

Interoperability in Public Sector: How Use of a Lightweight Approach Can Reduce the Gap between Plans and Reality

Svein Ølnes

Western Norway Research Institute
P.O. Box 163, N-6851 Sogndal, Norway
svein.olnes@vestforsk.no
<http://www.vestforsk.no>

Abstract. Better interoperability between systems, vocabularies, and organizations is considered necessary to most public organizations in order to better meet the demands from the users. The rapid growth of the Internet has been a driving force for both the user expectations and the enabling of such exchange. But succeeding with interoperability initiatives is hard, and the risks of failing are high, mostly because the expectations are too high and the inherent challenges are often underestimated. Many interoperability projects are over-specified and their findings are under-implemented. This paper discusses the challenges of interoperability in public sector and argues for a lightweight approach in order to lower the gap between plans and reality. The *Los* system is illustrated as an example of this lightweight approach to interoperability.

Keywords: eGovernment, interoperability, semantics

1 Introduction: What Difference Interoperability Can Make

This paper discusses challenges to making public websites, and especially municipality websites, easier to use by improving their basic structure and at the same time be able to interoperate with other websites. The ICT system *Los* is described as an example of a lightweight approach to interoperability that is also based on a “bottom-up” approach of designing a shared, controlled vocabulary.

First we take a look at an example showing the immediate internal benefits of using a shared controlled vocabulary like *Los*. If we do a search for ‘waste’ (‘søppel’ in Norwegian) on the website of the two biggest municipalities in Norway, Oslo and Bergen, this is what we get:

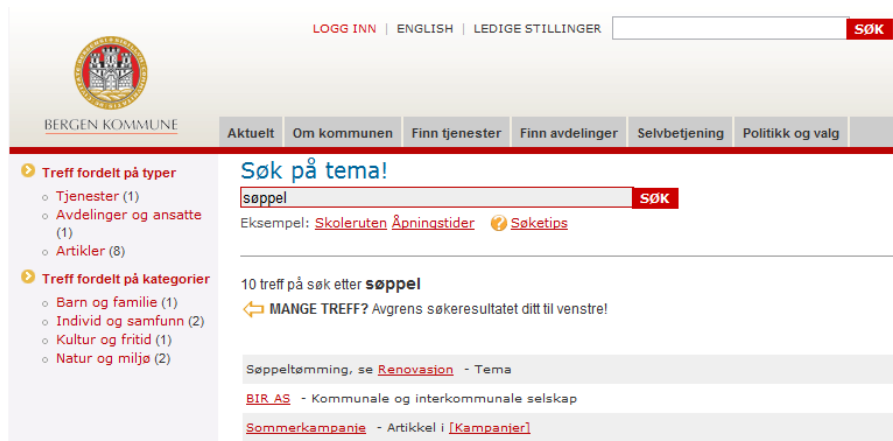


Fig. 1. Search for 'søppel' ('waste') on the municipality of Bergen's website (thematic search) as of March 2010

The search for 'søppel' gives the user a link to the theme 'renovasjon' ('waste handling') which is the official term used by the municipality. But 'søppel' is what most users likely will search for. The municipality's own waste management company BIR is also high on the list of results. In addition, when the user clicks on the term 'Renovasjon', she also gets relevant information from the central authorities regarding this term. This information covers laws and other regulations that affect the particular service. These external links are maintained by the Los system and the individual Los users (mostly municipalities) do not have to worry about it since the information on the municipality's website will be automatically updated.

The same search on the municipality of Oslo's website, gives this result:



Fig. 2. Search for 'waste' ('søppel') on the municipality of Oslo's website, March 2010

The result of the search on Oslo's website gives the user a seemingly random article which happens to mention 'søppel' but is far from leading the user to the core

service of waste handling. It is also quite old, as the rest of the returned articles on the first result page. The user is left confused and there is really no clue where to get more information, whether Oslo municipality's own information or relevant information from governmental bodies.

The improvements in internal information retrieval is important, but even more important is the provision of links to relevant external sources. Bergen will automatically get links to relevant governmental resources regarding the specific term (in this case handling of waste) from the Los system.

2 Interoperability

The examples shown in the introduction point first of all to the need for a well formed internal information architecture. A good information architecture is necessary in order to help users find what they look for. But it is not sufficient in order to make the individual website interoperable with other websites. In order to achieve interoperability there has to be agreed upon standards and terms or concepts.

Interoperability means the enabling ability of information and communication technology (ICT) systems and of the business processes they support to exchange data, and to enable them with the sharing of information and knowledge [1]. The Swedish public agency, Verva adds to this definition that this must be enabled without the need for any special efforts [2].

The European Commission, through their programme IDABC¹, further split interoperability into three layers in the first version of the European Interoperability Framework, EIF [1]: organizational interoperability, semantic interoperability, and technical interoperability

In the EIF 2.0 draft IDABC [3] puts forward this definition of interoperability:

“Interoperability is the ability of disparate and diverse organisations to interact towards mutually beneficial and agreed common goals, involving the sharing of information and knowledge between the organizations via the business processes they support, by means of the exchange of data between their respective information and communication technology (ICT) systems.”

Scholl [4] says that interoperability in essence leads to extensive information sharing among and between governmental entities. However, the obstacles, which prevent a rapid progress in that direction, are not merely technical. In fact, the technology side may prove the least difficult to address, while the organizational, legal, political, and social aspects may prove much more of a challenge.

The extraordinary growth and success of the Internet, and especially the web, has been a key driver in the need for interoperability. With Internet, communication has become a central part of almost every ICT system. And, because Internet represents an open platform, a heterogeneous digital environment has become the norm and not an exception.

¹ IDABC is a Community programme managed by the European Commission's [Directorate-General for Informatics](#). IDABC stands for Interoperable Delivery of European eGovernment Services to public Administrations, Business and Citizens.

In a more closed environment it was easier to look at information exchange as a bilateral challenge that could be solved between pairs of systems, one at the time. XML, the standard syntax for information exchange on the Net, was at first seen as the solution to the interoperability challenges before it was recognised as only the technical part of the interoperability stack. The agreement on using XML as the container of information did not solve the fundamental problems of interoperability. To say that different systems can communicate because they use XML is similar to saying that since you use the Latin alphabet you will be able to communicate. The *meaning* of the communication is vastly underestimated. As Harris et al. [5] states “.. these technologies will not be effective unless the meaning of the tabs, data items, and models of data are properly described, coordinated, communicated, and reused between designers, developers, and users of the information systems”.

Slowly, the challenges of interoperability are beginning to be fully understood, and there is a general acceptance in most governments, for the need of common frameworks to resolve this complex issue. Most European countries have implemented Governmental Interoperability Frameworks and they rely heavily on the recommendations in EIF coupled with the Service Oriented Architecture (SOA) described by OASIS² [6] among others.

In Norway, the Ministry of Government Administration and Reform has proposed an interoperability architecture for public sector in which the SOA principles are at the core [7]. The follow-up of the proposal has yet to be outlined, but a central part of the proposal is the idea of shared components. Public sector agencies should strive to build applications using shared components in order to facilitate openness and reuse. The use of open standards and open source software are also key parts of the architecture.

3 Challenges to Achieving Interoperability

Going back to the definition of interoperability from the European Commission and Verva, it means the ability of ICT systems (and the business processes they support) to exchange data and to enable the sharing of information and knowledge without any specific prearrangement. The “specific prearrangement” means that if a specific interoperability process functions between organisations A and B, it should also function between organisations A and C without any further amendments.

That does not mean that there is no prearrangement necessary. There is prearrangement necessary to facilitate interoperability between systems and processes. Primarily, the interoperability process involves the use of standards. Standards are agreements on terms, concepts and techniques (syntax), and are prerequisites for multilateral information exchange. These types of standards are what David and Greenstein [8] call compatibility standards and are separated from two other kinds of standards; the reference and the minimum quality standards. For the discussion of interoperability, only the compatibility standards are relevant.

² Organization for the Advancement of Structured Information Standards is a not-for-profit consortium that drives the development, convergence and adoption of open standards for the global information society.

The techniques or syntax part of standard agreements is the technical layer of the interoperability model. XML can serve as a good example of technical interoperability standards. The XML specification says how to format the code in your mark-up language and has a strict syntax definition contrary to HTML. But the name of the tags is at the discretion of the developer to define. That means that XML is a very good bearer of information where the names of tags have been agreed upon.

But, even if you have agreed that a specific XML should contain the tag <name>, there will probably be questions of how to interpret it. Is it both the first and the last name or just either one of them? Should, it be written with the last name first? And it gets worse! Consider the tag <address>. Is it visiting address? Is it postal address? What information should it contain?

Clearly, agreeing on terms and thus moving up to the semantic layer, is not enough either to avoid ambiguity. To avoid, or at least reduce, ambiguity there is a need to define the terms and concepts and there is also often a need to describe the relation between concepts. In other words there is a need for an ontology. Ontology is a wide term and in the ICT domain it is often common to distinguish between ontologies with a small or big 'o' [9].

Ontologies come in many flavours, from the simple controlled vocabulary to complex OWL models. The most familiar ontologies are taxonomies, hierarchical structures which essentially only have the relations 'broader' and 'narrower'. Thesauri go a bit further and expand the list of relations beyond mere hierarchical ones. Relations like 'Use', 'Used for', and 'Related Term' are added. But still there is only a finite list of relations.

If we want to create relations freely, the only possibility is to use semantic technology. Using semantic technology, you are no longer restricted to a set of given relations but can create them as you wish. The structure is no longer hierarchical as in a taxonomy and (partly) a thesauri, but a graph structure.

There are currently two standards for semantic technology; the RDF/OWL (Resource Description Framework/Web Ontology Language) suite from W3C [10] [11], also called semantic web, and the ISO 13250 Topic Maps standard [12].

Worldwide the semantic web standard from W3C draws the main attention, but in Norway the situation is special because of a very strong hold for Topic Maps. The latter is due to the relatively large Topic Maps community. Key developers of the Topic Maps standard are also Norwegian and they hold important positions as senior developers and architectures in major ICT consultancies. As a result, there exist several ready semantic solutions for Topic Maps, especially software supporting portal implementations, in contrast to the semantic web where lack of customer ready software currently is the Achilles heel.

Having the technology in place, the next challenge is to map information to the ontology of some kind. This task is difficult even if the ontology is as simple as its taxonomy.

Accepting that interoperability is more than agreeing on a common syntax (like XML), the next challenge is not to be overwhelmed of the complexities and almost try to encompass the whole world in the system model. The risk of doing things overly complex is just as serious as underestimating the challenges of interoperability. Many systems fall victim to the danger of being over-specified and under-implemented. The

clue is, to strike the right balance between accepting a complex challenge and seeking the simplest possible solution to the specific problem.

The development of the European Interoperability Framework from version 1.0 to the proposed version 2.0 can serve as an illustration of how interoperability issues often tend to get overly complex. In the EIF 1.0, the famous and much cited three-layer model with organizational, semantic, and technical interoperability was introduced. The framework was well received and many national interoperability frameworks (NIF) were developed on the basis of EIF.

In the proposed 2.0 version [3] the original three-layer model has been expanded to five layers. Also, two additional dimensions have been introduced to capture standards and interoperability chains.

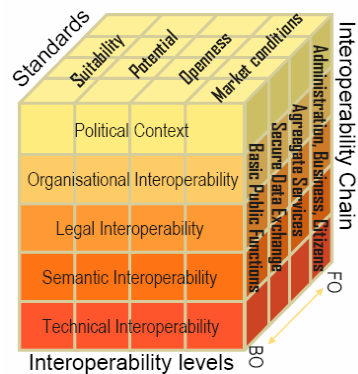


Fig. 3. The new interoperability model proposed in the EIF 2.0 draft (BO is Back Office and FO is Front Office).

What used to be a simple model has become a complex one and it is necessary to ask whether this really adds to the understanding of the interoperability issue in itself. Interoperability is complex but we should use models to simplify complex things, not make them more complex.

The introduction of new interoperability levels can be fruitful in order to reveal hidden knowledge or structures. The addition of a legal interoperability level was first seen in the Swedish report on a national framework for interoperability by Verva [2]. However, the political context should probably be seen as an implicit part of lawmaking. After all laws are the instruments enabling the transformation of political ideas into action.

As the challenges of interoperability finally seem to be fully understood, the huge gap between plans and reality in this field is also recognized [13]. Codagnone and Wimmer found that although a lot of attention has been paid to interoperability from the EU member states, the gap between intentions as expressed in various ICT plans and the actual achievements were assessed as very high.

Lowering the huge gap between plans and reality calls for the highlighting of good practices showing what is achievable with relatively low complexity. The Norwegian information sharing system Los is one of these examples and is explained in detail below.

4 Los – A Lightweight Approach to Interoperability

Los is the name of a system enabling seamless exchange of information between public organisations. The information exchange is based upon a list of keywords describing public services. At present the system is mostly used by municipalities in their information portals, as shown in the introductory chapter. The name Los means a pilot at sea (a navigator) in English and refers to easier navigation to public sector information and services.

4.1 The Los Information Architecture

Los is a controlled vocabulary with terms describing public services. The vocabulary is organized as a thesaurus following the ISO 2788 standard for monolingual thesauri construction [14] and the broader terms suggest a two-level navigation structure, shown in the figure below. To each preferred term a number of *help words* are attached. Help words are synonyms, expired terms, and so on; generally all the terms that are *not preferred*.

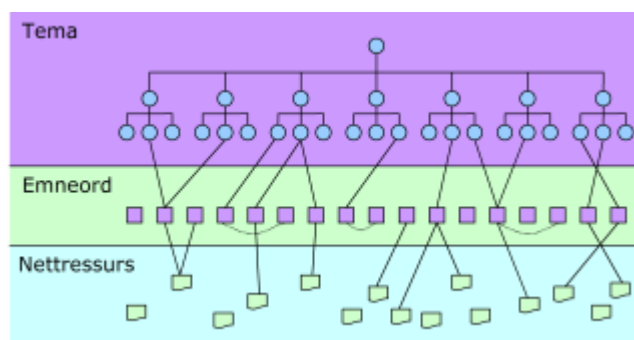


Fig. 4. The Los information architecture. “Tema” is the two-level thematic structure, “Emneord” are the topic words (terms), and “Nettressurs” are the information resources linked to the topic words.

The Los thesaurus was developed by doing a thorough examination of public websites and how they presented their information about their services [15]. Municipalities were also asked what services was most sought after by the users, and search logs, especially the search log at www.norge.no, were studied to identify the words used by the users when searching for public services.

Los has been available from 2007 and almost 100 municipalities have implemented it in their local portals. The key to this uptake is that the vendors of public sector portals have been interested in the project and implemented support for it in their portal systems. Through the key vendors of municipality portal system Los is therefore available to 80-90 % of the municipalities in Norway, although only a third of these have actually started using it.

4.2 Classification and Categorization

Los is a system for categorizing information about public sector services. The categorization process involves annotating information resources with Los topic words. The right words can be found either by browsing the Los structure or by browsing or searching the combined topic words and help words.

The users will meet Los either through navigating the structure itself or by searching and hitting a keyword or help word, as shown in the introduction chapter.

The concepts classification and categorization are often used interchangeably. Since both concepts deal with organizing information they are often mistaken for being synonymous. But as Elin K. Jacob [16] points out there is a difference between the two concepts, and to quote Jacob it is “a difference that makes a difference”.

According to Jacob traditional classification is characterized by rigor in that it mandates that an entity either is or is not a member of a particular class. The fact that the classes are mutually exclusive and non-overlapping makes the system itself stable. This contrasts to the system of categorization where nonbinding associations between entities are drawn. Unlike classification these associations are not based on a set of predetermined principles, but on the simple recognition of similarities that exist across a set of entities. This makes categorization a flexible but also unstable system.

The figure below shows some of the principle differences between classes/classification and categories/categorization from the municipality sector. The predefined classes in our example are ‘Person’, ‘Service’, and ‘Institution’ are all relevant building blocks in an ontology describing a municipality. But of course it is only a small part of this ontology. The hierarchical structure often associated with classes and classification is not shown here, but it is easy to imagine both ‘Institution’ and ‘Service’ as top nodes in a taxonomy (e.g. the class ‘Service’ could have a subclass of ‘Educational services’ and so on).

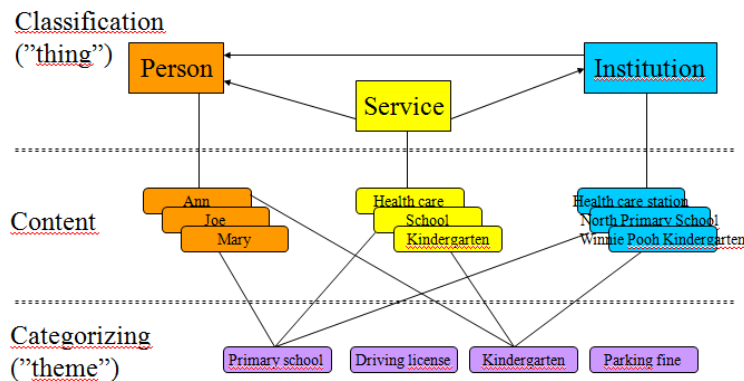


Fig. 5. Classification and categorization

The categorization part on the other deals with keywords that can be attached to content regardless of the class it belongs to. One keyword can easily be attached to different content and thus different classes. . But this “meeting” of classification and categorization is not troublefree. It is in some senses different worlds meeting, and

often colliding. However, they can play together if we regard the differences and turn them into strengths.

The classification and categorization concept division can also be found in today's main approaches to semantic technology. The semantic web, represented by the W3C standards rdf and owl is tightly coupled with the classification system in classes being the main concept and taxonomy structures denoting subclasses.

The other standard for semantic technology is Topic Maps. Although Topic Maps can be used to model typical classification systems, modelling using this standard normally tends to lean much on the categorization system. The reason for this can be that Topic Maps comes from the domain of library information science while semantic web has its roots in mathematics and logics [17].

Los is about categorization and therefore is only part of the answer to building a complete ontology for a municipality web portal. Los has to be integrated into a richer ontology in order to make the most out of it

4.3 Information resources, not services

A very important aspect of Los is the underlying semantics and the description of the key concepts. Los does not link to services but to information about services, whether they are electronically available or not. This distinction between *services* and *information that describes services* is very important in order to get a coherent semantic system where definitions of concepts are clear and unambiguous. The problems that can occur with unambiguous definitions were encountered during the test phase when the Los ontology met the ontology for the Bergen municipality web portal. It became clear that the municipality of Bergen, as one of the pilot users of Los, used a different definition of *service* than Los and that caused failures in the information integration process.

The following example can illustrate the difference between a service and a resource describing the service. Kindergarten is a service from the municipality (or a private company). Providing an online application form for kindergarten does not mean that there is an online service for kindergarten. The kindergarten is still the same physical thing, the online application form is not the service. There is considerable confusing and misuse of terms in regard to this. In our daily speaking we understand each other even if we are not very precise, but for a semantic system it is of utmost importance to clarify the terms and use them in a precise way.

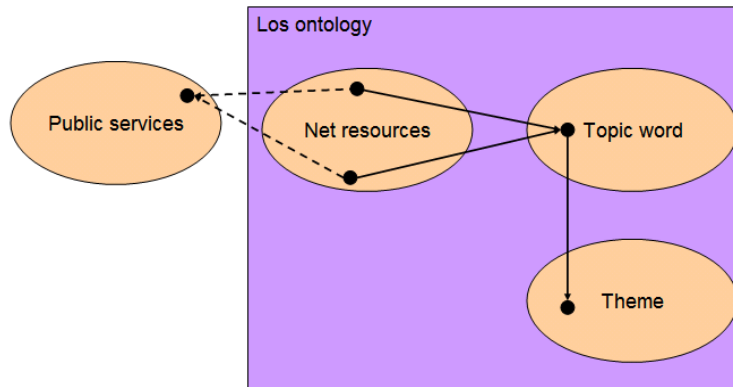


Fig. 6. The general Los ontology. Notice the distinction between services and information about services. Los is about the latter.

4.4 Lightweight Approach to Information Sharing

As shown in the introduction chapter Los is well suited to improve the information architecture on a public website. But the real value in the system lies in the ability to share information across organizations and across sectors. Los comes with an extensible collection of governmental resources annotated with the Los vocabulary. A subscriber to Los will automatically get access to these resources and can use the structure and vocabulary to adjust it to own information. By annotating one's own information the organization using Los can combine local and central information structured according to the predefined vocabulary. The result is that external links will be automatically linked and updated and there is no maintenance for the organization.

The vocabulary with the corresponding resources can be fetched from the Los hub by queries following the REST principle [18]. This way the information sharing process is completely open and foreseeable.

A local municipality using Los can share its annotated resources with others using the lightweight publishing protocol Atom [19]. Los resources are described using Qualified Dublin Core [20], but due to the dumbing-down principle the reader of the RSS based Los-information can skip all the DC details it does not understand and just show the known RSS elements.

Locally annotated Los resources are also expected to be expressed in the original document using Qualified Dublin Core. This way also systems not knowing of Los can make use of some of the added metadata. This way Los will also add value to the current initiatives for increased publishing of open data from public sector.

5 Conclusions and Further Research

Common guidelines are necessary to achieve interoperability. The proposed common architectural framework for public sector in Norway states that every (digital) service established should be designed for interoperability [7]. Interoperability is one of seven proposed principles laid down for the architectural framework, with the service-oriented approach as the overlaying principle.

But the report says little about how to achieve interoperability although it should be mandatory for every new (digital) service to at least discuss the needs and the possibilities. At the same time the challenges of interoperability can seem daunting to most organizations and this can itself delay the necessary development.

In the meantime, waiting for such guidelines, it is important to shed light on good practices like Los. It represents a lightweight approach to interoperability that is valuable as a way to lower the gap between plans and reality in interoperability. The immediate challenge for this and similar solutions, once they are fully implemented, is to open up and harmonise with each other and other initiatives both in public and private sector.

There is a need for a more thorough evaluation of interoperability projects or initiatives to better understand what the main challenges are. It would also be interesting to look at different methodological approaches in these projects; i.e. a “top-down” or a “bottom-up” approach.

Another interesting research issue is how lightweight approaches fit together with the more demanding interoperability initiatives and frameworks.

It would also be interesting to look at the use of Los from a municipality’s perspective and whether it has improved user satisfaction and reduced the need for physical contact or telephone or email contact with the municipality’s administration.

References

1. IDABC: European Interoperability Framework for Pan-European eGovernment Services v. 1.0, Luxembourg. <http://ec.europa.eu/idabc/servlets/Doc?id=19529> (2004)
2. Verva : Nationellt ramverk för interoperabilitet, <http://verva.24-timmarswebben.se/upload/verksamhetsstod/presentation-hearing.pdf>, Stockholm, (2008)
3. IDABC: European Interoperability Framework for Pan-European eGovernment Services, v. 2.0 draft, Luxembourg. <http://ec.europa.eu/idabc/servlets/Doc?id=31597> (2008)
4. Scholl, H. J.: Interoperability in E-Government: More than Just Smart Middleware. HICSS, vol. 5, pp.123, Proceedings of the 38th Annual Hawaii International Conference on System Sciences (HICSS’05) - Track 5 (2005).
5. Harris, S., J. Gibbons, J. Davies, A. Tsui, and C. Crichton: Semantic Technologies in Electronic Government, In Tomasz Janowski and Teresa Pardo, editors, ICEGOV 2008. Pages 45 — 51. Cairo. December, 2008. ACM Press.
6. OASIS: Reference Model for Service Oriented Architecture, <http://docs.oasis-open.org/soa-rm/v1.0/soa-rm.pdf> (2006)
7. Ministry of Government Administration, Reform and Church Affairs: Felles IKT-arkitektur i offentlig sektor. http://www.regjeringen.no/upload/FAD/Vedlegg/IKT-politikk/Felles_IKT_arkitektur_off_sektor.pdf Oslo (2007)

8. David, P. A. and S. Greenstein: The Economics of Compatibility Standards: An Introduction to Recent Research. *Economics of Innovation and New Technology*, 1(½) (1990)
9. Daconta, M. C., L. J. Orst and K. T. Smith: *The Semantic Web – A Guide to the Future of XML, Web Services, and Knowledge Management*. Wiley Publishing, Inc., Indianapolis (2003)
10. W3C: RDF Specification Development, <http://www.w3.org/RDF/#specs> (2004)
11. W3C: OWL Specification Development, <http://www.w3.org/2004/OWL/#specs> (2004)
12. ISO: Topic Maps. ISO 13250:2003
http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=38068
(2003)
13. Codagnone C. and M. A. Wimmer, editors: *Roadmapping eGovernment Research: Visions and Measurestowards Innovative Governments in 2020*. Guerinoni Marco, 2007. ISBN 978-88-95549-00-2. (2007)
14. ISO: ISO 2788:1986: Guidelines for the establishment and development of monolingual thesauri.
http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=7776
(1986)
15. Ølnes, S.: Nye LivsIT – Forslag til ny informasjonsstruktur for LivsIT. VF Report 2-2005, <http://www.vestforsk.no/www/download.do?id=547>, Vestlandsforskning, Sogndal (2005)
16. Jacob, E.K.: Classification and Categorization: A Difference that Makes a Difference. *LIBRARY TRENDS*, Vol. 52, No. 3, Winter 2004, pp. 515–540 (2004)
17. Pepper, S.: The TAO of Topic Maps. Finding the Way in Age of Infoglut. XML Europe 2000, Paris. <http://www.gca.org/papers/xml europe2000/papers/s11-01.html> (2000)
18. Fielding, R. T.: *Architectural Styles and the Design of Networked-based Software Architectures*, Dissertation for the degree of PhD, University of California, Irvine, <http://www.ics.uci.edu/~fielding/pubs/dissertation/top.htm>, (2000)
19. IETF: The Atom Publishing Protocol, RFC 5023, <http://bitworking.org/projects/atom/rfc5023.html> (2007)
20. Dublin Core Metadata Initiative: The Dublin Core Metadata Element Set.
<http://dublincore.org/documents/dcmi-terms/> (2008)