

# ORGANISATIONAL CHALLENGES OF THE SEMANTIC WEB IN DIGITAL LIBRARIES. A NORWEGIAN CASE STUDY

**Bendik Bygstad**

Norwegian School of IT

**Gheorghita Ghinea**

Brunel University

**Geir-Tore Klæboe**

G2 Consulting

## Abstract

**Purpose** – The Semantic Web initiative holds large promises for the future for digital libraries. There is, however, a considerable gap in Semantic Web research between the contributions in the technological field and the research done in the organisational field. This paper examines, from a socio-technical point of view the impact of Semantic Web technology on the strategic, organisational and technological levels.

**Design/methodology/approach** A comprehensive case study at the National Library in Norway was conducted, building on two major sources of information: a) the documentation of the digitising project in National Library of Norway and b) interviews with nine different stakeholders at three levels of NL's organisation during the summer of 2007. Top managers were interviewed on strategy, middle managers and librarians were interviewed regarding organisational issues and ICT-professionals were interviewed on technology issues.

**Findings** - Our findings indicate that the highest impact will be at the organisational level. The reason is mainly because inter-organisational and cross-organisational structures have to be established to address the problems of ontology engineering, and a development framework for ontology engineering in digital libraries must be examined.

**Originality/value** – IT professionals and library practitioners should be more mindful of organisational issues when planning and executing Semantic Web projects in digital libraries. In particular, practitioners should be aware that the ontology engineering process and the semantic meta-data production will affect the entire organisation. For public digital libraries this probably will also call for a more open policy towards user groups to properly manage the process of ontology engineering.

*Keywords: Semantic Web, Digital libraries, Organisational impact*

## 1 INTRODUCTION

For organisations with large amounts of non-numerical data the Semantic Web initiative holds promises for the future (Berners-Lee, Hendle and Lassila, 2001). The past decade gave us not only the Semantic Web and eXtensible Markup Language (XML), but also such tools as the Resource Description Framework (RDF) and ontology languages (Warren and Alsmeyer, 2005). All together, these concepts and tools provide us with a powerful environment to escape the lexical doldrums of the World Wide Web.

Organisations, however, are socio-technical systems; to work they need both technology and people. The progress of technologies has not been matched by an equal understanding of the organisational issues and challenges associated to the Semantic Web. This applies in particular to one of the largest potential users of this framework, namely large libraries. Whilst issues dealing with strategy (Borgman, 1999), user acceptance (Hill et al, 2000; Thong, Hong and Tam, 2002; Xie, 2008), as well as design and implementation (Champeny et al, 2004; Zhou and Stahl, 2007) are quite common themes in digital library research, with the notable exception of a contribution from Kim and Biehl (2005), little research has been done in the area of organisational implications of implementing Semantic Web in digital libraries. Some of the research done by Klischewski (2006) does have some relevance to digital libraries and his agenda for further research suggests focusing on cross-organisational adoption of the Semantic Web and the ability to set up and manage socio-technical infrastructures. This agenda may also be appropriate to research in digital libraries; however his focus on e-Government as a basis for his research represents a limitation because of the specific definition and use of ontologies in e-Government, and the structure of the organisation which is not necessarily transferable to digital libraries. This suggests that an investigation of the strategic and organisational implications implementing Semantic Web in Digital libraries would be appropriate, and represents the focus of the current paper.

Our research question is: *“what are the strategic, organisational and technological impacts of the Semantic Web on large digital libraries?”* The rest of the paper is structured as follows. In section 2 we discuss the opportunities for large libraries presented by the Semantic Web. In section 3 we describe our research approach and present our case, the National Library of Norway. We discuss our findings in section 4, and conclude briefly in section 5.

## 2 THE SEMANTIC WEB AND DIGITAL LIBRARIES

One of the main challenges in information retrieval from the World Wide Web is how to determine what is most relevant for a user's request. The basic elements in the most common and widespread search engines are purely statistical calculations and referential frequency from other sites. The need for a more semantic oriented information representation on the web has been acknowledged for some years, and considerable

research activities have been performed in the field. The benefit of the semantic approach based on Ontologies is that it gives us a framework for searching and browsing information objects on the web and gives more relevance and accuracy to search processes. It will also enable us to put machine-readable data on the Web, data which can be processed by automated tools as well as people (Berners-Lee, Hendle and Lassila, 2001).

The Semantic Web and semantic technologies are focused around knowledge representation through ontologies. An ontology is defined as a formal representation of the knowledge in a specific domain i.e. what exists can be formally represented by ontologies (Ferran and Mingullion, 2005). In the context of the Semantic Web, Ontologies can be queried and updated, both by computer and human users, to explicitly represent objects and concepts that exists in some context, together with the relationship that holds among them. Adams (2002) related ontologies to the librarian taxonomy term, calling both hierarchies of structured vocabularies. However, ontologies also include a set of semantic rules which is used to infer knowledge from a structured hierarchy of information, thus giving the complete structure a semantic meaning (Gruber, 1995).

## **2.1 The underlying technology**

The technology behind the Semantic Web is based on XML, which permits us to describe objects using tags. XML is widely used and creates the basis for a large number of standardized languages in a wide range of business activities. However, XML does not have any facilities to model the relationship between the entities described, and to overcome this limitation the Resource Description Framework (RDF) was designed (Warren and Alsmeyer 2005). Moreover, RDF enables data modeling; however it does not allow us to define any semantics. This is made by the description of an ontology using RDF Schema (RDFS) or an Ontology language, such as OWL which recently been described and adopted by the W3C. This language provides us with a graphical interface for ontology engineering in addition to an inference machine which helps us to detect new information relationships in the ontology as well as inconsistency in the ontology itself.

Other evolving standards for multimedia content description are the MPEG-7, which provides an interface for information meaning to be accessed by automated systems, and the MPEG-21 multimedia framework which aims to enable the transparent use of multimedia resources over a wide range of technological platforms, together with a define and interoperable framework for Intellectual Property Management and Protection (IPMP) (Bormans and Hill 2002) which is of special interest and concern for digital Libraries with multimedia content.

## **2.2 Digital libraries**

A digital library is a “particular kind of information system and consists of a set of components, typically a collection (or collections), a computer system offering diverse services on the collection (a technical infrastructure), people, and the environment (or

usage), for which the system is built” (Fuhr et al., 2007). Digital libraries today are mainly repositories of digitised documents and if they are to become repositories of knowledge, and thereby represent a strategic asset for the organisation, semantic annotation has to be connected to the digital content and semantic meaning has to be drawn from the documents (Chen, 2003). However many digital libraries hold not only digitised documents but also other several different media types such as digitised sound, film and pictures. The key challenges for digital libraries were already in the mid nineties identified as (Lynch and Garcia-Molina, 1995):

- *Interoperability*: The ability of digital libraries to share and relate information between different types of digital content across heterogeneous platforms.
- *Description of objects and repositories*: The need for a commonly accepted naming convention in the description of objects and repositories to facilitate search and information retrieval from different distributed resources.
- *The management of storage and collection of information*: The ability to store, index and retrieve non-textual and multimedia content.
- *User interface and human-computer interaction*. How information is visualized and presented to the user, and how a user is to navigate in large information repositories.

The Semantic Web contributes to a solution to these challenges (Warren and Alsmeyer, 2005). However, the challenges are not limited to the information structuring, but relate also to strategy and organisation. For instance, in the context of digital libraries, search capabilities are often limited by the general search technology used by traditional search engines, and are often off-the-shelf software from commercial search companies like Google and Fast Search and Transfer. Digital libraries, however, are usually more complex, with multiple media-types involved and meta-data stored in several databases, and the interoperability challenge is very much at the forefront in these libraries. Nevertheless, the search capabilities in these digital libraries are limited in much the same way as for more rudimentary digital libraries because of the lack of both semantic annotation and Ontologies. Indeed, one of the biggest challenges for organisations posed by the Semantic Web is no doubt the building and maintenance of Ontologies. This is a huge task which demands considerable resources even if one is to implement automated or semi-automated knowledge discovery systems (Grobelnik and Mladenčić, 2005 ).

These challenges are mainly technical. The problems of relevance and semantics and how information is connected are problems that organisations must address as the volume of information is increasing. To date few studies have addressed the organisational context that is needed to understand and manage them. In the next section we present an in-depth case study which allows us to discuss these aspects in more detail.

## **3 CASE STUDY**

### **3.1 Method**

In this section we first present our method, then proceed to describe our case. For the investigation of the research question we chose a qualitative and interpretive approach (Walsham, 1993), conducting a case study. The selected case was a large scale digitising project in the Norwegian National Library (NL). The NL has decided to digitise all its material covering 15 media types, and make this material available for users on the WWW. It has carried out a pilot project during the last two years to gain experience and is commencing a second phase of the project where new material types are to be digitised and made available.

#### **Data collection**

The case study builds on two major sources of information: a) the documentation of the Digitising project in NL and b) interviews with nine different stakeholders at three levels of NL's organisation during the summer of 2007. The informants were selected as follows: At each level three informants were identified, contacted and interviewed. Top managers were interviewed on strategy, middle managers and librarians were interviewed regarding organisational issues and ICT-professionals were interviewed on technology issues. The interviews were conducted in a semi-structured manner, and were taped for analysis of data. The interview guide is documented in Appendix A.

#### **Data analysis**

The data were analyzed at the three organisational levels, drawing on the nine interviews and on the substantial amount of project documentation. The analysis model defined for this case study aims to understand and explain the findings through the analysis of the impact Semantic Web technology will have on the following five areas at the NL:

- Technology choices
- Organisational Involvement
- Strategic choices
- Organisational consequences
- Resource availability

For each of these areas the respondents' answers were categorized into three levels of perceived impact; Low, Medium and High. The initial score was checked with written sources for consistency, and special care was given to cases where the three respondents expressed differing views. At the end, the scores were accumulated up into the three levels, as the overall results shown in table 1.

### **3.2 The digitising project**

The Digitising project was initially born as a result of a strategic process in NL focusing on The Digital National Library (NB 2006). NL had been digitising different media types for more than a decade before the advent of the digitising project. In 2005 NL's digital collection consisted of more than 50.000 hours of radio programs, 200.000 photos and more than 200.000 newspapers issues. The NL had already established a digitising production line for these material types, and has since gained considerable experience in the field. The new elements in the digitising project were an additional 11 material types to be added to the production line, and all material types were to be made available through a generic user interface using traditional search engine technology. A pilot project was initialized during the spring of 2006 adding books and magazines to the digitising production line, and a user interface was designed to make the digitised material available for users. NLs digital library was officially opened in April 2007 (see <http://www.nb.no>).

The critics soon arrived on the arena, and there were mainly two areas that were criticized: a) the lack of semantic context in the digitised material, and b) the lack of tools and instruments for users to interact and to add domain knowledge to the information objects (NBF, 2007). In its defense, NL stated that the released product was a Beta-product, and that the purpose of the pilot project was to gain experience in the field.

In parallel with the pilot project a report containing an overall plan was developed to define handling procedures for the remaining material types (Hauknes, 2007). The report's recommendations were that a new pilot project should be initialized and a production line be established to gain experience with four additional material types. Moreover, it recommended that a new evaluation should be carried out during the summer of 2008 to adjust the project plans and costs before the remaining material types are to be put into the production line. The report's estimate on man-hours was that 20% of the resources should be used on meta-data production, and the rest in the production line. Meta-data production was to be kept at a minimum to ensure the linkage between the digitised material and the original material, and to ensure a correct retrieval through the search engine.

When the second pilot project is completed the ambition is to commit the base organisation to take over the responsibility for the process. This, as the total time span of the project is ten years, and it is not suitable to organise this process as a single project. The process will therefore be organised as a program, and several smaller projects will be initiated during the next ten years, hopefully resulting in a total digitised collection at NL in 2018. How the commitment of the basis organisation is to be ensured is not stated in the project report; however the main challenges for the organisation in this process are identified as:

- To ensure that the responsibility and the decision areas and levels in the organisation are clearly defined and communicated.

- To establish good structures for personnel handling throughout many years of digitising activities demanding high production volumes from the organisation in parallel with ordinary NL tasks.
- To ensure that there are enough resources in different departments so the impact from the digitising process on ordinary resources to keep to a minimum.
- To ensure that the demand for high production volumes is not putting the quality of the information retrieval at risk.
- To ensure that the demand for quick-wins in the start-up phase is not putting the quality of the long-term software development involved at risk.

The meta-data challenges however are not defined as a part of the digitising program, but are to be handled by the basis organisation which will make the decisions regarding meta-data registration. Furthermore, the question of retrieval of information from the digital library is not defined as a part of the digitising program, but is to be handled by a separate project, directly controlled by the National Librarian, in parallel with the program. The time span for this project is not known, but the main focus for this project is to make the digitised content available on the web and thus the focus on improvements in the traditional search engine technology is high. The separation of the two projects may look unreasonable for both technological and organisational reasons, but on the other hand one may argue that it reduces the overall complexity of the digitising pilot project.

## **4 ANALYSIS AND DISCUSSION**

One of the observations during the interviews was that the participants had different perceptions of the concept of the Semantic Web. However, the need for more semantics in information retrieval was expressed by most of the participants, and the knowledge of topic-catalogues/topic-maps helped in the understanding of the need for ontologies in a semantic paradigm. The findings are structured according to the analysis model presented, and table 1 illustrates the perceived relative impact of the introduction of Semantic Web technology in a digital library setting.

<Table 1 here >

### **4.1 Strategic level**

Arguably, libraries have made their first critical strategic decision regarding the Semantic Web when they decide to start the process of digitising their information objects. The second strategic decision is how and what kind of meta-data is to be produced and how this is going to be organised (Lytras et al, 2006). The first decision is technology focused, while the second is socio-technical and cross-organisationally focused. The motivation for these

decisions is mainly political because they are set to solve the problem of availability and relevance as the information volume grows, often across several media types.

The findings at the strategic level indicate that the introduction of Semantic Web technologies will have a high impact on organisational involvement and resource availability. This is mainly based on the interviews and the experience from the digitising pilot project: if Semantic Web technology is to be implemented, the entire organisation must support the initiative. On resource availability the indicator is most likely affected by the resource availability on the current digitising pilot project, and the awareness that even semi-automated semantic annotation will be highly resource-consuming. One senior manager remarked:

*“The main challenge for the digitising pilot project is resource availability, and to make the organisation (NL) realize that this is the problem.”*

The findings also indicate a medium impact on organisational consequence. The interviews indicated that the current organisational structure is built to support meta-data production, and therefore there is no need for re-organising because of the introduction of the Semantic Web. However the socio-technical infrastructure needed to build and maintain ontologies might demand some changes in the organisational structure. Ontology building will be an inter- and cross-organisational effort, especially when there are different material types involved which are handled by different parts of the organisation.

The low impact indication areas are Technology choices and Strategic choices:

*“The technological choices made prior to the Digitising pilot project are based on a conservation perspective and an availability perspective”*

This quotation from the interviews may be explained by the fact that most of the strategic decisions made prior to the Digitising pilot project also cover the strategic decisions that have to be made prior to a Semantic Web introduction.

#### **4.2 Organisational level**

At the organisational level the score is high on four of the five areas. The informants perceive the magnitude of the organisational challenges facing NL of the Semantic Web as overwhelming. The basis for this seems to be the perception of the impact from ontology production and maintenance, and the strong focus on semantic meta-data production, all of which have to be established. For the digitising pilot project this would probably involve a total re-thinking of the role of semantic meta-data and how meta-data in general is produced, and the introduction of semi-automated Knowledge Discovery (KD) systems is likely to be the consequence. One middle manager said:

*“The meta-data production in the digitising pilot project is primarily a registration of meta-data for preservation purposes. Semantic annotation of the information objects would be an entirely new paradigm for the organisation.”*

The impact on technological choices registers a high score probably because the overall perception of the tools used today is that these are inadequate in a Semantic Web paradigm. This may be questionable because some of the tools used do have some support for semantic annotation; however this functionality is not used in meta-data production because there are no systems operational today that can make use of this semantic meta-data in the search process. The NL is reducing the total number of meta-data registration systems from 35 to 2. Nonetheless, the overall picture is justified.

As regards organisational involvement the interviews indicate that the Semantic Web will have a high impact on this area, as illustrated in this quotation from another middle manager:

*“The organisation will have to be more involved in the Digitising pilot project both in the planning phase and the operational phase of every new type of information objects put into production.”*

This score may be interpreted as the result of the perceived lack of organisational involvement in the Digitising pilot project today rather than a full understanding of the organisational involvement needed in a Semantic Web paradigm. However, a high organisational involvement for digital libraries transforming to Semantic Web technology this involvement is probably crucial, because the knowledge in Ontology engineering and maintenance and in the semantic annotation of documents will be an inter- and cross-organisational knowledge. Klischewski's (2006) call for cross-organisational infrastructures is highly appropriate as one of the critical success factors in Semantic Web implementations in digital libraries. The interviews also indicate concerns related to the present meta-data strategy, and the perception that a new semantic based meta-data strategy is needed in a Semantic Web paradigm.

On organisational consequences the interviews indicates the need for cross-organisational structures. On the other hand the informants express a certain fatigue in the organisation because of the re-organisations process the organisation recently have been through, wanting to solve the matter within the existing organisational structures. Not surprisingly, the interviews also indicate that the resource availability is a high-score area. This is a hot topic in the existing Digitising pilot project, and the perception of this as an area where the introduction of Semantic Web technology will have a high impact might be an expression of concern for the general resource availability in the organisation as well as a statement of what might be one of the critical success factors in the process if Semantic Web technology is to be implemented.

Any new information system will have some impact on the organisational level when introduced and implemented (Orlikowski and Walsham, 1996). This is also the case for the Semantic Web but informants highlight that when this technology is implemented in a digital library the organisational impact probably will be universal because the main organisational

activities in this organisation is the digitising of information objects and meta-data production (Hauknes, 2007). In a Semantic Web paradigm ontology engineering and maintenance will be an additional task which might possibly force the organisation into a more inter-organisational and cross-organisational structure. In this situation the question about organisational readiness and structural changes to meet the challenges from the Semantic Web is highly adequate.

#### 4.3 Technological level

At the technology level the interviews indicate that the overall impact from Semantic Web technology is considered to be low or medium. Said one IT manager:

*“NL is in front internationally regarding digitising of information objects. Actually no one, not even we, has done a lot in the digitising area when we look at its potential.”*

As illustrated, this might be related to the perception amongst the informants that the NL is among the top ten in the international library community when it comes to the use of technology in library systems in general, and in digital library setting in particular. NL has been digitising content for over a decade, and in such areas as the digitising of the non-textual material types, they have competent and experienced personnel. The impact Semantic Web technology will have on the Digitising project is therefore regarded by the informants as low, probably because of the high confidence in their own technological capabilities. However, the organisation lacks first-hand experience in ontology engineering, and the software tools available for this purpose are not known in the organisation. Although there are discussions about semantic structure, RDF is not mentioned as an applicable standard for creating these structures. However the XML experience and competence in the organisation is high and this would probably reduce the impact at this level.

Moreover, the impact from the Semantic Web on technology choices is perceived as low. This makes sense because the technology platforms in use at NL today are mainly based on open standards and will support the implementation of Semantic Web technology, and as RDF is highly connected to XML the knowledge needed is partially present in the organisation today. The interviews indicate that organisational involvement has a medium score and this is probably a result of the high involvement of technology in the organisation today. Introduction of Semantic Web technology will probably have to involve the organisation further on the technological level because the technology used in ontology engineering must be introduced and implemented in the organisation. However the impact will probably not be significant because the technology penetration in the organisation already is high.

The impact of strategic choices is perceived as medium high on the technology level by the informants. The strategic choices regarding technology platforms are, as described earlier, based on open standards and in this context the impact probably would be low. However,

the strategic choices regarding semantic annotation will have some impact on the technology level but there are no indications that strategic choices already made on the technology level will in any way obstruct the implementation of Semantic Web technology.

Organisational consequences are perceived by the stakeholders as having a low impact on the technology level. The informants think that there is no reason to assume that the introduction of Semantic Web technology will have any organisational consequences on the technology level i.e. that the technology in itself is forcing any change upon the organisation. Moreover, the resource availability also receives a low score on the technological level from the interview subjects. This might be a bit surprising because it indicates that there is no need for additional technological resources if one is to implement Semantic Web technology. However the availability of technological resources in NL today is good, and this score is probably an expression of this and the perception of the availability of technological resources in the future. This directly addresses Lynch and Garcia-Molina's (1995) third challenge regarding the management of storage and collection of information which, in a digital library, obviously is a highly technological one.

## **5 CONCLUSIONS**

There is a considerable gap in the Semantic Web research field between the research done in the technological field and the research done in socio-technical field. This paper is a contribution to understand and to explain how digital libraries are affected when introducing Semantic Web technology. The research documented in this paper has investigated the different strategic, organisational and technological impacts of the Semantic Web on a large national digital library. We offer three conclusions from this research:

- On the strategic level the impact from Semantic Web Technology in digital libraries is moderate. This conclusion is based on the fact that the strategic technology choices have already been made when the library have decided to move from a traditional library towards a digital library. However, strategic semantic meta-data issues involving the organisation have usually not been decided because of the lack of an Ontology engineering development framework and a semantic meta-data annotation tool which are able to reduce the cost of the meta-data annotation. This represents an unknown risk factor at the strategic level. It also illustrated that top-level support and funding depends on a broad understanding of the key issues.
- On the organisational level the impact from Semantic Web Technology will be high because both the ontology engineering process and the semantic meta-data production will affect the entire organisation both on the inter-organisational and the cross-organisational level. For public digital libraries this probably will call for a

more open policy towards user groups to manage the process of ontology engineering in a proper manner.

- On the technology level the impact will be relatively low because the technology needed to support the traditional digital library already is in place and the leap from this technology to Semantic Web technology is not a giant one.

We acknowledge that these conclusions may be constrained by the fact that a single case approach is chosen, and the fact that the knowledge of Semantic Web technology among the interview subjects was limited. The overall conclusion, however, that the largest impact of the Semantic Web is at the organisational level, is supported by other research. This conclusion might be assumed to be universal to the digital library community because they are all facing the same challenges regarding semantic meta-data production and ontology engineering and maintenance, regardless of organisational connection (public/enterprise) and regardless of language and geographic location.

Our results should be investigated by other research projects, perhaps also in digital libraries in a private enterprise. In addition there are three main areas of further research that may be pursued:

- How cross-organisational ontology engineering structures may be built and maintained
- The development of a digital library ontology engineering framework
- The development of a framework for semantic meta-data harvesting from digital content based on Knowledge discovery

All these three research directions illustrate that, for the Semantic Web to become a reality, it is necessary to understand and manage the socio-technical structures surrounding this initiative.

## REFERENCES

- Adams, K. (2002) The Semantic Web: Differentiating between taxonomies and Ontologies. *Online*, Vol. 26, No. 4, pp. 20-23
- Berners-Lee, T., Hendler, J. and Lassila, O. (2001) The Semantic Web: a new form of web content that is meaningful to computers will unleash a new revolution of possibilities, *Scientific American*, Vol. 5, No. 1.
- Borgman, C. L. (1999). What are digital libraries? Competing visions. *Information Processing & Management*, Vol. 35, No. 3, pp. 227-243.
- Bormans, J. and Hill, K. (2002) *MPEG-21 Overview v.5*, ISO/IEC JTC1/SC29/WG11/N5231: Coding of moving pictures and audio.
- Champeny, L., Borgman, C. L., Leazer, G. H., Gilliland-Swetland, A. J., Millwood, K. A., D'Avolio, L., Finley, J. R., Smart, L. J., Mautone, P. D., Mayer, R. E., and Johnson, R. A. 2004. Developing a digital learning environment: an evaluation of design and implementation processes. In *Proceedings of the 4th ACM/IEEE-CS Joint Conference on Digital Libraries*, Tuscon, AZ, USA, June 07 - 11, 2004.
- Chen, H. (2003) Towards building digital library as an institution of Knowledge, NSF Post Digital Library Futures Workshop, Chatham, MA, available at [www.sis.pitt.edu/%7Edlwkshop/paper\\_chen.html](http://www.sis.pitt.edu/%7Edlwkshop/paper_chen.html).
- Ferran, N., Mor, E. and Minguillón, J. (2005) Towards personalization in digital libraries through Ontologies, *Library Management*, Vol. 26, No. 4/5, pp 206 – 217.
- Fuhr, N., G.Tsakonas, T.Aalberg, M.Agosti, P.Hansen, S. Kapidakis, C. Klas, L. Kovács, M. Landoni, A. Micsik, C. Papatheodorou, C. Peters and I. Sølvsberg (2007) Evaluation of digital libraries, *International Journal of Digital Libraries*, Vol. 8, pp. 21–38.
- Grobelnik, M. and Mladenic, D. (2005) Automated knowledge discovery in advanced knowledge management, *Journal of Knowledge Management*, Vol. 9, No. 5, pp 132 -149.
- Gruber, T (1995). A translation approach to portable ontology specifications. *International Journal of Knowledge Acquisition and Knowledge-Based system*, Vol. 5, No. 2, pp 199-220.
- Hauknes, R. (2007) *Project report from the DigitALT project*, National library of Norway, internal document.
- Hill, L. L., Carver, L., Dolin, R., Smith, T. R., Frew, J., and Rae, M. 2000. Alexandria digital library: user evaluation studies and system design. *J. Am. Soc. Inf. Sci.* Vol. 51, No. 3, pp. 246-259.

Kim, H. M. and Biehl, M. (2005) Exploiting the Small-Worlds of the Semantic Web to Connect Heterogeneous, Local Ontologies, *Information Technology and Management*, Vol. 6, pp 89 – 96.

Klischewski, R. (2006) Ontologies for e-document management in public administration, *Business Process Management Journal*, Vol. 12, No. 1, pp 34 – 47.

Lynch, C. and Garcia-Molina, H. (1995) ITTA Digital Libraries Workshop, available at [diglib.stanford.edu/diglib/pub/reports/iita-dlw/main.html](http://diglib.stanford.edu/diglib/pub/reports/iita-dlw/main.html).

Lytras, M., Sicilia, M-A., Davies, J. and Kashyap, V. (2005) Digital Libraries in the knowledge era: Knowledge management and Semantic Web technologies. *Library Management*, Vol. 26, No. 4/5, pp 170-175.

NB (2006) *Digitising of the National Library's collection*, Project overview, National library of Norway, internal document, February 1<sup>st</sup>.

NBF (2007) NB Digital – mixed first impression, *Norwegian Librarian Journal*, published 25. April 2007, available at: <http://jaguar.intrapoint.no/bibliotekforeningen/article.php?id=1582>

Orlikowski, W.J. and Walsham, G. (Eds.) (1996) *Information Technology and Changes in The Organisational Work*, Chapman & Hall, London.

Thong, J. Y., Hong, W., and Tam, K. 2002. Understanding user acceptance of digital libraries: what are the roles of interface characteristics, organisational context, and individual differences?. *Int. J. Hum.-Comput. Stud.* Vol. 57, No. 3, pp. 215-242.

Walsham, G. (1993) *Interpreting Information Systems in organisations*, Wiley, Chichester, UK.

Warren, P. and Alsmeyer, D. (2005) Applying semantic technology to a digital library: a case study, *Library Management*, Vol. 26, No. 4/5, pp. 196 – 205.

Xie, H. I. 2008. Users' evaluation of digital libraries (DLs): Their uses, their criteria, and their assessment, *Information Processing & Management*, Vol. 44, No. 3, pp. 1346-1373.

Zhou, N. and Stahl, G. 2007. Information behavior of small groups: implications for design of digital libraries. In *Proceedings of the 7th ACM/IEEE-CS Joint Conference on Digital Libraries*, JCDL '07. ACM, New York, NY, pp. 183-184.

## APPENDIX A: INTERVIEW GUIDE

### Semi-structured interviews

The interviews was structured in three groups according to the three levels in the analysis model; Strategic, Organisational and Technological. The questions were likewise structured and presented for the interview objects before the interview was held. Additional questions were asked in the as the interview moved along to elaborate on issues brought up by the interview object. The initial questions on each level were as follows:

#### **Strategic level:**

- *The Digitising project is a large project which probably would affect the entire organisation in the National Library. What are the strategic thoughts behind the Digitising Project and the connection to National Library's basis organisation regarding organising, involvement and resource usage?*
- *One of the core components in a Digital library are the meta-data production and its quality. Are there any strategic leads on how the meta-data production is going to happen in the future?*
- *There has been some debate around the digitising project and the need for semantic annotation of the digitised material. Have this been a point of discussion in the strategic work with the digitising project?*
- *If the National Library were to move in the direction of the Semantic Web, what will be the main challenges for the organisation from your point of view?*

#### **Organisational level:**

- *The Digitising project is a large project which probably would affect the entire organisation in the National Library. How deep is the basis organisation in the National Library involved in this project, and have you observed any organisational consequences as such from this project?*
- *One of the core components in a digital library is the meta-data production and its quality. Are the procedures regarding meta-data production changed in any way as a result of this project?*
- *There has been some debate around the digitising project and the need for semantic annotation of the digitised material. Do you feel that this debate is present in the National Library, and what do you feel are the pro and cons for the digitising project regarding this matter?*
- *If the National Library were to move in the direction of the Semantic Web, what will be the main challenges for the organisation from your point of view?*

#### **Technological level:**

- *The Digitising project is a large project which probably would affect the entire organisation in the National Library. What are the reasons behind the choice of technological platform for Digitising Project?*
- *One of the core components in a Digital library is the meta-data production and its quality. How would you characterize the technology behind the meta-data harvesting you are performing today, and are this technology in your opinion relevant when it comes to harvesting semantic meta-data and the use of automated or semi-automated knowledge discovery systems?*
- *There has been some debate around the digitising project and the need for semantic annotation of the*

*digitised material. Do you feel that this debate is present in the National Library today, and how does the ICT professionals at the National Library in your opinion participate in this debate?*

- *If the National Library were to move in the direction of the Semantic Web, what will be the main challenges for the organisation from your point of view?*