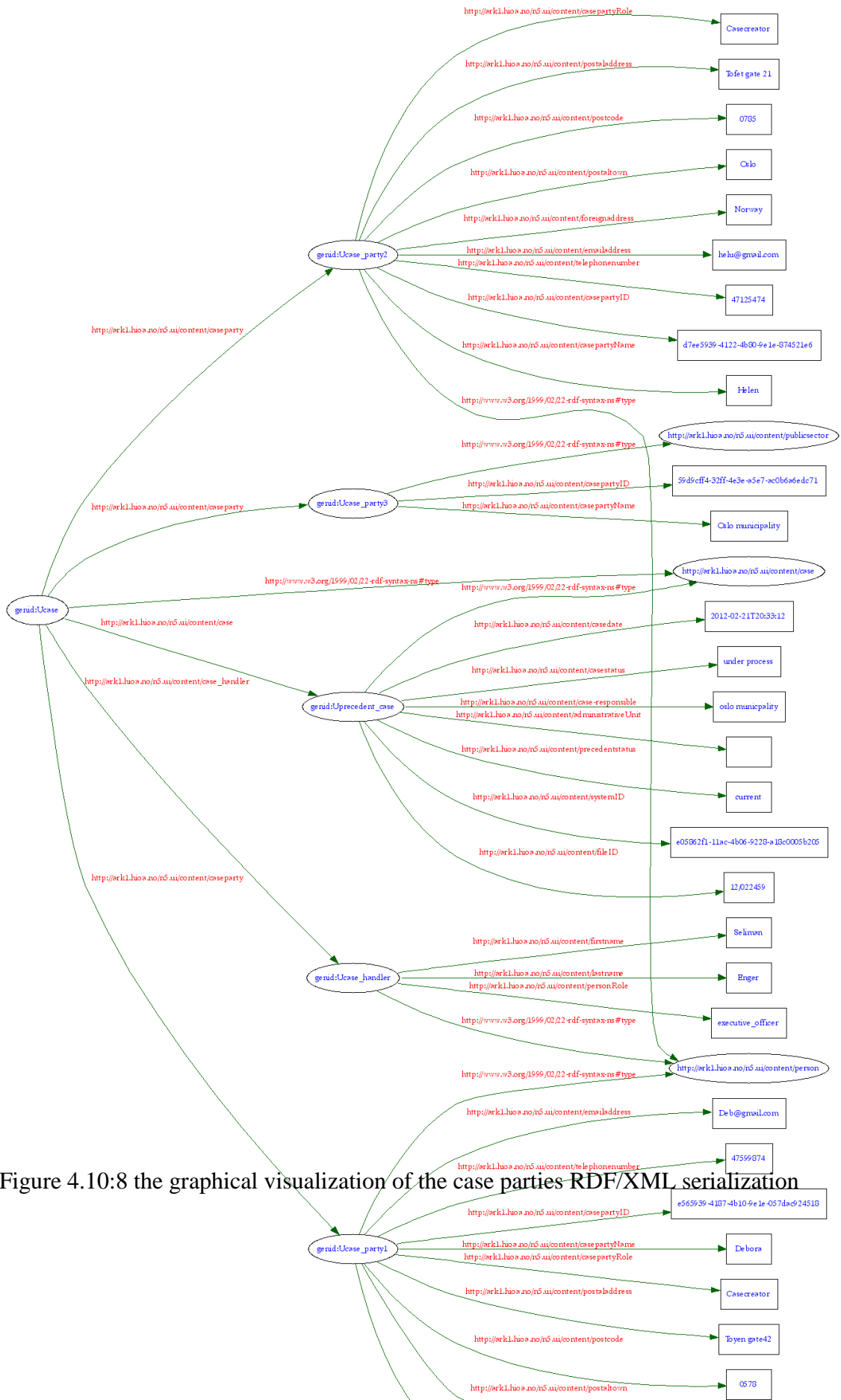
Figure 4.10:8 the graphical visualization of the case parties RDF/XML serialization

4.10.5.Relating the archive structure data with object of interest

The previous section 4.8.2 and 4.8.4 produced RDF data for the Noark5 archive structure, and Noark5 objects of interest, which is case parties. The main vision of semantic web in not limited to the creation of meaning among data where as it also contains the possibility to relate, link, formalize more data or databases. In this case, the above two resources subject (records structure) and object (case party) are mainly known as the Noark5 inner core and Noark5 outer core.

Therefore, linking, identifying, and creating meaningful relationships between these binaries can reduce the limitation for retrieving the resources in addition to the possibility of unlocking the object of interest from the system. Therefore, the modeling enables us to understand and access not only the records but also the object types for instance, the person, organization, department etc.

The RDF name properties: n5ui:hasCaseHandler, n5ui:hasCase, n5ui:isPartOf, n5ui:isRelated, n5ui:isResponsibleFor, n5ui:hasPart, and n5ui:hasCaseType.

```
1: <rdf:RDF
2:          xmlns:content="http://ark1.hioa.no/n5.ui/content/"
3:          xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
4:          xmlns:n5ui="http://ark1.hioa.n5.ui/"
5:          xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#">
6:
7:                   <n5ui:fonds rdf:nodeID="fonds">
8:                            <n5ui:systemID>d7ee5939-4122-4b80-9e1e-057dac4f3d29</n5ui:systemID>
9:                            <n5ui:title>new fond</n5ui:title>
10:                           <n5ui:description>Oslo manicupality</n5ui:description>
11:                           <n5ui:documentstatus>closed</n5ui:documentstatus>
12:                           <n5ui:documentmedium>electronic format</n5ui:documentmedium>
13:                           <n5ui:createdDate>2012-02-21T20:32:39</n5ui:createdDate>
14:                           <n5ui:createdBy>Oslo manicupality</n5ui:createdBy>
15:                           <n5ui:finalisedDate>2012-02-22T14:38:29</n5ui:finalisedDate>
16:                           <n5ui:finalisedBy>Oslo manicupality</n5ui:finalisedBy>
17:                           <n5ui:isCreatedBy rdf:nodeID="fondscreator"/>
18:                           <n5ui:isRelated rdf:nodeID="series"/>
19:                   </n5ui:fonds>
20:                   <n5ui:series rdf:nodeID="series">
```

```
21:                                <n5ui:systemID>b23fd11e-1860-4b84-a449-c77c86357330</n5ui:systemID>

22:                                <n5ui:title>series</n5ui:title>

23:                                <n5ui:description>series</n5ui:description>

24:                                <n5ui:documentmedium>electronic format</n5ui:documentmedium>

25:                                <n5ui:createdDate>2012-02-21T20:32:56</n5ui:createdDate>

26:                                <n5ui:createdBy>Oslo manicupality</n5ui:createdBy>

27:                                <n5ui:finalisedDate>2012-02-22T14:38:29</n5ui:finalisedDate>

28:                                <n5ui:finalisedBy>Oslo manicupality</n5ui:finalisedBy>

29:                                <n5ui:isRelated rdf:nodeID="file"/>

30:                 </n5ui:series>

31:             <n5ui:file rdf:nodeID="file">

32:                                <n5ui:systemID>e05862f1-11ac-4b06-9228-a18c0005b205</n5ui:systemID>

33:                                <n5ui:title>file</n5ui:title>

34:                                <n5ui:description>application for KG</n5ui:description>

35:                                <n5ui:fileID>10/629652</n5ui:fileID>
```

………

…………………………………………..

Figure 4.10:9 Noark5 archives and case parties RDF/XML serialization, see appendix 4 for the parsing details

So far we defined and make use of the following RDF properties: - n5ui:isCreated,

n5ui:isRelated, n5ui:hasCaseType, n5ui:hasPart, n5ui:hasCaseHandler, n5ui:isPartOf,

and n5ui:isResponsibleFor.
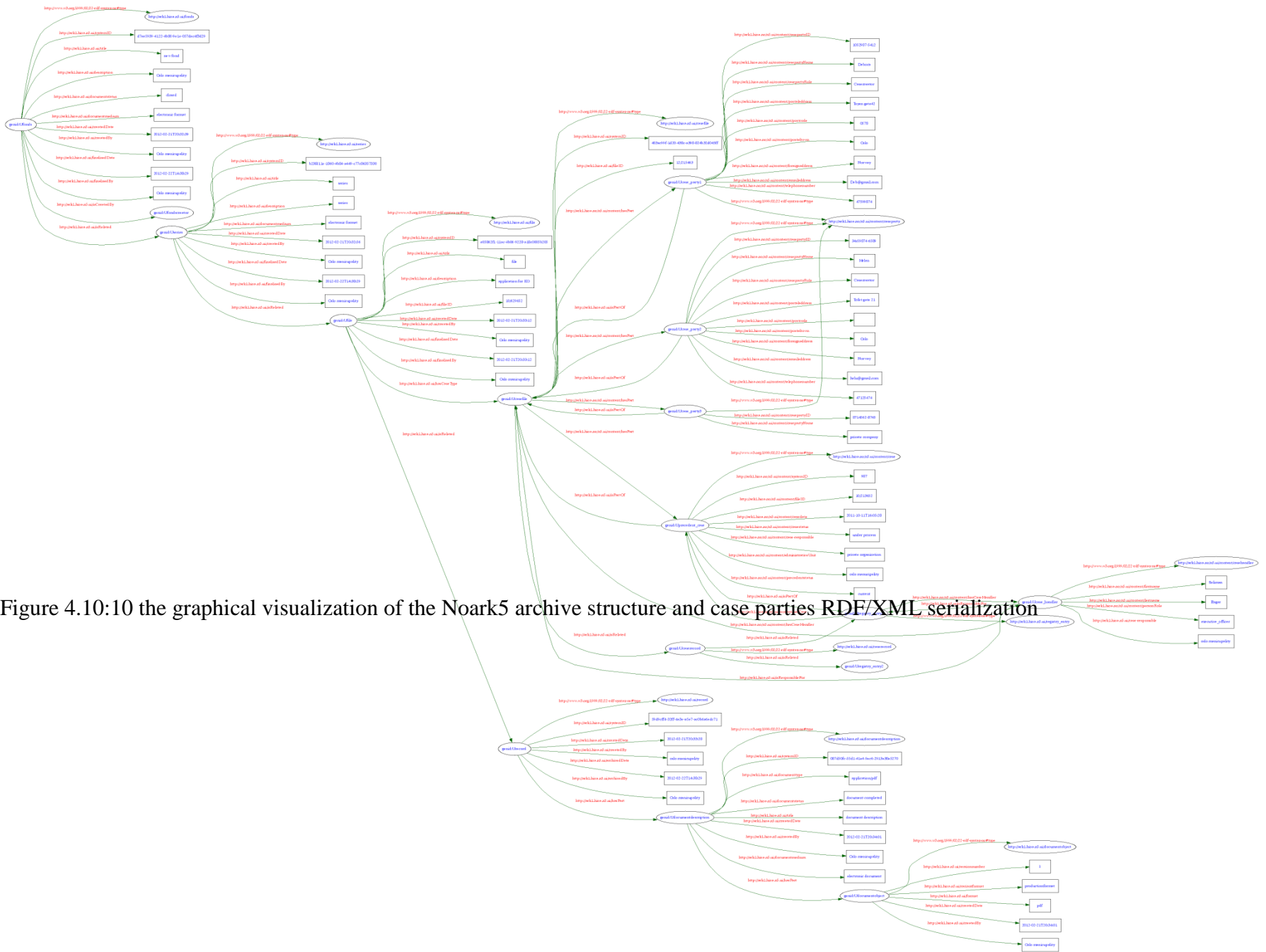
Figure 4.10:10 the graphical visualization of the Noark5 archive structure and case parties RDF/XML serialization

4.10.6.Creating semantic relationship between an organization and National archives of Norway

On the other hand, as we tried to mention it in the pervious section the semantic web could enable us to create more semantic based relationship in between different archive organizations or public administrations of Norway. For instance, let us take the national archive of Norway as an example. The national archive of Norway has a central office in Oslo. All the regional state archives together form the National archives of Norway, headed by the Director General. The National Archives of Norway is an independent government authority under the Ministry of Culture. There are eight regional archives in Norway and a Sámi Archives. Integrating the semantic web technology for Norwegian public administration and records management can have the benefit for the archives/ERM to use semantic based records from the different Noark5 databases, which is handled by the eight regional state archives.

The Organisation:-The central Office of the National archives of Norway in Oslo and the regional state archives together form the National Archives of Norway, headed by the Director General. The National Archives of Norway is an independent government authority under the ministry of culture.

The National Archives: - Central Office in Oslo is responsible for the records created by the government's central administration, i.e. ministries of the Supreme Court.

- The Regional State Archives in Oslo: - is a part of the National Archives of Norway, covering the district of Oslo, Østfold, and Akershus. The Regional State Archives in Oslo and The National Archives are both situated in the National Archives Building in Oslo.
- The Regional State Archives in Hamar: - is a part of the National Archives of Norway, covering the district of Hedmark and Oppland.
- The Regional State Archives in Kongsberg: - is a part of the National Archives of Norway, covering the district of Buskerud, Telemark and Vestfold.

- The Regional State Archives in Kristiansand: - is a part of the National Archives of Norway, covering the districts of Aust-Agder and Vest-Agder.

- The Regional State Archives in Stavanger: - is a part of the National Archives of Norway, covering the district of Rogaland.

- The Regional State Archives in Bergen: - is a part of the National Archives of Norway, covering the district of Hordaland og Sogn and Fjordane.

- The Regional State Archives in Trondheim: - is a part of the National Archives of Norway, covering the district of Møre og Romsdal, Sør-Trøndelag, Nord-Trøndelag og Nordland.

- The Regional State Archives in Tromso: - is a part of the National Archives of Norway, covering the district of Troms, Finnmark and Svalbard.

In addition, the Samisk arkiv/Sami arkiiva archives is situated in Kautokeino and is deposited public archives of importance to the Sámi community, history and culture archives is also responsible for Sami private archives (Noark-Arkivverket, 2011).

Using the following property names both the national and regional archives can adapt semantic based communication and relation between them. N5ui:isRelated, n5ui:hasCreator, n5ui:hasDepartment, n5ui:hasGeneralDirector, n5ui:hasSecertary, n5ui:hasRegionalBranch, n5ui:isOwnerOf, n5ui:isResponsibleFor, n5ui:hasOffice, n5ui:issubPartOf, and n5ui:issharedTheSame. Further, both have the possibility of using huge amount of data from different part of their databases.

```
1: <rdf:RDF
2:          xmlns:n5ui="http://ark1.hioa.no/n5.ui/"
3:          xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
4:          xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#">
5:                    <n5ui:fondentity rdf:nodeID="fondentity">
6:                              <n5ui:systemID>18b5f512-c780-499c-837a-0d05d4c73dcf</n5ui:systemID>
7:                              <n5ui:title>new fond</n5ui:title>
8:                              <n5ui:description>new fond creation</n5ui:description>
9:                              <n5ui:isRelated rdf:nodeID="series"/>
```

```
10:                                <n5ui:hasCreator rdf:nodeID="organisation"/>
11:                        </n5ui:fondentity>
12:                        <n5ui:fondscreator rdf:nodeID="organisation">
13:                                <n5ui:fondscreatorID>546738</n5ui:fondscreatorID>
14:                                <n5ui:hasDepartment rdf:nodeID="department"/>
15:                                <n5ui:hasGeneralDirector rdf:nodeID="generaldirector"/>
16:                                <n5ui:hasSecertary rdf:nodeID="secertary"/>
17:                                <n5ui:hasRegionalBranch rdf:nodeID="regionalstatearchives"/>
18:                                <n5ui:isOwnerOf rdf:nodeID="fondentity"/>
19:                        </n5ui:fondscreator>
20:                        <n5ui:generaldirector rdf:nodeID="generaldirector">
21:                                <n5ui:isResponsibleFor rdf:nodeID="organisation"/>
22:                        </n5ui:generaldirector>
23:                        <n5ui:regionalstatearchives rdf:nodeID="regionalstatearchives">
24:                                <n5ui:hasOffice rdf:nodeID="Bergen"/>
25:                                <n5ui:hasOffice rdf:nodeID="Stavanger"/>
26:                                <n5ui:hasOffice rdf:nodeID="Tromso"/>
27:                                <n5ui:hasOffice rdf:nodeID="Trondhiem"/>
28:                                <n5ui:hasOffice rdf:nodeID="Oslo"/>
29:                                <n5ui:hasOffice rdf:nodeID="Hamar"/>
30:                                <n5ui:hasOffice rdf:nodeID="kongsberg"/>
31:                                <n5ui:hasOffice rdf:nodeID="Kristiandsand"/>
32:                                <n5ui:hasOffice rdf:nodeID="samarkiiva"/>
33:                                <n5ui:issubPartOf rdf:nodeID="organisation"/>
34:                        </n5ui:regionalstatearchives>
35:                        <n5ui:regionalstatearchives rdf:nodeID="Bergen">
36:                                <n5ui:issharedTheSame rdf:nodeID="fondentity"/>
37:                                <n5ui:isPartOf rdf:nodeID="regionalstatearchives"/>
38:                        </n5ui:regionalstatearchives>
39:                        <n5ui:regionalstatearchives rdf:nodeID="Stavanger">
40:                                <n5ui:issharedTheSame rdf:nodeID="fondentity"/>
41:                                <n5ui:isPartOf rdf:nodeID="regionalstatearchives"/>
42:                        </n5ui:regionalstatearchives>
43:                        <n5ui:regionalstatearchives rdf:nodeID="Tromso">
44:                                <n5ui:issharedTheSame rdf:nodeID="fondentity"/>
45:                                <n5ui:isPartOf rdf:nodeID="regionalstatearchives"/>
46:                        </n5ui:regionalstatearchives>
47:                        <n5ui:regionalstatearchives rdf:nodeID="Trondhiem">
48:                                <n5ui:issharedTheSame rdf:nodeID="fondentity"/>
49:                                <n5ui:isPartOf rdf:nodeID="regionalstatearchives"/>
50:                        </n5ui:regionalstatearchives>
51:                        <n5ui:regionalstatearchives rdf:nodeID="Oslo">
52:                                <n5ui:issharedTheSame rdf:nodeID="fondentity"/>
```

```
53:                              <n5ui:description>head office</n5ui:description>
54:                              <n5ui:isPartOf rdf:nodeID="regionalstatearchives"/>
55:                      </n5ui:regionalstatearchives>
56:                      <n5ui:regionalstatearchives rdf:nodeID="Hamar">
57:                              <n5ui:issharedTheSame rdf:nodeID="fondentity"/>
58:                              <n5ui:isPartOf rdf:nodeID="regionalstatearchives"/>
59:                      </n5ui:regionalstatearchives>
60:                      <n5ui:regionalstatearchives rdf:nodeID="kongsberg">
61:                              <n5ui:issharedTheSame rdf:nodeID="fondentity"/>
62:                              <n5ui:isPartOf rdf:nodeID="regionalstatearchives"/>
63:                      </n5ui:regionalstatearchives>
64:                      <n5ui:regionalstatearchives rdf:nodeID="Kristiansand">
65:                              <n5ui:issharedTheSame rdf:nodeID="fondentity"/>
66:                              <n5ui:isPartOf rdf:nodeID="regionalstatearchives"/>
67:                      </n5ui:regionalstatearchives>
68:                      <n5ui:regionalstatearchives rdf:nodeID="samarkiiva">
69:                              <n5ui:issharedTheSame rdf:nodeID="fondentity"/>
70:                              <n5ui:isPartOf rdf:nodeID="regionalstatearchives"/>
71:                      </n5ui:regionalstatearchives>
72:                      <n5ui:series rdf:nodeID="series">
73:                              <n5ui:systemID>7a129a6f-a990-4424-89d2-8945e5859ff0</n5ui:systemID>
74:                              <n5ui:isRelated rdf:nodeID="files"/>
75:                      </n5ui:series>
76:                      <n5ui:files rdf:nodeID="files">
77:                              <n5ui:systemID>0282d31d-6a0f-4fb5-b027-b89d2f34b121</n5ui:systemID>
78:                              <n5ui:isRelated rdf:nodeID="records"/>
79:                      </n5ui:files>
80:                      <n5ui:records rdf:nodeID="records">
81:                              <n5ui:systemID>dc3381e4-9f1d-493f-9a69-a66903a0dea8</n5ui:systemID>
82:                      </n5ui:records>
83: </rdf:RDF>
```

Figure 4.10:11 an organization and the National archives of Norway RDF/XML serialization, see appendix 5 for graphical visualization

4.10.7. Files and Records of Noark5 extracted document

In Noark5 Files and records are organized with the help of classes. Classification system and classes are used in the archive structure to group the series and the files. Class is the sub part of the main classification system see section 4.10.9. for class and classification system. A group of files and documents that have the same type or that belongs to the same activity can be organized to the same file cases and records.

The following RDF/XML document express the semantic relationship between several files and records using the name property n5ui:isPartOf, n5ui:hasPart, and issubPartOf. The main advantage of relating more files and records that belongs to the same entity or activity allow us to access embedded data, in addition the more we create relationship the more we increase the possible for retrieving relevance data on the right time. This indicates that these property names or predicates will minimize wasting of time for navigating and searching time.

```
1: <rdf:RDF
2:         xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
3:         xmlns:n5ui="http://ark1.hioa.no/n5.ui/"
4:         xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#" >
5:                 <rdf:Description rdf:nodeID="fonds">
6:                         <n5ui:systemID>68f08d11-6796-409a-96da-1b403b406549</n5ui:systemID>
7:                         <n5ui:title>new fond</n5ui:title>
8:                         <n5ui:haspart rdf:nodeID="subfonds"/>
9:                         <n5ui:createdBy rdf:nodeID="fondscreator"/>
10:                         <n5ui:isRelated rdf:nodeID="series"/>
11:                 </rdf:Description>
12:                 <rdf:Description rdf:nodeID="fondscreator">
13:                         <n5ui:fondscreatorID>7261</n5ui:fondscreatorID>
14:                         <n5ui:fondscreatorName>Oslo municipality</n5ui:fondscreatorName>
15:                         <n5ui:description> new document creation</n5ui:description>
16:                         <n5ui:isCreated rdf:nodeID="fond"/>
17:                 </rdf:Description>
            ----------------------------------------

                ------------------------------------------------

188:         <n5ui:isPartOf rdf:nodeID="documentdescription3of3"/>
189:         </n5ui:documentobject>
190: </rdf:RDF>
```

Figure 4.10:12 The Noark5 files and records RDF/XML serialization see also appendices 6 and 7 for detailed RDF/XML serialization and graphical visualization

4.10.8.Case files, case records, registry entry with case parties, and clients

In addition to the pervious archive structure and types there are also file types called case files in Noark5 that correspond to a "case" in the Noark4. Case files that are received or dispatched as an activity or applications are usually grouped into the folders called files cases and record cases. According to the Noark5 description and specification, it is possible to create many other specialized file types based on the case files such as cases from Noark4 *appointment case, and building case* (Noark5, 2012, p. 60). Using semantic web a semantic relationship can be created between these cases types and their correspondence files for instance relating the files to the case files using the property name n5ui:hasCaseFile (the file resource hasCaseFile whose value is case file).

The case files and case records that belongs together then will related to each other, in similar way documents or records that belongs to the type of files and records then will grouped or related to files etc. The modeling part here shows the possibility of relating these kinds of case files, case records, registry entry, records, and case records with their clients, case parties and their executives. Doing so will make easier life to the record managers or organizers in general and to users in particular.

The following properties are defined n5ui:hasCasefile, n5ui:hasCaseparty, n5ui:hasPrecedent, n5ui:hasClient, n5ui:hasCaseResponsible, n5ui:isPartOf, n5ui:isRelated and n5ui:hasPart.

On the other hand, records must or should have document descriptions element and/or document objects. Document description and document object contains metadata information of a given electronic documents. n5ui:isPartOf and n5ui :hasPart would related the document description and document object with its records/documents. However, sometimes there are some systems such as task systems, which do not contain document description element, in this case a direct semantic relationship between records

and document objects is possible using the n5ui:isPartOf property. See appendices 5, and 7 for RDF/XML formalization.

```
1: <rdf:RDF
2:          xmlns:n5ui="http://ark1.hioa.no/n5.ui/"
3:          xmlns:content="http://ark1.hioa.no/n5.ui/content/"
4:          xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
5:          xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#" >
6:
7:                    <n5ui:fonds rdf:nodeID="fonds">
8:                              <n5ui:systemID>3be6bd40-cc8a-4694-8f91-b0c61827cf14</n5ui:systemID>
9:                              <n5ui:isRelated rdf:nodeID="series"/>
10:                   </n5ui:fonds>
11:                   <n5ui:series rdf:nodeID="series">
12:                             <n5ui:systemID>f97c8684-7da1-4d5f-96ac-0fdbcf233fd1</n5ui:systemID>
13:                             <n5ui:isRelated rdf:nodeID="files"/>
14:                   </n5ui:series>
15:                   <n5ui:files  rdf:nodeID="files">
16:                             <content:hasCaseFile rdf:nodeID="casefile"/>
17:                             <n5ui:isRelated rdf:nodeID="records"/>
18:                   </n5ui:files>
19:                   <n5ui:records rdf:nodeID="records">
20:                             <n5ui:systemID>009e95e9-d12e-49f5-a3c1-bde05fba1f4f</n5ui:systemID>
21:                   </n5ui:records>
22:                   <n5ui:casefile rdf:nodeID="casefile">
```
……………………………………………………………………………

Figure 4.10:13 The Noark5 files, case files, case records, and records RDF/XML serialization see appendix 8 for further details, and appendix 9 the Noark5 and document object RDF/XML serialization

4.10.9. Classification system and Classes

The classification system has been widely known in the library and archives. The Noark5 standard also contains the classification systems and classes to classify the Noark5 records. In Noark5, there are two types of classification system called function-based classification system and object-based classification system. An organization that uses

87

functional based classification system can have one main function and subdivided functions. There are some main points for the reason of using function-based classification system to organize and arrange records entity, Noark5 has listed some of the reasons as following:-

- Documents that have been created as a result of the same activities are linked together. This gives the documents important contextual information.

- It simplifies the retrieval of files and documents.

- It can control access to documents. Certain classes can for example contain documents which must be screened.

- It can be a starting point for preservation and disposal. It is currently generally accepted that disposal decisions should be based on the organisation‟s functions and activities, and not on the content of the documents (macrodisposal).

The other main type of classification system is *object-based* classification. "The objects" will often be people, but they can also be companies, properties, etc. Unlike function-based classification systems, object-based systems are often flat, i.e. they consist of one level only (Noark5, 2012, p. 51). Moreover, the Classification system has sub part called class (see section 4.10.10), which is used for describing the data or records on the ERM and archives.

4.10.10.   Class

A classification system consists of classes. In the case of function-based (topic-based) classification, the classes will normally form part of a hierarchy, in which three or four levels are the norm. In the conceptual model, the sublevels are called "subclasses" and appear as a self-relation in Class[1]. ISO 15489 recommends that the classes describe the organisation's functions and activities (business processes). The uppermost level will then typically describe the main functions. The second level can describe sub functions,

---

[1] In the state administration‟s archive key, the levels below the classes are called main groups, groups and subgroups. In Noark 5, the common term „subclasses‟ is used for all sublevels.(Noark5, 2012, p. 51)

88

while the third level describes the processes (i.e. activities that are constantly repeated) (Noark5, 2012, pp. 51–52).

The classes must have a separate identification, which is unique within the classification system. This corresponds to what is called *order value* or *file code* in Noark4. Identifications from higher classes must be inherited downwards within the hierarchy, so that it is easy to see which level you are at[2]. Therefore, the classification system is found on the top level and the class is the subsidiary or subclass of the classification system.
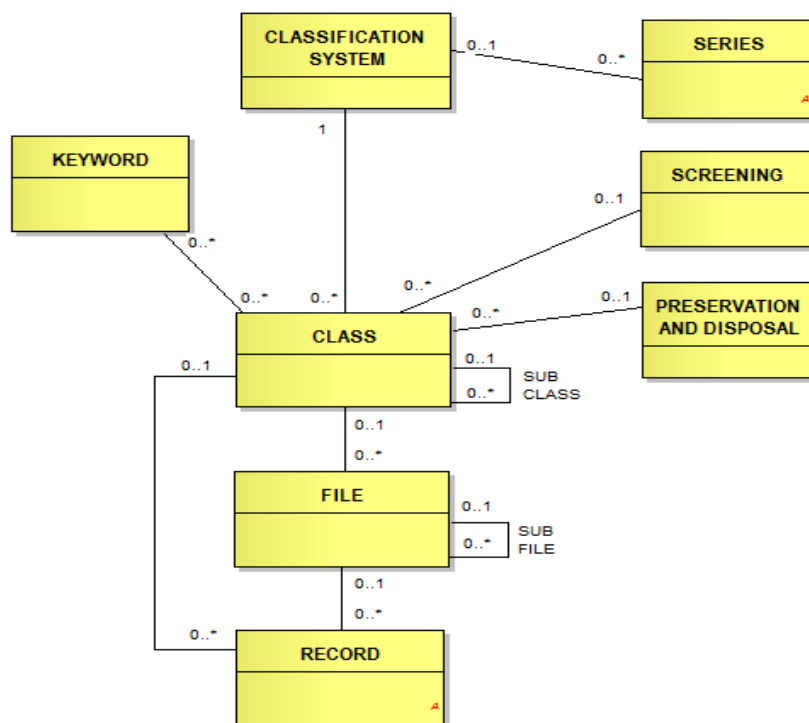
Figure 4.10:14Conceptual model for Classification System (Noark5, 2012, p. 53).

The class has significant role in classifying and arranging the records or documents based on their subject categories for instance from the K-codes for schools classification system, the high level is "0 – organization" and administration and the subclasses are "00 FELLES", "006 Likestilling", and "1 plan og Budsjet" etc. Therefore, using semantic

---

[2] Example of identification and title of classes and subclasses in the state administration"s standard key: First level: 2 Positions and personnel Second level: 2.3 Salaries and pensions Third level: 2.3.6 Employer"s contributions(Noark5, 2012, p. 52)

web for classification system could provide more suitable, flexible, and easy way of organizing the digital records. In this case, let us take one example of the "K-koder for skoler" and see how the semantic benefit the classifications systems.

K-koder for skoler

0 organisasjon og Administrasjon

000 FELLES

006 Likestilling

01 VALG, FOLKEAVSTEMNINGER

02 OPPGAVEFORDELING/FORHOLD STAT- KOMMUNE

03 ORGANISASJON

04 SAKSBEHANDLING

05 KONTORTJENESTER

050 Felles

054 kopiering, trykking

056 IT-tjeneste

06 INFORMASJON, ARKIV, BIBLIOTEK

060 Felles

07 FORHOLD UTAD

070 Felles

  etc.... (Arkivplan, 2012).

RDF/XML format of the classification system and classes

```
1: <rdf:RDF
2:        xmlns:n5ui="http://ark1.hioa.no/n5.ui/"
3:        xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
4:        xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#">
5:
6:        <n5ui:classificationsystem rdf:nodeID="k_koder">
7:                <n5ui:systemID>db2df75-c1af-43e8-839e-b3281cb18a70</n5ui:systemID>
8:                <n5ui:classificationtype>k-koder for skoler</n5ui:classificationtype>
```

```
9:                    <n5ui:createdDate>2011-10-11T17:11:26</n5ui:createdDate>
10:                   <n5ui:createdBy> Oslo maunicipality </n5ui:createdBy>
11:                   <n5ui:hasElement rdf:nodeID="class"/>
12:          </n5ui:classificationsystem>
13:          <n5ui:class rdf:nodeID="OrganisationogAdministration">
14:                   <n5ui:systemID>ac2d3375-b5b9-4359-8db7-b383397f4534 </n5ui:systemID>
15:                   <n5ui:classID>0 ORGANISASJON OG ADMINISTRASJON </n5ui:classID>
16:                   <n5ui:title>organisation og Administration</n5ui:title>
17:                   <n5ui:createdDate>2011-10-11T17:12:35</n5ui:createdDate>
18:                   <n5ui:createdBy>Oslo maunicipality</n5ui:createdBy>
19:                   <n5ui:isParentClassOf rdf:nodeID="Oppgavefordeling"/>
20:                   <n5ui:isParentClassOf rdf:nodeID="KONTORTJENESTER"/>
21:                   <n5ui:isElementOf rdf:nodeID="k_koder"/>
22:          </n5ui:class>
23:          <n5ui:Oppgavefordeling rdf:nodeID="Oppgavefordeling">
24:                   <n5ui:systemID>a686e335-6729-4677-ae86-8dde783ce29a</n5ui:systemID>
25:                   <n5ui:classID>02OPPGAVEFORDELING/FORHOLD STAT -
KOMMUNE</n5ui:classID>
26:                   <n5ui:title>OPPGAVEFORDELING/FORHOLD STAT - KOMMUNE
</n5ui:title>
27:                   <n5ui:createdDate> 2011-10-11T19:26:56 </n5ui:createdDate>
28:                   <n5ui:createdBy> Oslo maunicipality </n5ui:createdBy>
29:                   <n5ui:hasChildClass rdf:nodeID="Felles"/>
30:                   <n5ui:isSubClassOf rdf:nodeID="OrganisationogAdministration"/>
31:          </n5ui:Oppgavefordeling>
32:          <n5ui:KONTORTJENESTER rdf:nodeID="KONTORTJENESTER">
33:                   <n5ui:systemID>c27a5-5497-450a-83c2-5bf6fcd0f8bb</n5ui:systemID>
34:                   <n5ui:classID>05 KONTORTJENESTER</n5ui:classID>
35:                   <n5ui:title>KONTORTJENESTER</n5ui:title>
36:                   <n5ui:createdDate>2011-10-11T19:26:56  </n5ui:createdDate>
37:                   <n5ui:createdBy> Oslo maunicipality </n5ui:createdBy>
38:                   <n5ui:hasChildClassOf rdf:nodeID="kopiering"/>
39:                   <n5ui:isSubClassOf rdf:nodeID="OrganisationogAdministration"/>
40:          </n5ui:KONTORTJENESTER>
41:          <n5ui:Felles rdf:nodeID="felles">
42:                   <n5ui:classID>020 Felles</n5ui:classID>
43:                   <n5ui:title>Felles</n5ui:title>
44:                   <n5ui:hasFileType rdf:nodeID="files1"/>
45:                   <n5ui:hasFileType rdf:nodeID="files2"/>
46:                   <n5ui:hasFileType rdf:nodeID="files3"/>
47:                   <n5ui:isChildClassOf rdf:nodeID="Oppgavefordeling"/>
48:          </n5ui:Felles>
49:          <n5ui:kopiering rdf:nodeID="kopiering">
```

```
50:                    <n5ui:classID>054 Kopiering</n5ui:classID>
51:                    <n5ui:title>Kopiering</n5ui:title>
52:                    <n5ui:hasFileType rdf:nodeID="files"/>
53:                    <n5ui:isChildClassOf rdf:nodeID="KONTORTJENESTER"/>
54:        </n5ui:kopiering>
55:        <rdf:Description rdf:nodeID="fonds">
56:                    <n5ui:isRelated rdf:nodeID="series"/>
57:        </rdf:Description>
58:        <n5ui:series rdf:nodeID="series">
59:                    <n5ui:isRelated rdf:nodeID="files1"/>
60:                    <n5ui:isRelated rdf:nodeID="files2"/>
61:                    <n5ui:isRelated rdf:nodeID="files3"/>
62:                    <n5ui:isPartOf rdf:nodeID="fonds"/>
63:                    <n5ui:hasClassificationType rdf:nodeID="k_koder"/>
64:        </n5ui:series>
65: </rdf:RDF>
```

Figure 4.10:15 the classification system with classes RDF/XML serializations

Figure 4.10:16 graphical visualization of the classification system with classes RDF/XML serialization

4.10.11.   Summary

The main objective of this research is to explore and discover the possible options of integrating, mapping, and representing the Noark5 ERMS in the form of semantic web. The integration and modeling of Noark5 employs two approaches i.e. framework of the Noark5 records management and semantic web, and proposed prototype for guidance and better implementation. The proposed model, framework elaborates and explains the component of Noark5 and the semantic web. In addition, the framework approach and the prototype provides a general overview of the research work, including the Noark5 data outlook before the RDF modeling and mapping process and the standalone semantic web components (see Figure 3.6:1and Figure 3.7:1 for further details).

Moreover, the implementation makes use of the RDF serialization and RDF visualizations of the semantic web to map and represent the Noark5 extracted data as well as metadata elements. The main input source for this research was Noark5 documents, Noark5 records, Noark5 metadata elements, and Noark5 extracted XML elements. The RDF data modeling follows the triple rules, which is *subject-predicate-object* form. For instance in this case, the Noark5 records are the resources (subjects) part, the Noark5 metadata elements are the properties (predicates), which is used as a means to create semantic relationships, and the Noark5 outer core *object of interests* (the people for example the creator of the records, case files and registry) are the objects.

Therefore, the Noark5 extracted data are the sample data within the archive structure, and metadata elements of the Noark5 are modeled using the RDF/XML serialization and the parsing results were produced in terms of RDF data modeling and RDF graphic representation, which are both in terms of triple (subject-predicate-object) format. The RDF/XML serialization is more for machines consumption, it is readable, and understandable by machines whereas the graphical visualization is more valuable for human utilization.

 Furthermore, the research makes use of blank nodes identifier. The blank nodes are used in both subject and object of the RDF statements. Using Blank nodes vocabulary as subject and object helps to avoid ambiguity and confusion for the future work as this

implementation or research work lays the foundation for the Noark5 actual data representation and integration.

Moreover, the Noark5 recordkeeping standard has its own records or archives structure. Modeling the data or converting the Noark5 data into semantic web is possible without changing the overall Noark5 architecture such as the basic archives structure format i.e. fonds->series->files->records as all the records are organized and stored within this context see chapter five discussion part for further elaboration. Therefore, the implementation and integration of semantic web into the records management standard has been done without making change of the basic record structure format, which is more straightforward modeling and mapping from the Noark5 extracted data into the RDF data.

 In addition, the overall semantic integration and modeling process is able to ensure and help the archive structure to maintain and organize the electronic records in more easy and flexible way.

In practice, the records management has rules and regulations about the use of Noark5 inner core and Noark5 outer core solution. At some point, the requirements and standards of Noark5 allow flexible use of records and documents. That means, in Noark5 there are possibilities to recall and retrieve the Noark5 outer core overall data when it is necessary, however digging information about the records and the owner of the records can consume time.

Therefore, in general the *object of interest* is not visible or accessible alike the Noark5 records. Therefore, the implementation result shows the semantic RDF capability of creating meanings to the data, identifying and relating entities, exposing and make use of nested data in this case (the object of interest), which is never been the focus of the Noark5 system.

 In addition, mapping all the Noark5 extracted records, relating files with their case types, clients, and connecting and relating an archive and public administration organizations etc are some of the part that are modeled into the form of semantic.

In general, the following Bnodes and properties are used throughout the modeling.

Bnodes	:ufondscreator, :ufonds, :usubfonds, :useries, :ufiles, :urecords, :udeocumentdescription, :udocumentobject, :ufondentity, :ucase_party, :uprecedent_case, :ucase_handler, :ucase, :ucasefile, :ucaserecord, :uregistry_entry, :ucase_responsibility, :uorganisation, :uclient, :ugeneraldirector, :uregionalstatearchives, :uBergen, :uStavanger, :uTromso, :uTrondhiem, :uOslo, :uHamar, :ukongsberg, :ukristiansand, :uSamarkiiva,:u_koder,:classificationsystem, :UorganisationogAdministration, :class, :Uoppgavefordeling, :UKONTORTJENESTER, :Ufelles :Ukopiering.
Properties	n5ui:isCreated,n5ui:isCreatedBy,n5ui:isRelated,n5ui:hasPart,n5ui:issubPartOf,n5ui:belongsTo,n5ui:isPartOf,n5ui:hasCaseType,n5ui:isResponsibleFor, n5ui:hasCase,n5ui:hasCreator,n5ui:isOwnerOf,n5ui:hasOffice,n5ui:issharedTheSame,n5ui:hasDepartment,n5ui:hasGeneralDirector, n5ui:hasSecertary, n5ui:hasRegionalBranch,n5ui:hasCaseParty,n5ui:hasPrecedent,n5ui:hasClient,n5ui:hasCaseFile,n5ui:hasCaseResponsible,n5ui:hasCaseHandler, n5ui:hasElement,n5ui:isParentClassOf,n5ui:isElementOf, n5ui:hasChildClassOf,n5ui:isSubClassOf,n5ui:hasFileType,n5ui:isChildClassOf, n5ui:hasClassificationType.

Table 4.10.11-1   RDF Bnodes and properties

Moreover, there are much more capacities of semantic web that could enables us to create, expose, combine, and distribute meaningful data and information over the web. James Hendler *an artificial intelligence researcher and one of the originators of the semantic web*, has used the slogan "A little semantics goes a long way" in his several presentations and speeches in various conferences. I could not agree more, making use of little RDF/Ontology, would lead to use of huge amount of structured and semi structured data over the web, in this case various data such as public administration, government, and other related data can be searchable and accessible on the web.

Finally yet importantly, the study make use of RDF keywords, and describe Noark5 elements in terms of bnodes such as around 38 bnodes, and 32 properties were used to create semantic based data.

Last but not least, the overall implementation and modeling procedure has been accomplished with the main ground of the literature reviews, proposed framework, proposed prototype, and methodological approaches. In addition, the main concepts and

ideas of expressing, defining, identifying, and providing well-organized data can be achieved using the RDF (Resource Description Framework language) in more straightforward way, which do not contains more complex transformation and conversion of data. In addition, the research opens an opportunity for further study and use of additional language see Chapter six, section 6.2., for future work. Moreover, see chapter five for the discussion part that covered detailed explanation of why preserving or retaining the provenance and context of archive is necessary and important.

Chapter five

5. Discussion

In this discussion chapter, an explanation of the implementation, approach, and summarization of the research work results is provided. In addition, there is a detailed description of semantic web technologies and its relevant issues for the Noark5 ERMS and archives. Moreover, the discussion part of this study addresses the significance and essence of semantic web, the archive structure quality, provenance, and context, the Noark5 ERM and the benefit of semantic web technologies. In addition, the discussion provides an elaboration of the importance of retaining or preserving the provenance and context of the archives during the implementation period.

On the other hand, there shall be some points from the archivist and record manager point of view. Finally and importantly, it is very important to recall the research questions and objectives of the research, methodological approach, and discuss if the chosen methodological approach and framework is the right choice to achieve this study. Moreover, there shall be discussion with regards to how and what has been achieved throughout the thesis.

5.1. The ERM Noark5 and the benefit of Semantic Web Technologies

The Noark5 systems, standards, and requirements are developed for ensuring the electronic records quality and authenticity. The Director General of National archive has approved the records management standard (Noark5) as the current standard for electronic recordkeeping. For that reason, regional and National public administration and archives of Norway built their record collections based on the Noark5 solution. This indicates that the current official requirement and standard of Noark5 is a necessary and must for any public and private administration who aimed at organizing and providing an electronic record. Furthermore, there are roles and regulation that are designed globally by organizations such as ISO and ICA that should or must be followed in order to ensure the quality of several ERM systems as well as the electronic records.

At present, in Norway digital archives or electronic recordkeeping organizations are collecting, organizing, and making their digital born or digitized records available via the

Noark5 standards. Furthermore, the national archives of Norway have also update their own standards and requirements, which is Noark5 in order to preserve, secure, and ensure ERM systems, archives, and their record qualities.

Therefore, this research study is able to introduce the use and benefit of semantic web for records keeping management using the RDF data modeling. The main reason for integrating semantic web for Noark5 is first that, the semantic web technology is nowadays used for facilitating, expanding, merging, and distributing data/information on the web. In addition, the capability of semantic web enables several governments, public agencies, and other information providers to open their data for better usability and accessibility by their intended users'. Secondly, because the researcher believed that most software applications that are developed for particular or general purposes have the nature of securing and locking the data, therefore this default problem of most applications prevent users from accessing the right information at the right time, therefore the semantic web has capability of solving such issues by unlocking information from the system. Thirdly, using semantic web benefit an organization one-way to enhance data usability by their users the other way to improve their services. The fourth point is that, using such kinds of technology in the archive and records management areas can increase the possibility of retrieving relevant information using semantic based search engines. Finally yet importantly, the semantic web can be used to solve interoperability problems between systems and applications for instance *semantic interoperability* could enable various ERM systems to exchange and share diverse sources between them.

In general, the RDF language is used to model the records that are extracted from the Noark5 core. These electronic records are extracted from Noark5 for the sake of long-term preservation. Therefore, the XML extracted records was mapped into the form of RDF using the RDF/XML serialization. The RDF modeling was implemented while by retaining the provenance and context aspects of the Noark5 record structure, which are fonds-series-files-records. The main reason for retaining the archives traditional structure is initially, the objective of the study is not to change the Noark5 archive structure rather it aims to introduce the use of semantic web tools and techniques in the area of ERM and archives. For that reason, the implementation tries to model the data without influencing the archive structure and provenance (see section 5.3.), for more explanation and reason

for why retaining the traditional archival structure such as provenance and context is very essential. Secondly, all public records and archives in Norway built their digital and physical records based on the Noark5 archives structures and therefore, changing the context might bring difficulties and problems such as data loss. Therefore, the modeling procedure of the Noark5 sample data has achieved with no change of the basic structure, which is more straightforward implementation.

5.2.   The archive structure quality, provenance, and context

As the research attempts to address some of the main points for retaining the archive structure quality, provenance, and context is that first and foremost, completely changing or converting the archive structure into a semantic web one can have a unknown impact and more demerit than its merit in every aspects of the ERM and archives. For instance, losing data and metadata description, misuse of format, more investment of time, effort, money etc. In addition, it is also true that the archivist and the records managers or the ERM and archive organizations, they cannot simply change and accept the usage of semantic web technology especially if the semantic web takes the data out of the archival structure. However, there shall be the possibility of building the semantic layer on top of the archive structure or using the semantic web in parallel with the archive structure in similar way as we tried to represent and model the Noark5 extracted data into the RDF data. In general, it is essential to maintain and ensure the overall traditional quality of archives, roles, and regulations, as well as archival structures. As a result, with respect of all the above facts it is possible to make use of semantic web technology.

On the other hand, retaining the archival provenance and context might brought its own challenges and problems into the semantic web arena where we aim to make use of additional languages of the semantic web such as LOD which is mainly focus in linking, and publishing RDF data on the web. For that reason, there constraints of the archive structure provenance and context can be a challenge for semantic web.

5.3.   The essence of semantic web technology

The overall implementation and result of the current study are able to verify and show the importance of semantic web for archives and electronic records managements (ERMs). The overall modeling and representing data allows the Noark5 record to be more usable and understandable by both humans and machines. In addition to that, representing the Noark5 records in a semantic web based form could make life easier for both records managers and archivists in particular and for the archive and ERM organizations in general through allowing them to use or integrate other semantic based sources either from internal or external source. That means, the ERM and archive organizations would be able to make use of external semantic sources and can easily integrate these external sources with Noark5 systems. In addition, maintaining and ensuring semantic interoperability in the ERM could enable them to interchange data between applications without any interruption for instance exchanging data between the different regional public administrations of Norway. Furthermore, it introduces and enables an organization to model, link, publish, and access the ERM and archival data. For instance, there shall be a possibility of using some part of the Noark5 records or data on the web mainly information that are not sensitive or records that are officially available for users on the web. It is extremely important to note that, data or information that are sensitive will continue to be secured and protected and shall not be accessed or used over the web. Therefore, that means we are addressing the resources that are made accessible and usable by the ERM organizations.

In general, the main benefit of semantic web is to enable ERM and archive organizations to make use of semantic web technologies, semantic web external sources and to enhance the use of related technologies. Moreover, data exchangeability, usability, interoperability, accessibility, and retrieval of relevant and well-defined documents are highly supported by the semantic web technology.

5.4. The general implementation and output of the thesis

In this section of chapter five, it is very important to recall the research questions and objectives of the research, methodological approach, and discuss how the question has been addressed and if the chosen methodological approach, prototype, and framework is the right choice to achieve this thesis. In addition, the discussion emphasis more in the research questions and how and in what context has been able the researcher to respond the questions and address the statement problems.

5.4.1. How can Noark5 metadata be mapped to the RDF

Mapping Noark5 metadata to RDF is achieved using the appropriate principle and roles of RDF semantics. RDF is extensively known for describing and expressing metadata elements. For that reason, it was relatively straightforward to add semantic into the metadata or to make the metadata elements semantic based sources. Furthermore, the Noark5 metadata was mapped to the RDF based data using the RDF/XML serialization and adding the semantic to each Noark5 metadata elements are achieved using the Noark5 URI identifier. In addition, during the mapping procedure, more than sixty metadata elements are mapped into the form of RDF.

5.4.2. How can RDF be integrated in electronic records management system

Electronic records management systems are systems that are used for storing data. Records managers are the one who create data, organize, and sort data on the ERMS. Therefore, applying semantic web tools, techniques, and languages are able to facilitate and improve the ERM services. The overall mapping of metadata and sample data of Noark5 are made manually and has shown positive results, for instance Noark5 data before the RDF implementation was only known as electronic records that have metadata description and users for instance can retrieve and access the digital resource from the Noark5 ERMS.

Nevertheless, as we tried to provide an example in (section 1.5., and 2.1.) there are no options and ways of retrieving detailed information of an artist or a writer, which in our case is the nested data *objects of interest*. Or for instance, the writer his/her books, literatures, and what the writer were doing last summer and if he/she was fun of Manchester united in addition if he/she has been registered record/information in any other ERM organizations or if the writer was just only have registered information in the Norwegian ERM.

Therefore, throughout the implementation the RDF modeling and mapping procedure are able to show and confirm that the actual Noark5 ERMS and their data can make use of semantic web by mapping, modeling, and integrating semantic so that users will be able to retrieve relevant and detailed information from the Noark5. For instance, a parent applicant who have a case for kindergarten, which is in our case (the object of interest) and the files of this applicants was embedded within the system, and therefore, now the use of RDF enables us to retrieve information in relation to his/her application case as well as the objects. Therefore, making the Noark5 data semantic based then can easily enable them to trigger data from their system as well as could enable them to use external semantic sources.

5.4.3. *How can RDF be used to represent Noark5 metadata and records and retain the traditional archive qualities such as provenance and context*

The classical archive structure, Noark5 metadata, Noark5 extracted data are modeled using the RDF/XML data modeling as well as graphical representation into the RDF. In addition, it is easiest to identify, and understand the type and meaning of the Noark5 data and we are able to address issues in relation to the data that are nested within the archival structure and system. It is possible to access electronic records using Noark5 based system but it is harder to access the objects of interests. In fact, there is no concept of object-of-interest. Therefore, the implementation step modeled the *objects of interest* and the Noark5 extracted data along with the fonds structure.

As a result, the RDF based objects of interest and Noark5 records are combined and related to each other using the new defined RDF property name or predicate. In general,

the overall process is not only model and map the data and the people (objects of interest) but it also able to create semantic relations between the subject, which is in this case the electronic records of Noark5 and the object, which is the people for instance can be the person who create the records or cases.

Moreover, the semantic based metadata and the new defined properties enable us to relate these two types of resources, and these properties are known as predicate in RDF triples. Therefore, the provenance and context of the archives are preserved and the overall implementation, i.e. representation of Noark5 metadata and records are handled in straightforward ways.

5.4.4. *The methodological approach and contributions*

The research are able to introduce semantic web technologies for the ERMS and the overall results indicates that the archive and ERM material can easily be identified and accessed by end users through semantic web. In addition, the present RDF data model result can be used for facilitating data integration and data exchangeability within various systems and can be used as a basis for creation of better search algorithms. On the other hand, there shall be a possibility of creating RDF base extraction tools that can help the ERM organizations to transfer their Noark5 semantic based electronic records into the archives (see chapter six for detailed and future work).

Therefore, the methodological approach enables the researcher to explore, discover, and understand the existing documents, data, codes, systems, standards, and requirements of the Noark5 and the semantic web technology elements. As a result, understanding these data, constructing, and interpreting data are able to accomplish using the proposed framework and prototype. This indicates that, the chosen methodology, the proposed framework and prototype was the right choice for the research theme, which was to introduce and integrate semantic web technology for the Noark5 ERMS and archive institutions.

Chapter six

6. Conclusions and future work

6.1. Conclusion

In keeping with the times, the public administrations and archives are now changing their classical records into the digital format by digitizing and scanning their psychical records into the form of electronics, therefore the current generations are able to accesses and use digital resources. In addition, organizations such as record keeping, RM, libraries, and archives are facilitating and increasing their data usability through collecting and organizing digital born resources and are working hard in making these data available and accessible on the web. Thanks to the rapid growth of technology, nowadays even the knowledge societies are also able to use several types of information that are oldest and latest on the web using information communication technologies (ICT). Several ICT mediums are used to communicate, interact, upload, share, and access digital information and knowledge online with families, friends, colleagues, with the rest of the world societies. Consequently, several commercial and non-commercial software and applications are also able to facilitate data distribution and dissemination over the web using semantic web and related technologies.

The Noark5 system is an electronic records management system that is used by Norwegian public administration for storing, organizing, and updating their electronic records. For that reason, all regional and national electronic records management and archives are used the Noark5 system, which was approved by the General Directory of Norway for managing and handling public and private electronic records.

The use of semantic web technologies for Electronic records management allows the public administrations and archives to organize, define, and assign meaning to the data. In addition, the technology facilitates and enhances the use of system, which is more user and information centric than system centric that means introducing semantic web not only make data more usable but also create easy interaction between the information providers and the users, between the users, information providers and the machines. That means, for instance, users can request information or access to records that was already

archived ten years ago and the information provider or the users him/her self are able to generate well-defined and relevant data from the archive databases.

More or less, in order to achieve easy semantic based data sources, semantic based search engines, and semantic based services, the public and private administration both should be able to consider of using open source software and applications in addition to the semantic web technologies.

The research work describes the possibility of using semantic web technologies (RDF) for Noark5 records management standard. The overall RDF modeling and implementation process is able to achieve RDF based data using the Noark5 extracted data and the open source software system. The result then verifies the benefit and applicability of semantic web technology for Noark5 ERM and archives. The semantic web languages such as RDF and some Owl are used to create semantic based relationships between people, organizations, and objects. Therefore, the overall result indicates that the ERM/archives can use semantic web technologies for describing, expressing, conceptualizing, modeling, and publishing electronic records or digital resources that are entitle by the public administration for access and use on the web.

In general, the semantic web has the following benefits for both ERM and archives.

- Using the RDF language in ERM provides an opportunity to use semantic based Noark5 records
- Semantic web technologies and its languages could enable ERM/archives to incorporate external semantic web sources into the Noark5 system such as DC, FOAF and others.
- It allows organizations to allocate, link, express, and create semantic based relationships between people (object of interest), resources, place, and events.
- The current modeled data, which is Noark5 extracted data can be used to model the actual Noark5 data in addition, could enable them to establish semantic interoperability between various application and databases.

- The present status can enable records managers or archivists to map the RDF data into more rich, advanced, and globally used languages such as ontology and linked open data (LOD).

Last but not least, the semantic web enables us to expose nested Noark5 object data from the system i.e. *object of interest,* which was embedded deep down in between the fonds archive structure and the systems. In addition to that, allow us to identify and understand the semantic relationship between those people, data, and places.

Therefore, despite the above fact the Norwegian records management organizations and archives can map and model their actual data into the form of semantic web. The current open source software development for Noark5 could enable them to adapt semantic based sources and semantic interoperability between several Noark5 public and private databases.

Moreover, the research work provides detailed description and explanation of the Semantic web for the ERMS within appropriate methodological approach, theoretical framework, literature review, and data modeling procedure. Therefore, to conclude the actual data of Noark5 can be represented, modeled, shared, merged, distributed, exchanged, linked, and published using the semantic web technologies and languages such as RDF, OWL, and LOD. Section 6.2., provides detailed description of future work that contains important points in regard the use of additional semantic languages and other suggested frameworks.

6.2. Future work

The research result is able to explore and describe the importance of semantic web technology for ERM and archives using the RDF. Representing the extracted data and metadata of Noark5 has been mapped and visualized accordingly. The general concept and vision of semantic web, which is creating, combing, relating, and modeling the Noark5 extracted data verifies the essence of semantic web in increasing data understandability, usability, and accessibility.

The overall results demonstrate the importance of semantic web that enable both the archives and ERMs to create semantic relations between entities and semantic interoperability between different databases. As a result, there shall be possibility of opening and modeling the actual resource of public administrations, so that their data can be more open and accessible. Moreover, additional languages for instance RDFS, Owl, and LOD can be used to expand the current RDF based data. Therefore using these languages the RDF data can be mapped into ontology-based data that is richer than RDF, which can enable the digital resource globally published and linked on the web with other external sources.

In addition, the public administration could think of adapting policy and regulations for opening and accessing data, and can enhance open source software and applications usability in order to provide quality data and service. The following (Figure 6.2:1) depicts an example of the eight regional state public administration and archives of Norway. In addition, the (Figure 6.2:1) describe that, integrating SW for ERMS can benefit record managers, archivists, and in general ERM organizations for better data access.

 In addition, the use of semantic web in this area will provide the capacity for transferring semantic based electronic records into the archives. As a result, the archives are then will be able to gather and preserve semantic based digital sources from the ERMs, in this case archive organizations are able to centralize, combine semantic based data from the electronic records management. In general, both ERMs and archives are able to create strong collaboration and interaction nationally as well as internationally with several ERM organizations. In addition, they can maintain, centralize, decentralize, merge, distribute, interchange, and retrieve data that has been archived in the past or data that are currently in use.
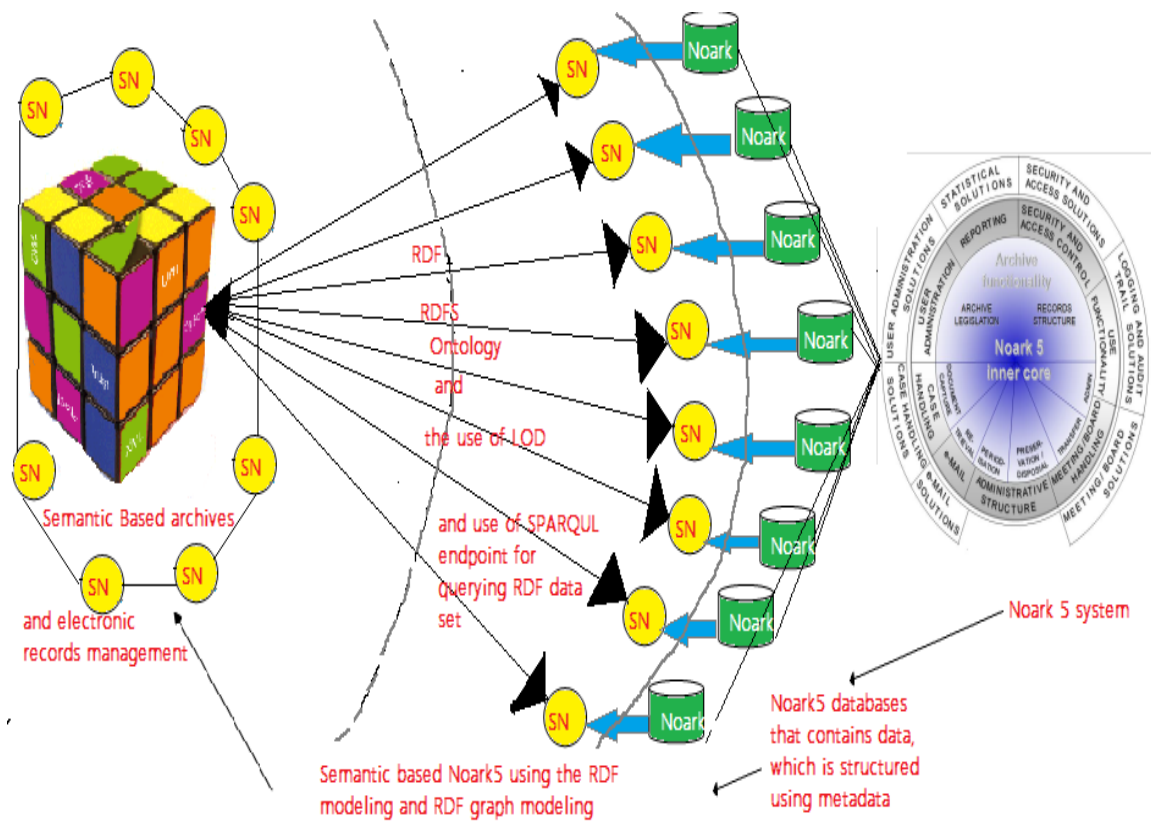
Figure 6.2:1 Semantic based ERMs and archives preview of the semantic data and database when it centralized and decentralized (Noark5 and SW picture are taken from Noark5, 2012; Semantic web, 2012)

Moreover, the following the above semantic based ERMs and archives framework can be adapted using the current RDF data. The following are some of the future works that are suggested based on the result of the present study.

- Designing semantic web approach at organizational level in order to use and expand semantic web technology for the records management.

- Mapping, modeling, representing, and publishing the actual Noark5 data using the semantic web languages RDF, ontology, and LOD in order to expose, share and use meaningful and structured Noark5 RDf based data.

- Using automated semantic web tools and techniques for better modeling, representing, and formalizing huge and complex data such as using Jena framework, KOAN, sesame, and JRDF semantic tools.

- Designing or adapting RDF based extraction and querying tools in order to integrate easy search and retrieval functions to the archives/ERMS.

- Designing an evaluation approach for evaluating and analyzing the RDF data that are modeled or mapped into RDF.

- Integrate external semantic web sources into the Noark5 systems

- Using the linked open data (LOD) for linking, connecting, sharing the RDF data into the web with other external data.

- Designing semantic based search engine algorithms and querying tools for searching, and retrieving Noark data.

7. References

Allemang, D., & Hendler, J. (2008). *Semantic Web for the Working Ontologist: Effective Modeling in RDFS and OWL* (1st ed.). Morgan Kaufmann.

Antoniou, G., & Harmelen, F. van. (2008). *A semantic web primer*. Cambridge, Mass.: MIT Press.

Arkivplan. (2012). K-koder for skoler / Oppsett saksarkiv / Ordningssystemer / Reglement / Arkivplan Randaberg kommune / Randaberg kommune / IKA Rogaland / Arkivplan - Arkivplan. Retrieved June 7, 2012, from http://stranda.arkivplan.no/content/view/full/9953

Australia, N. A. of. (2012). Glossary. Retrieved June 21, 2012, from http://www.naa.gov.au/records-management/publications/glossary.aspx

Badr, Y., Chbeir, R., Abraham, A., & Hassanien, A.-E. (2010). *Emergent Web Intelligence: Advanced Semantic Technologies*. Springer.

Barbera, M., Cortese, C., Zitarosa, R., & Groppo, E. (2009). Building a Semantic Web Digital Library for the Municipality of Milan. *Rethinking Electronic Publishing: Innovation in Communication Paradigms and Technologies-Proc. 13th International Conference on Electronic Publishing* (pp. 133–154). Retrieved from http://elpub.scix.net/data/works/att/156_elpub2009.content.pdf

Benjamins, V., Contreras, J., Blázquez, M., Dodero, J., Garcia, A., Navas, E., Hernandez, F., et al. (2004). Cultural Heritage and the Semantic Web. In C. Bussler, J. Davies, D. Fensel, & R. Studer (Eds.), *The Semantic Web: Research and Applications*, Lecture Notes in Computer Science (Vol. 3053, pp. 433–444). Springer Berlin / Heidelberg. Retrieved from http://www.springerlink.com/content/nqbe4me9rvtwfwf4/abstract/

Berners-Lee, T. (1998). Semantic Web: Why RDF is more than XML.

    http://www.w3.org/DesignIssues/RDF-XML.html. Retrieved June 3, 2012, from

    http://www.w3.org/DesignIssues/RDF-XML.html

Bonanos, P. (2009, May 19). How the Semantic Web Will Redefine Digital Music. *GigaOM*.

    Retrieved May 16, 2012, from http://gigaom.com/2009/05/19/at-sfmusictech-the-

    semantic-web-sneaks-onstage/

Breslin, J. G., Passant, A., & Decker, S. (2009). *The Social Semantic Web* (1st ed.). Springer.

Carminati, B., Ferrari, E., & Thuraisingham, B. (2004). Using RDF for policy specification and

    enforcement. *15th International Workshop on Database and Expert Systems*

    Applications, 2004. Proceedings (pp. 163– 167). Presented at the 15th International

    Workshop on Database and Expert Systems Applications, 2004. Proceedings, IEEE.

    doi:10.1109/DEXA.2004.1333467

Cassell, C., & Symon, G. (2004). *Essential guide to qualitative methods in organizational*

    research. SAGE.

Davies, J., Fensel, D., & Harmelen, F. van. (c2003). *Towards the semantic web: ontology-driven*

    knowledge management. Chichester: Wiley.

Engels, R. H. P., ESIS, & Tønnesen, J. R. (2007). Case Study: A Digital Music Archive (DMA)

    for the Norwegian National Broadcaster (NRK) using Semantic Web techniques.

    Retrieved May 16, 2012, from http://www.w3.org/2001/sw/sweo/public/UseCases/NRK/

Fensel, D., Van Harmelen, F., Ding, Y., Klein, M., Akkermans, H., Broekstra, J., Kampman, A.,

    et al. (2002). On-to-knowledge: Semantic web enabled knowledge management.

    submitted to IEEE Computer. Retrieved from

    http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.17.1246&rep=rep1&type=pdf

Han, L., Finin, T., Parr, C., Sachs, J., & Joshi, A. (2008). RDF123: from Spreadsheets to RDF.

    The Semantic Web-ISWC 2008, 451–466.

Keeran, P., & Levine-Clark, M. (c2007). *Research within the disciplines: foundations for reference and library instruction*. Lanham, Md.: Scarecrow Press.

Kruk, S. R., & McDaniel, B. (2008). *Semantic Digital Libraries*. Springer. Retrieved from http://www.springer.com/computer/database+management+%26+information+retrieval/book/978-3-540-85433-3

Lee, T. B., Hendler, J., & Lassila, O. (2001). The semantic web. *Scientific American*, *284*(5), 34–43.

Mazzini, S., & Ricci, F. (2011). EAC-CPF Ontology and Linked Archival Data.

MusicOntology. (2010). Music Ontology Specification. Retrieved May 16, 2012, from http://musicontology.com/

Myers, M. D. (2008). *Qualitative Research in Business & Management*. SAGE Publications Ltd.

Mylopoulos, J. (1980). An overview of Knowledge Representation. *SIGPLAN Not.*, *16*(1), 5–12. doi:10.1145/960124.806869

Noark 5. (2012). Noark 5 English / Noark / Public Sector / Home - Riksarkivet og statsarkivene - Arkivverket. Retrieved February 13, 2012, from http://www.arkivverket.no/eng/Public-Sector/Noark/Noark-5-English

Noark-Arkivverket. (2011). Noark - Arkivverket. Retrieved March 24, 2012, from http://www.arkivverket.no/eng/Public-Sector/Noark

Passin, T. B. (2004). *Explorer's guide to the semantic web*. Greenwich, Conn.: Manning.

Patel-Schneider, P. F. (2004). The semantic web and knowledge representation. *Computer Science, 2004. ENC 2004. Proceedings of the Fifth Mexican International Conference in* (p. 7). doi:10.1109/ENC.2004.1342582

Peter, M. (2007). *Social Networks and the Semantic Web*. Retrieved from http://www.springer.com/computer/database+management+%26+information+retrieval/book/978-0-387-71000-6

Pickard, A. J. (2007). *Research Methods in Information* (illustrated ed.). Facet Publishing.

113

Raimond, Y., Abdallah, S., Sandler, M., & Giasson, F. (2007). The music ontology. *Proceedings of the International Conference on Music Information Retrieval* (pp. 417–422).

Semanticweb. (2012). Learning - semanticweb.com. Retrieved June 20, 2012, from http://semanticweb.com/category/learning

Styles, R., Ayers, D., & Shabir, N. (2008). Semantic MARC, MARC21 and the Semantic Web. *WWW 2008 17th International World Wide Web Conference*.

The national archives. (2007). Noark-4 / Noark / Public Sector / Home - Riksarkivet og statsarkivene - Arkivverket. Retrieved June 3, 2012, from http://www.arkivverket.no/eng/Public-Sector/Noark/Noark-4

The national archives. (2011). Public Sector - Arkivverket. Retrieved June 3, 2012, from http://www.arkivverket.no/eng/Public-Sector

Thefreedictionary. (2012). pyramid - definition of pyramid by the Free Online Dictionary, Thesaurus and Encyclopedia. Retrieved June 21, 2012, from http://www.thefreedictionary.com/pyramid

Unicode. (2012). What is Unicode? Retrieved March 16, 2012, from http://unicode.org/standard/WhatIsUnicode.html

W3C. (2004a). RDF Primer. Retrieved February 21, 2012, from http://www.w3.org/TR/rdf-primer/

W3C. (2004b). Resource Description Framework (RDF): Concepts and Abstract Syntax. *http://www.w3.org/TR/rdf-concepts/#section-data-model*. Retrieved June 2, 2012, from http://www.w3.org/TR/rdf-concepts/#section-data-model

W3C. (2004c). Architecture of the World Wide Web, Volume One. Retrieved March 16, 2012, from http://www.w3.org/TR/webarch/#identification

W3C. (2004d). Resource Description Framework (RDF): Concepts and Abstract Syntax. *w3c recommendation*. Retrieved June 2, 2012, from http://www.w3.org/TR/rdf-concepts/#section-data-model

W3C. (2004e). RDF/XML Syntax Specification (Revised). *w3c recommendation*. Retrieved May

24, 2012, from http://www.w3.org/TR/REC-rdf-syntax/

W3C. (2007). Web Ontology Language OWL / W3C Semantic Web Activity. Retrieved May 23,

2012, from http://www.w3.org/2004/OWL/

W3C. (2011). W3C Semantic Web Activity. Retrieved February 28, 2012, from

http://www.w3.org/2001/sw/

Yu, L. (c2007). *Introduction to the Semantic Web and Semantic Web services*. Boca Raton, Fla.:

Chapman & Hall/CRC.

Zaihisma Che Cob, & Abdullah, R. (2008). Ontology-based Semantic Web services framework

for knowledge management system. *Information Technology, 2008. ITSim 2008.

International Symposium on* (Vol. 2, pp. 1 –8). doi:10.1109/ITSIM.2008.4631671

8. Appendices

8.1. Appendix 1: Noark5 arkiv structure in the form of XML (Noark5 XML)

```
<arkiv

xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"

xmlns="http://www.arkivverket.no/standarder/noark5/arkivstruktur"

xmlns:n5mdk="http://www.arkivverket.no/standarder/noark5/metadatakatalog">

<systemID>d7ee5939-4122-4b80-9e1e-057dac4f3d29</systemID>

<tittel>Arkiv</tittel>

<beskrivelse>Arkiv</beskrivelse>

<arkivstatus>Avsluttet</arkivstatus>

<dokumentmedium>Elektronisk arkiv</dokumentmedium>

<opprettetDato>2012-02-21T20:32:39</opprettetDato>

<opprettetAv>system</opprettetAv>

<avsluttetDato>2012-02-22T14:38:29</avsluttetDato>

<avsluttetAv>system</avsluttetAv>

<arkivskaper>

  <arkivskaperID>964963193</arkivskaperID>

  <arkivskaperNavn>Test Name</arkivskaperNavn>

  <beskrivelse>Test Name Archive</beskrivelse>

</arkivskaper>

<arkivdel>

  <systemID>b23fd11e-1860-4b84-a449-c77c86357330</systemID>

  <tittel>Arkivdel</tittel>

  <beskrivelse>Arkivdel</beskrivelse>

  <arkivdelstatus>Avsluttet periode</arkivdelstatus>

  <dokumentmedium>Elektronisk arkiv</dokumentmedium>

  <opprettetDato>2012-02-21T20:32:56</opprettetDato>

  <opprettetAv>system</opprettetAv>

  <avsluttetDato>2012-02-22T14:38:29</avsluttetDato>

  <avsluttetAv>system</avsluttetAv>

  <mappe>

   <systemID>e05862f1-11ac-4b06-9228-a18c0005b205</systemID>

   <mappeID>uninitialized</mappeID>

   <tittel>Mappe</tittel>

   <opprettetDato>2012-02-21T20:33:12</opprettetDato>

   <opprettetAv>system</opprettetAv>

   <avsluttetDato>2012-02-22T14:38:29</avsluttetDato>

   <avsluttetAv>system</avsluttetAv>

   <registrering>

    <systemID>59d9cff4-32ff-4e3e-a5e7-ac0b6a6edc71</systemID>

    <opprettetDato>2012-02-21T20:33:28</opprettetDato>

    <opprettetAv>system</opprettetAv>

    <arkivertDato>2012-02-22T14:38:29</arkivertDato>

    <arkivertAv>system</arkivertAv>

    <dokumentbeskrivelse>

      <systemID>087d50fc-55d1-41e4-9ec6-2913a38e5270</systemID>
```

```
<dokumenttype>application/pdf</dokumenttype>
<dokumentstatus>Dokumentet er ferdigstilt</dokumentstatus>
<tittel>TestDok</tittel>
<beskrivelse>TestDok</beskrivelse>
<opprettetDato>2012-02-21T20:34:01</opprettetDato>
<opprettetAv>system</opprettetAv>
<tilknyttetRegistreringSom>Vedlegg</tilknyttetRegistreringSom>
<dokumentnummer>1</dokumentnummer>
<tilknyttetDato>2012-02-22</tilknyttetDato>
<tilknyttetAv>system</tilknyttetAv>
<dokumentobjekt>
  <versjonsnummer>1</versjonsnummer>
  <variantformat>Produksjonsformat</variantformat>
  <format>pdf</format>
  <opprettetDato>2012-02-21T20:34:01</opprettetDato>
  <opprettetAv>system</opprettetAv>
  <referanseDokumentfil>f4489063-2aa5-4c5f-b0ab-1c0535e41e3c.pdf</referanseDokumentfil>
 <sjekksum>3778d80a72b68f85b7b8f5a811c4168e72bcba1d4e275fcc0680788f84831483</sjekksum>
    <sjekksumAlgoritme>SHA256</sjekksumAlgoritme>
    <filstoerrelse>90708</filstoerrelse>
  </dokumentobjekt>
 </dokumentbeskrivelse>
 </registrering>
 </mappe>
 </arkivdel>
</arkiv>
```

The above XML data representation translated

```
<!—fonds and fondscreator in xml format →
<?xml version="1.0" encoding="UTF-8"?>
<!—xml declaration -- >
<fonds description="fonds">
        <systemID>d7ee5939-4122-4b80-9e1e-057dac4f3d29</SystemID>
        <title>fond creation</title>
        <description>new fond document</description>
        <fondsstatus>created</fondsstatus>
        <documentmedium>electronic format</documentmedium>
        <createdDate>2012-02-21T20:32:39</createdDate>
        <createdBy>Oslo municipality</createdBy>
        <finalisedDate>2012-02-22T14:38:29</finalisedDate>
        <finalisedBy>Oslo municipality</finalisedBy>
<fondscreator>
        <fondscreatorID>964963193</fondscreatorID>
        <fondscreatorName> Oslo municipality</fondscreatorName>
        <description> new document creation</description>
</fondscreator>
        <series>
```

117

```
<systemID>b23fd11e-1860-4b84-a449-c77c86357330</systemID>
<title>series</title>
<description> series creation</description>
<recordsectionstatus>closed</recordsectionstatus>
<documentmedium>electronic format</documentmedium>
<createdDate>2012-02-21T20:32:56</createdDate>
<createdBy>Oslo municipality</createdBy>
<finalisedDate>2012-02-22T14:38:29</finalisedDate>
<finalisedBy>Oslo municipality</finalisedBy>
<recordsperiodStartDate>2012-02-21T20:32:56</recordsperiodStartDate>
<recordsperiodEndDate>2012-02-22T14:38:29</recordsperiodEndDate>
<file>
        <systemID>e05862f1-11ac-4b06-9228-a18c0005b205</systemID>
        <fileID>12/125463</fileID>
        <title>file</title>
        <description>new file</description>
        <createdDate>2012-02-21T20:33:12</createdDate>
        <createdBy>Oslo municipality</createdBy>
        <finalisedDate>2012-02-22T14:38:29</finalisedDate>
        <finalisedBy>Oslo municipality</finalisedBy>
        <record>
                <systemID>59d9cff4-32ff-4e3e-a5e7-ac0b6a6edc71</systemID>
                <createdDate>2012-02-21T20:33:28</createdDate>
                <createdBy>Oslo municipality</createdBy>
                <archiveDate>2011-02-25</archviedDate>
                <archviedBy>Oslo municipality</archivedBy>
                <documentdescription>
                <systemID>087d50fc-55d1-41e4-9ec6-2913a38e5270</systemID>
                <documenttype>application/pdf</documenttype>
                <documentstatus>completed</documentstatus>
                <title>record</title>
                <description>application document</description>
                <createdDate>2012-02-21T20:34:01</createdDate>
                <createdBy>Oslo municipality</createdBy>
                <linkedrecordAs>external</linkedrecordAs>
                <documentnumber>1</documentnumber>
                <linkedDate>2012-02-22T20:56:00</linkedDate>
                <linkedBy>Oslo municipality</linkedBy>
                <documentobject>
                        <versionnumber>2.2</versionnumber>
                        <varianformat></variantformat>
                        <format>pdf</format>
                        <formatDetails>production format</formatDetails>
                        <createdDate>2012-02-21T20:34:01</createdDate>
                        <createdBy>Oslo municipality</createdBy>
        <referncedocumentdescription>document object</referencedocumentdescription>
                        <referncedocumetfile>124535fy</referncedocumentfile>
```

```xml
<referenceDocumentfile>f4489063-2aa5-4c5f-b0ab-1c0535e41e3c.pdf</referenceDocumentfile>
<checksum>3778d80a72b68f85b7b8f5a811c4168e72bcba1d4e275fcc0680788f84831483</checksum>
<checksumAlgoritme>SHA256</checksumAlgoritme>
<filesize>90708</filesize>
</documentobject>
</documentdescription>
</record>
</file>
</series>
</fonds>
```

8.2. Appendix 2 : Fonds and Fondscreator: Parsing output 31 triples

Subject	Predicate	Object
genid:Ufondscreator	http://www.w3.org/1999/02/22-rdf-syntax-ns#type	http://ark1.hioa.no/n5.ui/fondscreator
genid:Ufondscreator	http://ark1.hioa.no/n5.ui/fondscreatorID	"964963193"
genid:Ufondscreator	http://ark1.hioa.no/n5.ui/fondscreatorName	"Oslo municipality"
genid:Ufondscreator	http://ark1.hioa.no/n5.ui/description	"creating new fond"
genid:Ufondscreator	http://ark1.hioa.no/n5.ui/isCreated	genid:Ufonds_1
genid:Ufondscreator	http://ark1.hioa.no/n5.ui/isCreated	genid:Ufonds_2
genid:Ufondscreator	http://ark1.hioa.no/n5.ui/isCreated	genid:Ufonds_3
genid:Ufonds_1	http://www.w3.org/1999/02/22-rdf-syntax-ns#type	http://ark1.hioa.no/n5.ui/fonds
genid:Ufonds_1	http://ark1.hioa.no/n5.ui/systemID	"d7ee5939-4122-4b80-9e1e-057dac4f3d29"
genid:Ufonds_1	http://ark1.hioa.no/n5.ui/title	"fond1"
genid:Ufonds_1	http://ark1.hioa.no/n5.ui/description	"New fond"
genid:Ufonds_1	http://ark1.hioa.no/n5.ui/fondsstatus	"closed"
genid:Ufonds_1	http://ark1.hioa.no/n5.ui/createdDate	"2012-02-21T20:32:39"
genid:Ufonds_1	http://ark1.hioa.no/n5.ui/createdBy	"Oslo municipality"
genid:Ufonds_1	http://ark1.hioa.no/n5.ui/isCreatedBy	genid:Ufondscreator
genid:Ufonds_2	http://www.w3.org/1999/02/22-rdf-syntax-ns#type	http://ark1.hioa.no/n5.ui/fonds
genid:Ufonds_2	http://ark1.hioa.no/n5.ui/systemID	"d7ee5939-4122-4b80-9e1e-057dac4f3d27"
genid:Ufonds_2	http://ark1.hioa.no/n5.ui/title	"fond2"
genid:Ufonds_2	http://ark1.hioa.no/n5.ui/description	"New fond creation"
genid:Ufonds_2	http://ark1.hioa.no/n5.ui/fondsstatus	"closed"
genid:Ufonds_2	http://ark1.hioa.no/n5.ui/createdDate	"2012-02-21T20:32:39"
genid:Ufonds_2	http://ark1.hioa.no/n5.ui/createdBy	"Oslo municipality"
genid:Ufonds_2	http://ark1.hioa.no/n5.ui/isCreatedBy	genid:Ufondscreator
genid:Ufonds_3	http://www.w3.org/1999/02/22-rdf-syntax-ns#type	http://ark1.hioa.no/n5.ui/fonds
genid:Ufonds_3	http://ark1.hioa.no/n5.ui/systemID	"d7ee5939-4122-4b80-9e1e-057dac4f3d28"
genid:Ufonds_3	http://ark1.hioa.no/n5.ui/title	"fond3"
genid:Ufonds_3	http://ark1.hioa.no/n5.ui/description	"New fond3"
genid:Ufonds_3	http://ark1.hioa.no/n5.ui/fondsstatus	"closed"
genid:Ufonds_3	http://ark1.hioa.no/n5.ui/createdDate	"2012-02-21T20:32:39"
genid:Ufonds_3	http://ark1.hioa.no/n5.ui/createdBy	"Oslo municipality"
genid:Ufonds_3	http://ark1.hioa.no/n5.ui/isCreatedBy	genid:Ufondscreator

8.3. Appendix 3 : Archives structure: output 37 triples

Subject	Predicate	Object
genid:Ufonds_1	http://www.w3.org/1999/02/22-rdf-syntax-ns#type	http://ark1.hioa.no/n5.ui/fonds
genid:Ufonds_1	http://ark1.hioa.no/n5.ui/systemID	"4e1b554c-dcc7-4b89-820f-7°8e0c44a672"
genid:Ufonds_1	http://ark1.hioa.no/n5.ui/description	"fond1"
genid:Ufonds_1	http://ark1.hioa.no/n5.ui/hasPart	genid:Usubfonds
genid:Ufonds_1	http://ark1.hioa.no/n5.ui/isRelated	genid:Useries_1
genid:Usubfonds	http://www.w3.org/1999/02/22-rdf-syntax-ns#type	http://ark1.hioa.no/n5.ui/subfonds
genid:Usubfonds	http://ark1.hioa.no/n5.ui/systemID	"0a7a9c5a-f39a-43f4-a531-8fe6344c2188"
genid:Usubfonds	http://ark1.hioa.no/n5.ui/title	"subfonds"
genid:Usubfonds	http://ark1.hioa.no/n5.ui/issubPartOf	genid:Ufonds_1
genid:Useries_1	http://www.w3.org/1999/02/22-rdf-syntax-ns#type	http://ark1.hioa.no/n5.ui/series
genid:Useries_1	http://ark1.hioa.no/n5.ui/systemID	"de01097f-03d1-4032-b787-8062264ecf17"
genid:Useries_1	http://ark1.hioa.no/n5.ui/description	"series1"
genid:Useries_1	http://ark1.hioa.no/n5.ui/isRelated	genid:Ufile_1
genid:Ufile_1	http://www.w3.org/1999/02/22-rdf-syntax-ns#type	http://ark1.hioa.no/n5.ui/files
genid:Ufile_1	http://ark1.hioa.no/n5.ui/systemID	"83029702-519a-4933-aeda-89ba6cfcee2b"
genid:Ufile_1	http://ark1.hioa.no/n5.ui/fileID	"11/012465"
genid:Ufile_1	http://ark1.hioa.no/n5.ui/description	"file1"
genid:Ufile_1	http://ark1.hioa.no/n5.ui/isRelated	genid:Urecord_1
genid:Urecord_1	http://www.w3.org/1999/02/22-rdf-syntax-ns#type	http://ark1.hioa.no/n5.ui/records
genid:Urecord_1	http://ark1.hioa.no/n5.ui/systemID	"403ac99f-1d20-438c-a398-824b32d045ff"
genid:Urecord_1	http://ark1.hioa.no/n5.ui/recordtype	"document"
genid:Urecord_1	http://ark1.hioa.no/n5.ui/recordID	"11/012466"
genid:Urecord_1	http://ark1.hioa.no/n5.ui/description	"record1"
genid:Urecord_1	http://ark1.hioa.no/n5.ui/hasPart	genid:Udocumentdescription
genid:Urecord_1	http://ark1.hioa.no/n5.ui/hasPart	genid:Udocumentobject
genid:Udocumentdescription	http://www.w3.org/1999/02/22-rdf-syntax-ns#type	http://ark1.hioa.no/n5.ui/documentdescription
genid:Udocumentdescription	http://ark1.hioa.no/n5.ui/systemID	"29743769-edf4-46f5-be47-51a0167fbfeb"

genid:Udocumentdescription	http://ark1.hioa.no/n5.ui/documenttype	"electronic record"
genid:Udocumentdescription	http://ark1.hioa.no/n5.ui/documentstatus	"closed"
genid:Udocumentdescription	http://ark1.hioa.no/n5.ui/createdDate	"2011-02-21T20:32:56"
genid:Udocumentdescription	http://ark1.hioa.no/n5.ui/createdBy	"Oslo municipality"
genid:Udocumentobject	http://www.w3.org/1999/02/22-rdf-syntax-ns#type	http://ark1.hioa.no/n5.ui/documentobject
genid:Udocumentobject	http://ark1.hioa.no/n5.ui/versionnumber	"bf1dfa10-243a-4a7c-9904-89ec9325cca0"
genid:Udocumentobject	http://ark1.hioa.no/n5.ui/variantformat	"production format"
genid:Udocumentobject	http://ark1.hioa.no/n5.ui/format	"pdf format"
genid:Udocumentobject	http://ark1.hioa.no/n5.ui/createdDate	"2011-02-21T20:32:56"
genid:Udocumentobject	http://ark1.hioa.no/n5.ui/createdBy	"Oslo municipality"

8.4. Appendix 4: Archives and case parties RDF/XML document

```
1: <rdf:RDF
2:          xmlns:content="http://ark1.hioa.no/n5.ui/content/"
3:          xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
4:          xmlns:n5ui="http://ark1.hioa.n5.ui/"
5:          xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#">
6:
7:                    <n5ui:fonds rdf:nodeID="fonds">
8:                              <n5ui:systemID>d7ee5939-4122-4b80-9e1e-057dac4f3d29</n5ui:systemID>
9:                              <n5ui:title>new fond</n5ui:title>
10:                             <n5ui:description>Oslo manicupality</n5ui:description>
11:                             <n5ui:documentstatus>closed</n5ui:documentstatus>
12:                             <n5ui:documentmedium>electronic format</n5ui:documentmedium>
13:                             <n5ui:createdDate>2012-02-21T20:32:39</n5ui:createdDate>
14:                             <n5ui:createdBy>Oslo manicupality</n5ui:createdBy>
15:                             <n5ui:finalisedDate>2012-02-22T14:38:29</n5ui:finalisedDate>
16:                             <n5ui:finalisedBy>Oslo manicupality</n5ui:finalisedBy>
17:                             <n5ui:isCreatedBy rdf:nodeID="fondscreator"/>
18:                             <n5ui:isRelated rdf:nodeID="series"/>
19:                    </n5ui:fonds>
20:                    <n5ui:series rdf:nodeID="series">
21:                             <n5ui:systemID>b23fd11e-1860-4b84-a449-c77c86357330</n5ui:systemID>
```

```
22:                              <n5ui:title>series</n5ui:title>
23:                              <n5ui:description>series</n5ui:description>
24:                              <n5ui:documentmedium>electronic format</n5ui:documentmedium>
25:                              <n5ui:createdDate>2012-02-21T20:32:56</n5ui:createdDate>
26:                              <n5ui:createdBy>Oslo manicupality</n5ui:createdBy>
27:                              <n5ui:finalisedDate>2012-02-22T14:38:29</n5ui:finalisedDate>
28:                              <n5ui:finalisedBy>Oslo manicupality</n5ui:finalisedBy>
29:                              <n5ui:isRelated rdf:nodeID="file"/>
30:                      </n5ui:series>
31:                      <n5ui:file rdf:nodeID="file">
32:                              <n5ui:systemID>e05862f1-11ac-4b06-9228-a18c0005b205</n5ui:systemID>
33:                              <n5ui:title>file</n5ui:title>
34:                              <n5ui:description>application for KG</n5ui:description>
35:                              <n5ui:fileID>10/629652</n5ui:fileID>
36:                              <n5ui:createdDate>2012-02-21T20:33:12</n5ui:createdDate>
37:                              <n5ui:createdBy>Oslo manicupality</n5ui:createdBy>
38:                              <n5ui:finalisedDate>2012-02-21T20:33:12</n5ui:finalisedDate>
39:                              <n5ui:finalisedBy> Oslo manicupality</n5ui:finalisedBy>
40:                              <n5ui:hasCaseType rdf:nodeID="casefile"/>
41:                              <n5ui:isRelated rdf:nodeID="record"/>
42:                      </n5ui:file>
43:                      <n5ui:record rdf:nodeID="record">
44:                              <n5ui:systemID>59d9cff4-32ff-4e3e-a5e7-ac0b6a6edc71</n5ui:systemID>
45:                              <n5ui:createdDate>2012-02-21T20:33:28</n5ui:createdDate>
46:                              <n5ui:createdBy>oslo manicupality</n5ui:createdBy>
47:                              <n5ui:archivedDate>2012-02-22T14:38:29</n5ui:archivedDate>
48:                              <n5ui:archivedBy>Oslo manicupality</n5ui:archivedBy>
49:                              <n5ui:hasPart rdf:nodeID="documentdescription"/>
50:                      </n5ui:record>
51:                              <n5ui:documentdescription rdf:nodeID="documentdescription">
52:                              <n5ui:systemID>087d50fc-55d1-41e4-9ec6-2913a38e5270</n5ui:systemID>
53:                                      <n5ui:documenttype>application/pdf</n5ui:documenttype>
54:                                      <n5ui:documentstatus>document completed</n5ui:documentstatus>
55:                                      <n5ui:title>document description</n5ui:title>
56:                                      <n5ui:createdDate>2012-02-21T20:34:01</n5ui:createdDate>
57:                                      <n5ui:createdBy>Oslo manicupality</n5ui:createdBy>
58:                                      <n5ui:documentmedium>electronic document</n5ui:documentmedium>
59:                                      <n5ui:hasPart rdf:nodeID="documentobject"/>
60:                              </n5ui:documentdescription>
61:                                      <n5ui:documentobject rdf:nodeID="documentobject">
62:                                              <n5ui:versionnumber>1</n5ui:versionnumber>
63:                                              <n5ui:variantformat>productionformat</n5ui:variantformat>
64:                                              <n5ui:format>pdf</n5ui:format>
```

```
65:                                              <n5ui:createdDate>2012-02-21T20:34:01</n5ui:createdDate>
66:                                              <n5ui:createdBy>Oslo manicupality</n5ui:createdBy>
67:                                          </n5ui:documentobject>
68:                      <n5ui:casefile rdf:nodeID="casefile">
69:                              <n5ui:systemID>403ac99f-1d20-438c-a398-824b32d045ff</n5ui:systemID>
70:                              <n5ui:fileID>12/215463</n5ui:fileID>
71:                              <content:hasPart rdf:nodeID="case_party1"/>
72:                              <content:hasPart rdf:nodeID="case_party2"/>
73:                              <content:hasPart rdf:nodeID="case_party3"/>
74:                              <content:hasPart rdf:nodeID="precedent_case"/>
75:                              <content:hasCaseHandler rdf:nodeID="case_handler"/>
76:                              <n5ui:isRelated rdf:nodeID="caserecord"/>
77:                      </n5ui:casefile>
78:                      <content:caseparty rdf:nodeID="case_party1">
79:                              <content:casepartyID>1052987-5412</content:casepartyID>
80:                              <content:casepartyName>Debora</content:casepartyName>
81:                              <content:casepartyRole>Casecreator</content:casepartyRole>
82:                              <content:postaladdress>Toyen gate42</content:postaladdress>
83:                              <content:postcode>0578</content:postcode>
84:                              <content:postaltown>Oslo</content:postaltown>
85:                              <content:foreignaddress>Norway</content:foreignaddress>
86:                              <content:emailaddress>Deb@gmail.com</content:emailaddress>
87:                              <content:telephonenumber>47599874</content:telephonenumber>
88:                              <n5ui:isPartOf rdf:nodeID="casefile"/>
89:                      </content:caseparty>
90:                      <content:caseparty rdf:nodeID="case_party2">
91:                              <content:casepartyID>34r59874-6589</content:casepartyID>
92:                              <content:casepartyName>Helen</content:casepartyName>
93:                              <content:casepartyRole>Casecreator</content:casepartyRole>
94:                              <content:postaladdress>Tofet gate 21</content:postaladdress>
95:                              <content:postcode></content:postcode>
96:                              <content:postaltown>Oslo</content:postaltown>
97:                              <content:foreignaddress>Norway</content:foreignaddress>
98:                              <content:emailaddress>helu@gmail.com</content:emailaddress>
99:                              <content:telephonenumber>47125474</content:telephonenumber>
100:                              <n5ui:isPartOf rdf:nodeID="casefile"/>
101:                      </content:caseparty>
102:                      <content:caseparty rdf:nodeID="case_party3">
103:                              <content:casepartyID>8714562-8745</content:casepartyID>
104:                              <content:casepartyName>private company</content:casepartyName>
105:                              <n5ui:isPartOf rdf:nodeID="casefile"/>
106:                      </content:caseparty>
107:                      <content:case rdf:nodeID="precedent_case">
```

```
108:                                        <content:systemID>987</content:systemID>
109:                                        <content:fileID>10/213652</content:fileID>
110:                                        <content:casedate>2011-10-11T16:05:28</content:casedate>
111:                                        <content:casestatus>under process</content:casestatus>
112:                                        <content:case-responsible>private organization</content:case-responsible>
113:                                        <content:administrativeUnit>oslo manucipality</content:administrativeUnit>
114:                                        <content:precedentstatus>current</content:precedentstatus>
115:                                        <n5ui:isPartOf rdf:nodeID="casefile"/>
116:                                        <n5ui:isPartOf rdf:nodeID="registry_entry1"/>
117:                        </content:case>
118:                <content:casehandler rdf:nodeID="case_handler">
119:                                <content:firstname>Seliman</content:firstname>
120:                                <content:lastname>Enger</content:lastname>
121:                                <content:personRole>executive_officer</content:personRole>
122:                                <n5ui:case-responsible>oslo manicupality</n5ui:case-responsible>
123:                                <n5ui:isResponsibleFor rdf:nodeID="casefile"/>
124:                                <n5ui:isResponsibleFor rdf:nodeID="registry_entry1"/>
125:                        </content:casehandler>
126:                <n5ui:caserecord rdf:nodeID="caserecord">
127:                                <n5ui:isRelated rdf:nodeID="registry_entry1"/>
128:                                <n5ui:isRelated rdf:nodeID="registry_entry2"/>
129:                        </n5ui:caserecord>
130:                <n5ui:registry_entry rdf:nodeID="registry_entry1">
131:                                <content:hasCase rdf:nodeID="precedent_case"/>
132:                                <content:hasCaseHandler rdf:nodeID="case_handler"/>
133:                        </n5ui:registry_entry>
134: </rdf:RDF>
```

Appendix 6

8.6. Appendix 6 : The Noark 5 files and records RDF/XML serialization

```
1: <rdf:RDF
2:          xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
3:          xmlns:n5ui="http://ark1.hioa.no/n5.ui/"
4:          xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#" >
5:                    <rdf:Description rdf:nodeID="fonds">
6:                              <n5ui:systemID>68f08d11-6796-409a-96da-1b403b406549</n5ui:systemID>
7:                              <n5ui:title>new fond</n5ui:title>
8:                              <n5ui:haspart rdf:nodeID="subfonds"/>
9:                              <n5ui:createdBy rdf:nodeID="fondscreator"/>
10:                             <n5ui:isRelated rdf:nodeID="series"/>
11:                  </rdf:Description>
12:                  <rdf:Description rdf:nodeID="fondscreator">
13:                             <n5ui:fondscreatorID>7261</n5ui:fondscreatorID>
14:                             <n5ui:fondscreatorName>Oslo municipality</n5ui:fondscreatorName>
15:                             <n5ui:description> new document creation</n5ui:description>
16:                             <n5ui:isCreated rdf:nodeID="fond"/>
17:                  </rdf:Description>
18:                  <n5ui:subfonds rdf:nodeID="subfonds">
19:                             <n5ui:systemID>abe90af6-5858-4472-8e8f-a4e9c0099628</n5ui:systemID>
20:                             <n5ui:title> sub fonds</n5ui:title>
21:                             <n5ui:issubPartOf rdf:nodeID="fonds"/>
22:                             <n5ui:isRelated rdf:nodeID="series1"/>
23:                  </n5ui:subfonds>
24:                  <n5ui:series rdf:nodeID="series">
25:                             <n5ui:systemID>70efc42c-d1af-4326-9dfd-4c823be0a6e1</n5ui:systemID>
26:                             <n5ui:title>series</n5ui:title>
27:                             <n5ui:isPartOf rdf:nodeID="fonds"/>
28:                             <n5ui:isRelated rdf:nodeID="files1"/>
29:                             <n5ui:isRelated rdf:nodeID="files2"/>
30:                             <n5ui:isRelated rdf:nodeID="files3"/>
31:                  </n5ui:series>
32:                             <n5ui:files rdf:nodeID="files1">
33:                             <n5ui:systemID>adb72850-837e-4f97-91f3-4374ecfd7491</n5ui:systemID>
34:                               <n5ui:fileID>12/459685</n5ui:fileID>
35:                                              <n5ui:isPartOf rdf:nodeID="series"/>
36:                                              <n5ui:isRelated rdf:nodeID="records_1"/>
37:                                              <n5ui:isRelated rdf:nodeID="records_2"/>
38:                                              <n5ui:isRelated rdf:nodeID="records_3"/>
```

127

39:                                  </n5ui:files>
40:                                      <n5ui:records rdf:nodeID="records_1">
41:                          <n5ui:systemID>24317cc589</n5ui:systemID>
42:                          <n5ui:createdDate>2011-10-11T17:34:53</n5ui:createdDate>
43:   <n5ui:isPartOf rdf:nodeID="files1"/>
44:   <n5ui:hasPart rdf:nodeID="documentdescription_1"/>
45:                                  </n5ui:records>
46:                            <n5ui:documentdescription rdf:nodeID="documentdescription_1">
47:                              <n5ui:systemID>4907c78b-fc25-40a8-a957-efca1f3359fb</n5ui:systemID>
48:                          <n5ui:documenttype>application/pdf</n5ui:documenttype>
49:                                      <n5ui:hasPart rdf:nodeID="documentobject_1"/>
50:   </n5ui:documentdescription>
51:   <n5ui:documentobject rdf:nodeID="documentobject_1">
52:                          <n5ui:versionnumber>1</n5ui:versionnumber>
53:                          <n5ui:isPartOf rdf:nodeID="documentdescription_1"/>
54:                          </n5ui:documentobject>
55:                  <n5ui:records rdf:nodeID="records_2">
56:                          <n5ui:systemID>7bfeea2d-acf9-470b-a591-08ea30b85b0f</n5ui:systemID>
57:                          <n5ui:createdDate>2011-10-11T17:34:22</n5ui:createdDate>
58:                                      <n5ui:isPartOf rdf:nodeID="files1"/>
59:                                      <n5ui:hasPart rdf:nodeID="documentdescription_2"/>
60:                                  </n5ui:records>
61:                            <n5ui:documentdescription rdf:nodeID="documentdescription_2">
62:                              <n5ui:systemID>7°129a6f-a990-4424-89d2-8945e5859ff0</n5ui:systemID>
63:                          <n5ui:documenttype>application/pdf</n5ui:documenttype>
64:                          <n5ui:hasPart rdf:nodeID="documentobject_2"/>
65:                                      </n5ui:documentdescription>
66:                                      <n5ui:documentobject rdf:nodeID="documentobject_2">
67:                          <n5ui:versionnumber>1</n5ui:versionnumber>
68:                          <n5ui:variantformat>new format</n5ui:variantformat>
69:                          <n5ui:issubPartOf rdf:nodeID="documentdescription_2"/>
70:                          </n5ui:documentobject>
71:                          <n5ui:records rdf:nodeID="records_3">
72:                          <n5ui:systemID>0282d31d-6a0f-4fb5-b027-b89d2f34b121</n5ui:systemID>
73:                          <n5ui:isPartOf rdf:nodeID="files1"/>
74:                          <n5ui:hasPart rdf:nodeID="documentdescription_3"/>
75:                  </n5ui:records>
76:                            <n5ui:documentdescription rdf:nodeID="documentdescription_3">
77:                          <n5ui:systemID>dc3381e4-9f1d-493f-969o-a66903a0dea8</n5ui:systemID>
78:                                      <n5ui:hasPart rdf:nodeID="documentobject_3"/>
79:   </n5ui:documentdescription>
80:   <n5ui:documentobject rdf:nodeID="documentobject_3">
81:   <n5ui:isPartOf rdf:nodeID="documentdescription_3"/>

128

```xml
82:                          </n5ui:documentobject>
83:                       <n5ui:files rdf:nodeID="files2">
84:                  <n5ui:systemID>2304a4f0-b765-4be7-a620-a795866bd084</n5ui:systemID>
85:                    <n5ui:fileID>12/547896</n5ui:fileID>
86:                                  <n5ui:isPartOf rdf:nodeID="series"/>
87:                                  <n5ui:isRelated rdf:nodeID="records1"/>
88:                                  <n5ui:isRelated rdf:nodeID="records2"/>
89:                                  <n5ui:isRelated rdf:nodeID="records3"/>
90:                                  <n5ui:isPartOf rdf:nodeID="series"/>
91:                  </n5ui:files>
92:                          <n5ui:records rdf:nodeID="records1">
93:                  <n5ui:systemID>3be6bd40-cc8a-4694-8f91-b0c61827cf14</n5ui:systemID>
94:                                  <n5ui:isPartOf rdf:nodeID="files2"/>
95:                                  <n5ui:hasPart rdf:nodeID="documentdescription1"/>
96:                          </n5ui:records>
97:                    <n5ui:documentdescription rdf:nodeID="documentdescription1">
98:                  <n5ui:systemID>009e95e9-d12e-49f5-a3c1-bde05fba1f4f</n5ui:systemID>
99:                                      <n5ui:linkedBy>system</n5ui:linkedBy>
100:                                     <n5ui:issubPartOf rdf:nodeID="records"/>
101:                                     <n5ui:hasPart rdf:nodeID="documentobject1"/>
102:                            </n5ui:documentdescription>
103:                        <n5ui:documentobject rdf:nodeID="documentobject1">
104:                  <n5ui:versionnumber>1</n5ui:versionnumber>
105:                                  <n5ui:isPartOf rdf:nodeID="documentdescription1"/>
106:                                      </n5ui:documentobject>
107:                       <n5ui:records rdf:nodeID="records2">
108:                  <n5ui:systemID>f4bc0d89-48b6-4c52-a149-10074f24e01c</n5ui:systemID>
109:                  <n5ui:archivedBy>Oslo municipality</n5ui:archivedBy>
110:                                  <n5ui:isPartOf rdf:nodeID="files2"/>
111:                                  <n5ui:hasPart rdf:nodeID="documentdescription2"/>
112:                          </n5ui:records>
113:                     <n5ui:documentdescription rdf:nodeID="documentdescription2">
114:                  <n5ui:systemID>379°6e0f-e2e3-49f3-a2bd-e0a913788d88</n5ui:systemID>
115:                                      <n5ui:linkedBy>system</n5ui:linkedBy>
116:                                     <n5ui:hasPart rdf:nodeID="documentobject2"/>
117:                            </n5ui:documentdescription>
118:                    <n5ui:documentobject rdf:nodeID="documentobject2">
119:                                  <n5ui:issubPartOf rdf:nodeID="documentdescription2"/>
120:                                       </n5ui:documentobject>
121:                       <n5ui:records rdf:nodeID="records3">
122:                  <n5ui:systemID>87c4c78c-8370-4649-9e16-3b0bb5866563</n5ui:systemID>
123:                                  <n5ui:isPartOf rdf:nodeID="file2"/>
124:                                  <n5ui:hasPart rdf:nodeID="documentdescription3"/>
```

129

```
125:                                      </n5ui:records>
126:                                       <n5ui:documentdescription rdf:nodeID="documentdescription3">
127:                                  <n5ui:systemID>432978ea-eded-4e20-89ae-c129a8baee20</n5ui:systemID>
128:                                            <n5ui:linkedBy>system</n5ui:linkedBy>
129:                                             <n5ui:isPartOf rdf:nodeID="record3"/>
130:                                             <n5ui:hasPart rdf:nodeID="documentobject3"/>
131:                                        </n5ui:documentdescription>
132:                                          <n5ui:documentobject rdf:nodeID="documentobject3">
133:                            <n5ui:versionnumber>1</n5ui:versionnumber>
134:                            <n5ui:filesize>1232</n5ui:filesize>
135:                                            <n5ui:isPartOf rdf:nodeID="documentdescription3"/>
136:                                               </n5ui:documentobject>
137:              <n5ui:files rdf:nodeID="files3">
138:              <n5ui:systemID>baffe63d-354a-405f-9b0d-135bee4a113f</n5ui:systemID>
139:                           <n5ui:fileID>12/965836</n5ui:fileID>
140:                                     <n5ui:isPartOf rdf:nodeID="series"/>
141:                                     <n5ui:isRelated rdf:nodeID="records3of1"/>
142:                                     <n5ui:isRelated rdf:nodeID="records3of2"/>
143:                                     <n5ui:isRelated rdf:nodeID="records3of3"/>
144:              </n5ui:files>
145:                                   <n5ui:records rdf:nodeID="records3of1">
146:                      <n5ui:systemID>c0411f76-7d44-4fdc-bba9-0d6aa61f7557</n5ui:systemID>
147:                                          <n5ui:isPartOf rdf:nodeID="files3"/>
148:                                          <n5ui:hasPart rdf:nodeID="documentdescription3of1"/>
149:                                      </n5ui:records>
150:                                       <n5ui:documentdescription rdf:nodeID="documentdescription3of1">
151:                      <n5ui:systemID>d741dd93-9dab-4e6c-a874-a9d1654f1033</n5ui:systemID>
152:                                          <n5ui:linkedBy>system</n5ui:linkedBy>
153:                                               <n5ui:issubPartOf rdf:nodeID="records3of1"/>
154:                                          <n5ui:hasPart rdf:nodeID="documentobject3of1"/>
155:                                          </n5ui:documentdescription>
156:                                          <n5ui:documentobject rdf:nodeID="documentobject3of1">
157:                            <n5ui:versionnumber>1</n5ui:versionnumber>
158:                                          <n5ui:isPartOf rdf:nodeID="documentdescription3of1"/>
159:                                              </n5ui:documentobject>
160:                           <n5ui:records rdf:nodeID="records3of2">
161:                      <n5ui:systemID>b0c1d434-6afb-4cd9-8924-4f721cf97303</n5ui:systemID>
162:                      <n5ui:archivedBy>Oslo municipality</n5ui:archivedBy>
163:                                          <n5ui:isPartOf rdf:nodeID="files3"/>
164:                                          <n5ui:hasPart rdf:nodeID="documentdescription3of2"/>
165:                                      </n5ui:records>
166:                                       <n5ui:documentdescription rdf:nodeID="documentdescription3of2">
167:                                  <n5ui:systemID>503dfa9c-da1e-4146-ab26-b79f4bebf113</n5ui:systemID>
```

168:   <n5ui:linkedBy>system</n5ui:linkedBy>
169:   <n5ui:hasPart rdf:nodeID="documentobject3of2"/>
170:   </n5ui:documentdescription>
171:   <n5ui:documentobject rdf:nodeID="documentobject3of2">
172:   <n5ui:issubPartOf rdf:nodeID="documentdescription3of2"/>
173:   </n5ui:documentobject>
174:   <n5ui:records rdf:nodeID="records3of3">
175:                                   <n5ui:systemID>2c7ea559-03c6-45e1-b930-05b457dc42e5</n5ui:systemID>
176:   <n5ui:isPartOf rdf:nodeID="file3"/>
177:   <n5ui:hasPart rdf:nodeID="documentdescription3of3"/>
178:   </n5ui:records>
179:   <n5ui:documentdescription rdf:nodeID="documentdescription3of3">
180:                                   <n5ui:systemID>c4b9586d-672b-4781-b4b2-988717f19e57</n5ui:systemID>
181:   <n5ui:linkedBy>system</n5ui:linkedBy>
182:   <n5ui:isPartOf rdf:nodeID="records3"/>
183:   <n5ui:hasPart rdf:nodeID="documentobject3of3"/>
184:   </n5ui:documentdescription>
185:   <n5ui:documentobject rdf:nodeID="documentobject3of3">
186:                              <n5ui:versionnumber>1</n5ui:versionnumber>
187:                              <n5ui:filesize>1269</n5ui:filesize>
188:   <n5ui:isPartOf rdf:nodeID="documentdescription3of3"/>
189:   </n5ui:documentobject>
190: </rdf:RDF>

8.8. Appendix 8: The Noark5 files, case files, case records, and records RDF/XML serialization

```
1: <rdf:RDF

2:          xmlns:n5ui="http://ark1.hioa.no/n5.ui/"

3:          xmlns:content="http://ark1.hioa.no/n5.ui/content/"

4:          xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"

5:          xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#" >

6:

7:                    <n5ui:fonds rdf:nodeID="fonds">

8:                              <n5ui:systemID>3be6bd40-cc8a-4694-8f91-b0c61827cf14</n5ui:systemID>

9:                              <n5ui:isRelated rdf:nodeID="series"/>

10:                   </n5ui:fonds>

11:                   <n5ui:series rdf:nodeID="series">

12:                              <n5ui:systemID>f97c8684-7da1-4d5f-96ac-0fdbcf233fd1</n5ui:systemID>

13:                              <n5ui:isRelated rdf:nodeID="files"/>

14:                   </n5ui:series>

15:                   <n5ui:files  rdf:nodeID="files">

16:                              <content:hasCaseFile rdf:nodeID="casefile"/>

17:                              <n5ui:isRelated rdf:nodeID="records"/>

18:                   </n5ui:files>

19:                   <n5ui:records rdf:nodeID="records">

20:                              <n5ui:systemID>009e95e9-d12e-49f5-a3c1-bde05fba1f4f</n5ui:systemID>

21:                   </n5ui:records>

22:                   <n5ui:casefile rdf:nodeID="casefile">

23:                              <n5ui:systemID>f4bc0d89-48b6-4c52-a149-10074f24e01c</n5ui:systemID>

24:                              <n5ui:casefileID>12/875496</n5ui:casefileID>

25:                              <n5ui:hasCaseParty rdf:nodeID="caseparty"/>

26:                              <n5ui:hasPrecedent  rdf:nodeID="precedent"/>

27:                              <n5ui:isRelated rdf:nodeID="caserecord"/>

28:                   </n5ui:casefile>

29:                   <n5ui:caserecord rdf:nodeID="caserecord">

30:                              <n5ui:systemID>379a6e0f-e2e3-49f3-a2bd-e0a913788d88</n5ui:systemID>

31:                              <n5ui:recordID>12/547821-5482</n5ui:recordID>

32:                              <n5ui:isRelated rdf:nodeID="registry_entry1"/>

33:                              <n5ui:isRelated rdf:nodeID="registry_entry2"/>

34:                   </n5ui:caserecord>

35:                   <n5ui:registry_entry rdf:nodeID="registry_entry1">

36:                              <n5ui:systemID>87c4c78c-8370-4649-9e16-3b0bb5866563</n5ui:systemID>

37:                              <n5ui:recordID>12/238734-9832</n5ui:recordID>

38:                              <n5ui:serialnumber>5698632/12</n5ui:serialnumber>

39:                              <n5ui:registryentrytype>received</n5ui:registryentrytype>
```

40:                              <n5ui:description>incoming document recieved</n5ui:description>
41:                              <n5ui:status>registered</n5ui:status>
42:                              <n5ui:hasClient rdf:nodeID="client1"/>
43:                              <n5ui:hasCaseResponsible rdf:nodeID="Soliana"/>
44:                      </n5ui:registry_entry>
45:                      <n5ui:registry_entry rdf:nodeID="registry_entry2">
46:                              <n5ui:systemID>432978ea-eded-4e20-89ae-c129a8baee20</n5ui:systemID>
47:                              <n5ui:recordID>12/874598-2478</n5ui:recordID>
48:                              <n5ui:serialnumber>7412583/12</n5ui:serialnumber>
49:                              <n5ui:registryentrytype>approved</n5ui:registryentrytype>
50:                              <n5ui:status>dispatched</n5ui:status>
51:                              <n5ui:hasClient rdf:nodeID="client2"/>
52:                              <n5ui:hasCaseResponsible rdf:nodeID="Soliana"/>
53:                      </n5ui:registry_entry>
54:                      <n5ui:case_responsibility rdf:nodeID="case_responsibility">
55:                              <n5ui:administrativeUnit>Oslo municipality </n5ui:administrativeUnit>
56:                              <n5ui:executiveofficer>Soliana</n5ui:executiveofficer>
57:                              <n5ui:isResponsibleFor rdf:nodeID="registry_entry1"/>
58:                              <n5ui:isResponsibleFor rdf:nodeID="registry_entry2"/>
59:                      </n5ui:case_responsibility>
60:                      <n5ui:client rdf:nodeID="client1">
61:                              <n5ui:clienttype>private</n5ui:clienttype>
62:                              <n5ui:clientName>Helen</n5ui:clientName>
63:                              <n5ui:description>client</n5ui:description>
64:                              <n5ui:isPartOf rdf:nodeID="registry_entry1"/>
65:                      </n5ui:client>
66:                      <n5ui:client rdf:nodeID="client2">
67:                              <n5ui:clienttype>new user</n5ui:clienttype>
68:                              <n5ui:clientName>Domas</n5ui:clientName>
69:                              <n5ui:description>client</n5ui:description>
70:                              <n5ui:isPartOf rdf:nodeID="registry_entry2"/>
71:                      </n5ui:client>
72: </rdf:RDF>

8.9. Appendix 9 : The Noark5 records with document object RDF/XML serialization

```
1: <rdf:RDF
2:           xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
3:           xmlns:n5ui="http://ark1.hioa.no/n5.ui/"
4:           xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#">
5:                 <rdf:Description rdf:nodeID="fonds">
6:                         <n5ui:systemID>d741dd93-9dab-4e6c-a874-a9d1654f1033</n5ui:systemID>
7:                         <n5ui:title>new fonds</n5ui:title>
8:                         <n5ui:description>fonds created</n5ui:description>
9:                         <n5ui:fondsstatus>closed</n5ui:fondsstatus>
10:                        <n5ui:documentmedium>electronic format</n5ui:documentmedium>
11:                        <n5ui:createdDate>2011-10-11T17:35:44</n5ui:createdDate>
12:                        <n5ui:createdBy>Oslo municipality</n5ui:createdBy>
13:                        <n5ui:finalisedDate>2011-10-21</n5ui:finalisedDate>
14:                        <n5ui:finalisedBy>Oslo municipality</n5ui:finalisedBy>
15:                        <n5ui:isCreatedBy rdf:nodeID="fondscreator"/>
16:                        <n5ui:isRelated rdf:nodeID="series"/>
17:                        </rdf:Description>
18:                 <rdf:Description rdf:nodeID="fondscreator">
19:                        <n5ui:fondscreatorID>b0c1d434-6afb-4cd9-8924-4f721cf97303</n5ui:fondscreatorID>
20:                        <n5ui:fondscreatorName>Oslo municipality</n5ui:fondscreatorName>
21:                        <n5ui:description> new document creation</n5ui:description>
22:                        <n5ui:isCreated rdf:nodeID="fonds"/>
23:                 </rdf:Description>
24:                 <n5ui:series rdf:nodeID="series">
25:                        <n5ui:systemID>503dfa9c-da1e-4146-ab26-b79f4bebf113</n5ui:systemID>
26:                        <n5ui:title>series</n5ui:title>
27:                        <n5ui:description>series</n5ui:description>
28:                        <n5ui:recordsectionstatus>closed</n5ui:recordsectionstatus>
29:                        <n5ui:documentmedium>electronic document</n5ui:documentmedium>
30:                        <n5ui:createdDate>2011-10-11T17:51:58</n5ui:createdDate>
31:                        <n5ui:createdBy>Oslo municpality</n5ui:createdBy>
32:                        <n5ui:finalisedDate>2011-10-11T19:26:38</n5ui:finalisedDate>
33:                        <n5ui:finalisedBy>Oslo municipality</n5ui:finalisedBy>
34:                        <n5ui:isPartOf rdf:nodeID="fonds"/>
35:                        <n5ui:isRelated rdf:nodeID="files1"/>
36:                        <n5ui:isRelated rdf:nodeID="files2"/>
37:                        <n5ui:isRelated rdf:nodeID="files3"/>
38:                 </n5ui:series>
39:                 <n5ui:files rdf:nodeID="files1">
40:                 <n5ui:systemID>2c7ea559-03c6-45e1-b930-05b457dc42e5</n5ui:systemID>
41:                        <n5ui:fileID>10/854236</n5ui:fileID>
```
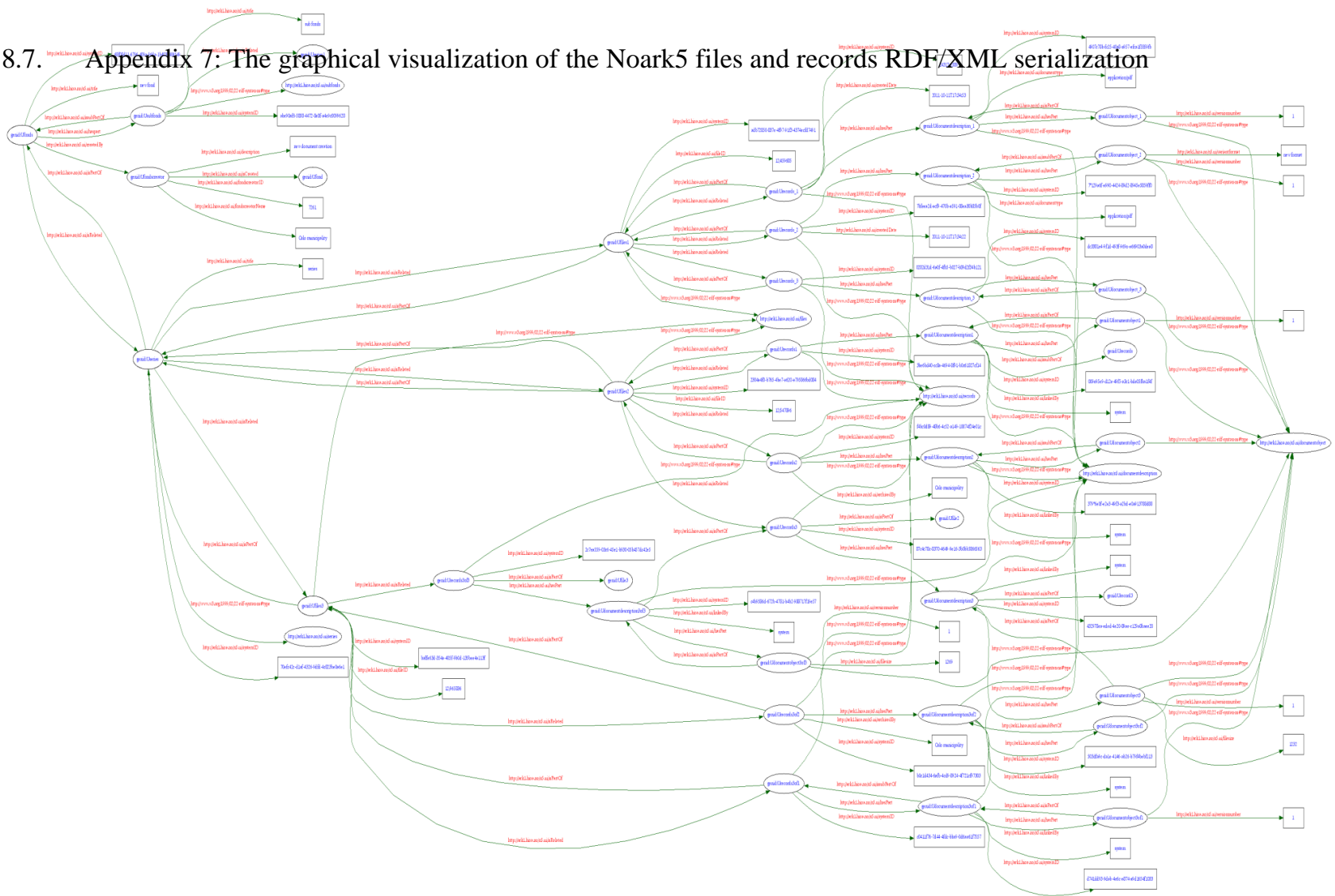
135

```
42:                              <n5ui:title>Files1</n5ui:title>
43:                     <n5ui:createdDate>2011-10-11T17:51:49</n5ui:createdDate>
44:                     <n5ui:createdBy>Oslo municpality</n5ui:createdBy>
45:                     <n5ui:finalisedDate>2011-10-11T19:26:38</n5ui:finalisedDate>
46:                     <n5ui:finalisedBy>Oslo municpality</n5ui:finalisedBy>
47:                                     <n5ui:isPartOf rdf:nodeID="series"/>
48:                                     <n5ui:isRelated rdf:nodeID="records_1"/>
49:                                     <n5ui:isRelated rdf:nodeID="records_2"/>
50:                                     <n5ui:isRelated rdf:nodeID="records_3"/>
51:             </n5ui:files>
52:                          <n5ui:records rdf:nodeID="records_1">
53:                     <n5ui:systemID>517d06b7-ed83-482f-8f55-dc934a4957a1</n5ui:systemID>
54:                     <n5ui:createdDate>2010-10-11</n5ui:createdDate>
55:                     <n5ui:createdBy>Oslo municipality</n5ui:createdBy>
56:                     <n5ui:archivedDate>2011-10-11T19:26:38</n5ui:archivedDate>
57:                     <n5ui:archivedBy>Oslo municpality</n5ui:archivedBy>
58:                                     <n5ui:isPartOf rdf:nodeID="files1"/>
59:                                     <n5ui:hasPart rdf:nodeID="documentobject_1"/>
60:                     </n5ui:records>
61:                                     <n5ui:documentobject rdf:nodeID="documentobject_1">
62:                     <n5ui:versionnumber>1</n5ui:versionnumber>
63:                                             <n5ui:variantformat>Production format</n5ui:variantformat>
64:         <n5ui:format>application/pdf</n5ui:format>
65:                     <n5ui:createdDate>2011-10-11T17:36:27</n5ui:createdDate>
66:                     <n5ui:createdBy>Oslo municipality</n5ui:createdBy>
67:                     <n5ui:referencedocumentfile>f4b45bc.pdf</n5ui:referencedocumentfile>
68:                     <n5ui:checksum>107e83aa</n5ui:checksum>
69:                     <n5ui:checksumAlgorithm>SHA256</n5ui:checksumAlgorithm>
70:                     <n5ui:filesize>123622</n5ui:filesize>
71:                                             <n5ui:isPartOf rdf:nodeID="records_1"/>
72:                                         </n5ui:documentobject>
73:         <n5ui:records rdf:nodeID="records_2">
74:                     <n5ui:systemID>b4bc49be-af51-412f-811f-8c19213806df</n5ui:systemID>
75:                     <n5ui:createdDate>2011-10-11T17:52:31</n5ui:createdDate>
76:                     <n5ui:createdBy>Oslo municipality</n5ui:createdBy>
77:                     <n5ui:archivedDate>2011-10-11T19:26:38</n5ui:archivedDate>
78:                     <n5ui:archivedBy>Oslo municpality</n5ui:archivedBy>
79:                                     <n5ui:isPartOf rdf:nodeID="files1"/>
80:                                     <n5ui:hasPart rdf:nodeID="documentobject_2"/>
81:                     </n5ui:records>
82:                                     <n5ui:documentobject rdf:nodeID="documentobject_2">
83:                     <n5ui:versionnumber>1</n5ui:versionnumber>
84:                                             <n5ui:variantformat>Production format</n5ui:variantformat>
```

```
85:                              <n5ui:format>pdf</n5ui:format>
86:                              <n5ui:createdDate>2011-10-11T17:52:42</n5ui:createdDate>
87:                              <n5ui:createdBy>Oslo municipality</n5ui:createdBy>
88:                              <n5ui:referencedocumentfile>f4b458275bc.pdf</n5ui:referencedocumentfile>
89:                              <n5ui:checksum>104786</n5ui:checksum>
90:                              <n5ui:checksumAlgorithm>SHA256</n5ui:checksumAlgorithm>
91:                              <n5ui:filesize>12783622</n5ui:filesize>
92:                                                        <n5ui:isPartOf rdf:nodeID="records_2"/>
93:                                      </n5ui:documentobject>
94:                         <n5ui:records rdf:nodeID="records_3">
95:                 <n5ui:systemID>6ad7da06-e311-403e-8fd7-8a2c9e09f251</n5ui:systemID>
96:                 <n5ui:createdDate>2011-10-11T17:36:21</n5ui:createdDate>
97:                 <n5ui:createdBy>Oslo municpality</n5ui:createdBy>
98:                 <n5ui:archivedDate>2011-10-11T19:26:38</n5ui:archivedDate>
99:                 <n5ui:archivedBy>Oslo municpality</n5ui:archivedBy>
100:                                      <n5ui:isPartOf rdf:nodeID="files1"/>
101:                                 <n5ui:hasPart rdf:nodeID="documentobject_3"/>
102:                         </n5ui:records>
103:                                  <n5ui:documentobject rdf:nodeID="documentobject_3">
104:                         <n5ui:versionnumber>3</n5ui:versionnumber>
105:                                             <n5ui:variantformat>Production format</n5ui:variantformat>
106:                 <n5ui:format>pdf</n5ui:format>
107:                          <n5ui:createdDate>2011-10-11T17:36:16</n5ui:createdDate>
108:                          <n5ui:createdBy>Oslo municpality</n5ui:createdBy>
109:                          <n5ui:referencedocumentfile>792n.pdf</n5ui:referencedocumentfile>
110:                 <n5ui:checksum>107e8nm3aa</n5ui:checksum>
111:        <n5ui:checksumAlgorithm>SHA256</n5ui:checksumAlgorithm>
112:                 <n5ui:filesize>123645</n5ui:filesize>
113:                                             <n5ui:isPartOf rdf:nodeID="records_3"/>
114:                                      </n5ui:documentobject>
115: </rdf:RDF>
```

8.10. Appendix 10: RDF/XML serialization of the Noark5fonds, subfonds (Files and case files, records, case record, and registry entry).

```
1: <rdf:RDF
2:         xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
3:         xmlns:n5ui="http://ark1.hioa.no/n5.ui/"
4:         xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#">
5:                 <rdf:Description rdf:nodeID="fonds">
```

```
6:                              <n5ui:systemID>74c828ac-7834-49c5-a868-687b324b9460</n5ui:systemID>
7:                              <n5ui:title>new fonds</n5ui:title>
8:                              <n5ui:description>new fonds document</n5ui:description>
9:                              <n5ui:fondsstatus>closed</n5ui:fondsstatus>
10:                             <n5ui:documentmedium>electronic format</n5ui:documentmedium>
11:                             <n5ui:createdDate>2011-10-11T17:33:35</n5ui:createdDate>
12:                             <n5ui:createdBy>Oslo municipality</n5ui:createdBy>
13:                             <n5ui:finalisedDate>2011-10-11T19:26:39</n5ui:finalisedDate>
14:                             <n5ui:finalisedBy>Oslo municipality</n5ui:finalisedBy>
15:                             <n5ui:hasPart rdf:nodeID="subfonds"/>
16:                             <n5ui:createdBy rdf:nodeID="fondscreator"/>
17:                             <n5ui:isRelated rdf:nodeID="series"/>
18:                             </rdf:Description>
19:             <rdf:Description rdf:nodeID="fondscreator">
20:                             <n5ui:fondscreatorID>7261254</n5ui:fondscreatorID>
21:                             <n5ui:fondscreatorName>Oslo municipality</n5ui:fondscreatorName>
22:                             <n5ui:description> new document creation</n5ui:description>
23:                             <n5ui:isCreated rdf:nodeID="fonds"/>
24:             </rdf:Description>
25:             <n5ui:subfonds rdf:nodeID="subfonds">
26:                             <n5ui:systemID>9e2cf47e-d06d-43a5-812f-5ac3c5350f51</n5ui:systemID>
27:                             <n5ui:title>subfonds</n5ui:title>
28:                             <n5ui:description>subfonds</n5ui:description>
29:                             <n5ui:isPartOf rdf:nodeID="fonds"/>
30:                             <n5ui:isRelated rdf:nodeID="series1"/>
31:             </n5ui:subfonds>
32:             <n5ui:series rdf:nodeID="series">
33:                             <n5ui:systemID>f068c9c6-6d1e-4229-874e-634280993340</n5ui:systemID>
34:                             <n5ui:title>series</n5ui:title>
35:                             <n5ui:description>series</n5ui:description>
36:                             <n5ui:recordsectionstatus>completed</n5ui:recordsectionstatus>
37:                             <n5ui:documentmedium>electronic document</n5ui:documentmedium>
38:                             <n5ui:createdDate>2011-10-11T17:35:30</n5ui:createdDate>
39:                             <n5ui:createdBy>Oslo municipality</n5ui:createdBy>
40:                             <n5ui:finalisedDate>2011-10-11T19:26:39</n5ui:finalisedDate>
41:                             <n5ui:finalisedBy>Oslo municpality</n5ui:finalisedBy>
42:                             <n5ui:isRelated rdf:nodeID="files1"/>
43:                             <n5ui:isRelated rdf:nodeID="files2"/>
44:                             <n5ui:isRelated rdf:nodeID="files3"/>
45:             </n5ui:series>
46:             <n5ui:files rdf:nodeID="files1">
47:                             <n5ui:systemID>2fa53b7f-8f66-4372-b501-7b52cc881d5b</n5ui:systemID>
48:                             <n5ui:fileID>11/215684</n5ui:fileID>
```

49:                              <n5ui:title>Files1</n5ui:title>
50:                                  <n5ui:createdDate>2011-10-11T17:50:41</n5ui:createdDate>
51:                                  <n5ui:createdBy>Oslo municipality</n5ui:createdBy>
52:                                  <n5ui:finalisedDate>2011-10-11T19:26:39</n5ui:finalisedDate>
53:                                  <n5ui:finalisedBy>Oslo municipality</n5ui:finalisedBy>
54:   <n5ui:isPartOf rdf:nodeID="series"/>
55:   <n5ui:isRelated rdf:nodeID="casefile_1"/>
56:   <n5ui:isRelated rdf:nodeID="casefile_2"/>
57:   <n5ui:isRelated rdf:nodeID="appointment_case"/>
58:   <n5ui:isRelated rdf:nodeID="meeting_case"/>
59:   <n5ui:isRelated rdf:nodeID="division_case"/>
60:   <n5ui:isRelated rdf:nodeID="records1"/>
61:   <n5ui:isRelated rdf:nodeID="records2"/>
62:   <n5ui:isRelated rdf:nodeID="records3"/>
63:                      </n5ui:files>
64:                      <n5ui:files rdf:nodeID="files2">
65:                                      <n5ui:systemID>105f898b-9af1-4957-978c-92bc8f848243</n5ui:systemID>
66:                      <n5ui:fileID>11/587421</n5ui:fileID>
67:                      <n5ui:title>File2</n5ui:title>
68:                      <n5ui:createdDate>2011-10-11T17:50:28</n5ui:createdDate>
69:                      <n5ui:createdBy>Oslo municipality</n5ui:createdBy>
70:                      <n5ui:finalisedDate>2011-10-11T19:26:36</n5ui:finalisedDate>
71:                      <n5ui:finalisedBy>Oslo municipality</n5ui:finalisedBy>
72:                                      <n5ui:hasPart rdf:nodeID="subfiles"/>
73:                      </n5ui:files>
74:                          <n5ui:casefile rdf:nodeID="casefile_1">
75:                                      <n5ui:systemID>da9b5210-9faf-440f-98ef-2d3745a553d4</n5ui:systemID>
76:                  <n5ui:fileID>11/453213</n5ui:fileID>
77:                          <n5ui:title>case file</n5ui:title>
78:                                  <n5ui:description>case file</n5ui:description>
79:                  <n5ui:createdDate>2011-10-11T17:35:16</n5ui:createdDate>
80:                  <n5ui:createdBy>Oslo municipality</n5ui:createdBy>
81:                  <n5ui:finalisedDate>2011-10-11T19:26:39</n5ui:finalisedDate>
82:                  <n5ui:finalisedBy>Oslo municipality</n5ui:finalisedBy>
83:                                  <n5ui:casedate>2011-10-20</n5ui:casedate>
84:                                  <n5ui:administrativeUnit>Oslo municipality</n5ui:administrativeUnit>
85:                                  <n5ui:case-responsible>executive officer</n5ui:case-responsible>
86:                                  <n5ui:casestatus>being processed</n5ui:casestatus>
87:                                  <n5ui:hasPart rdf:nodeID="caserecord_1"/>
88:                      </n5ui:casefile>
89:
90:                                  <n5ui:records rdf:nodeID="records1">
91:                          <n5ui:systemID>f17170ff-807d-402b-95c7-d9a02709565a</n5ui:systemID>

```
92:                                              <n5ui:recordtype>records</n5ui:recordtype>
93:                            <n5ui:createdDate>2011-10-11T17:50:17</n5ui:createdDate>
94:                            <n5ui:createdBy>Oslo municipality</n5ui:createdBy>
95:                            <n5ui:archivedDate>2011-10-11T19:26:39</n5ui:archivedDate>
96:                            <n5ui:archivedBy>Oslo municipality</n5ui:archivedBy>
97:                                          <n5ui:isPartOf rdf:nodeID="files1"/>
98:                                          <n5ui:hasPart rdf:nodeID="documentdescription"/>
99:                               </n5ui:records>
100:                              <n5ui:caserecord rdf:nodeID="caserecord_1">
101:                    <n5ui:systemID>2927a4b5-9b61-4043-8ddd-fe6dc2da6d61</n5ui:systemID>
102:                                      <n5ui:recordID>11/524136-8745</n5ui:recordID>
103:                                      <n5ui:hasPart rdf:nodeID="registry_entry1"/>
104:                                      <n5ui:hasPart rdf:nodeID="registry_entry2"/>
105:                              </n5ui:caserecord>
106:                              <n5ui:registry_entry  rdf:nodeID="registry_entry1">
107:                                        <n5ui:serialnumber>2154632/11</n5ui:serialnumber>
108:                              <n5ui:registryentrytype>under process</n5ui:registryentrytype>
109:                                    <n5ui:recordstatus>being processing</n5ui:recordstatus>
110:                                        <n5ui:registrydate>2011-10-11T17:33:56</n5ui:registrydate>
111:                            <n5ui:documentDate>2011-10-11T17:35:55</n5ui:documentDate>
112:                            <n5ui:receivedDate>2011-10-11T19:26:39</n5ui:receivedDate>
113:                                        <n5ui:sentDate>2011-10-10-T19:28:42</n5ui:sentDate>
114:                                        <n5ui:isRelated rdf:nodeID="caserecord_1"/>
115:                              </n5ui:registry_entry>
116:                              <n5ui:registry_entry  rdf:nodeID="registry_entry2">
117:                                        <n5ui:serialnumber>2365984/11</n5ui:serialnumber>
118:                              <n5ui:registryentrytype>under process</n5ui:registryentrytype>
119:                                    <n5ui:recordstatus>being processing</n5ui:recordstatus>
120:                                        <n5ui:registrydate>2011-10-11T17:36:58</n5ui:registrydate>
121:                            <n5ui:documentDate>2011-10-11T19:28:55</n5ui:documentDate>
122:                            <n5ui:receivedDate>2011-10-11T19:26:39</n5ui:receivedDate>
123:                                        <n5ui:sentDate>2011-10-11T25:16:48</n5ui:sentDate>
124:                                        <n5ui:isRelated rdf:nodeID="caserecord_1"/>
125:                              </n5ui:registry_entry>
126:                              <n5ui:casefile rdf:nodeID="casefile_2">
127:                                    <n5ui:systemID>31ee43ed-bae1-4518-ada3-5209b2f0baf3</n5ui:systemID>
128:                                    <n5ui:fileID>11/453213</n5ui:fileID>
129:                    <n5ui:title>casefile</n5ui:title>
130:                                        <n5ui:description>casefile</n5ui:description>
131:                                        <n5ui:createdDate>2011-10-11T17:53:26</n5ui:createdDate>
132:                                        <n5ui:createdBy>caseowner</n5ui:createdBy>
133:                            <n5ui:finalisedDate>2011-10-11T17:53:26</n5ui:finalisedDate>
134:                                        <n5ui:finalisedBy>Oslo municipality</n5ui:finalisedBy>
```

```
135:                                    <n5ui:casedate>2011-10-11T17:37:04</n5ui:casedate>
136:                                    <n5ui:administrativeUnit></n5ui:administrativeUnit>
137:                                    <n5ui:case-responsible></n5ui:case-responsible>
138:                                    <n5ui:casestatus>being processed</n5ui:casestatus>
139:                                    <n5ui:hasPart rdf:nodeID="caserecord_2"/>
140:                    </n5ui:casefile>
141:
142:                        <n5ui:records rdf:nodeID="records2">
143:               <n5ui:systemID>cbe0f755-410c-41d1-9cf1-ede0b14c2228</n5ui:systemID>
144:                                    <n5ui:recordtype>record</n5ui:recordtype>
145:               <n5ui:createdDate>2011-10-11T17:37:09</n5ui:createdDate>
146:               <n5ui:createdBy>Oslo municipality</n5ui:createdBy>
147:               <n5ui:archivedDate>2011-10-11T19:26:40</n5ui:archivedDate>
148:               <n5ui:archivedBy>Oslo municipality</n5ui:archivedBy>
149:                                    <n5ui:isPartOf rdf:nodeID="files1"/>
150:                                    <n5ui:hasPart rdf:nodeID="documentdescription"/>
151:                        </n5ui:records>
152:                        <n5ui:caserecord rdf:nodeID="caserecord_2">
153:                        <n5ui:systemID>f4b9a6da-9f29-4bda-bbdb-cf5100ad77e5</n5ui:systemID>
154:                                    <n5ui:recordID>11/2114587-3215</n5ui:recordID>
155:                                    <n5ui:hasPart rdf:nodeID="registry_entry_1"/>
156:                                    <n5ui:hasPart rdf:nodeID="registry_entry_2"/>
157:                        </n5ui:caserecord>
158:                        <n5ui:registry_entry  rdf:nodeID="registry_entry_1">
159:                                    <n5ui:serialnumber>1265487/11</n5ui:serialnumber>
160:                        <n5ui:registryentrytype>under process</n5ui:registryentrytype>
161:                                    <n5ui:recordstatus>being processing</n5ui:recordstatus>
162:                                    <n5ui:registrydate>2011-10-11T17:53:48</n5ui:registrydate>
163:               <n5ui:documentDate>2011-10-11T17:53:48</n5ui:documentDate>
164:               <n5ui:receivedDate>2011-10-11T29:19:56</n5ui:receivedDate>
165:                                    <n5ui:sentDate>2011-10-11T29:55:48</n5ui:sentDate>
166:                                    <n5ui:isRelated rdf:nodeID="caserecord_2"/>
167:                        </n5ui:registry_entry>
168:                        <n5ui:registry_entry  rdf:nodeID="registry_entry_2">
169:                                    <n5ui:serialnumber>5642318/11</n5ui:serialnumber>
170:                        <n5ui:registryentrytype>electronic document</n5ui:registryentrytype>
171:                                    <n5ui:recordstatus>being processing</n5ui:recordstatus>
172:                                    <n5ui:registrydate>2011-10-11T17:53:48</n5ui:registrydate>
173:               <n5ui:documentDate>2011-10-11T17:53:48</n5ui:documentDate>
174:               <n5ui:receivedDate>2011-10-11T29:19:56</n5ui:receivedDate>
175:                                    <n5ui:sentDate>2011-10-11T29:55:48</n5ui:sentDate>
176:                                    <n5ui:isRelated rdf:nodeID="caserecord_2"/>
177:                        </n5ui:registry_entry>
```

```
178:                                    <n5ui:records rdf:nodeID="records3">
179:                          <n5ui:systemID>4e1b554c-dcc7-4b89-820f-7a8e0c44a672</n5ui:systemID>
180:                                        <n5ui:recordtype>record</n5ui:recordtype>
181:                          <n5ui:createdDate>2011-10-11T17:53:14</n5ui:createdDate>
182:                          <n5ui:createdBy>Oslo municipality</n5ui:createdBy>
183:                          <n5ui:archivedDate>2011-10-11T19:26:36</n5ui:archivedDate>
184:                          <n5ui:archivedBy>Oslo municipality</n5ui:archivedBy>
185:                                    <n5ui:isPartOf rdf:nodeID="files1"/>
186:                                    <n5ui:hasPart rdf:nodeID="documentobject"/>
187:                          </n5ui:records>
188: </rdf:RDF>
```