Proceedings of the First International Semantic Computing Conference (IEEE ICSC 2007), Irvine,California, September, 2007. IEEE Press.

Elements of a National Semantic Web Infrastructure —Case Study Finland on the Semantic Web

Eero Hyvönen, Kim Viljanen, Eetu Mäkelä, Tomi Kauppinen, Tuukka Ruotsalo,
Onni Valkeapää, Katri Seppälä, Osma Suominen, Olli Alm, Robin Lindroos, Teppo Känsälä,
Riikka Henriksson, Matias Frosterus, Jouni Tuominen, Reetta Sinkkilä, and Jussi Kurki
Semantic Computing Research Group (SeCo)
Helsinki University of Technology (TKK) and University of Helsinki
http://www.seco.tkk.fi/, firstname.lastname@tkk.fi

Abstract

This article presents the vision and results of creating the basis for a national semantic web content infrastructure in Finland in 2003–2007. The main elements of the infrastructure are shared and open metadata schemas, core ontologies, and public ontology services. Several practical applications testing and demonstrating the usefulness of the infrastructure are overviewed in the fields of eCulture, eHealth, eGovernment, eLearning, and eCommerce.

1 A Semantic Content Infrastructure

The Semantic Web[1] is based on a metadata layer that describes the contents and services on the web in a machine "understandable" way based on ontologies [5, 34]. The idea from the application viewpoint is simple: if the machine understands the contents and services it is dealing with, then better interoperability of web systems can be obtained and intelligent services provided to the end-users.

This papers argues that a conceptual "semantic content infrastructure" is needed for the semantic web, in the same way as roads are needed for traffic and transportation, power plants and electrical networks are needed for energy supply, or GSM standards and networks are needed for mobile phones and wireless communication. A solid, commonly shared infrastructure would make it much easier and cheaper for public organizations and companies to create interoperable, intelligent services on the coming semantic web. In our view, the infrastructure should be open source and its central components be maintained by the public sector in order to guarantee wide usage and interoperability across different application domains and user communities. Business applications and public services can be built most cost-effectively based on such an infrastructure. In order to facilitate inteoperability on the international level, open recommendations of the W3C and other similar organizations are employed and applied to national conventions and languages.

To test and realize this vision, the National Semantic Web Ontology project in Finland (FinnONTO)[2] was initiated by the Semantic Computing Research Group (SeCo)[3] at the Helsinki University of Technology (TKK) and the University of Helsinki. The work started in 2003 and lasts until the end of 2007. The funding consortium behind the project is exceptionally large—38 different organizations in the final phase of the project—and represents a wide spectrum of functions of the Finnish society, including museums, libraries, health organizations, government, media, and education.

The project aims at the following concrete results:

1. *Metadata standards*. Nationally adapted standards for representing metadata in various application fields are being created.

2. *Core ontologies*. A library of central national interlinked core ontologies is developed in order to initiate ontology development processes in Finland. The idea is that after a research period, the participating organizations could continue developing the ontologies collaboratively for machine and human usage, instead of traditional thesauri.

3. *Public ontology services*. An ontology library and web service framework ONKI[4] is being developed [24] to

[1] http://www.w3.org/2001/sw/

[2] http://www.seco.tkk.fi/projects/finnonto/

[3] http://www.seco.tkk.fi/

[4] http://www.seco.tkk.fi/services/onki/

enable ontology usage in ontology development, content indexing, and information retrieval through public web and mash-up services.

4. *Tools for metadata creation*. A key bottleneck of the proliferation of the semantic web is production of metadata. For this purpose, a number of semiautomatic content annotation tools are being developed [37].

5. *Tools for semantic portal building*. A framework for creating semantic search and browsing services based on the multi-facet search paradigm has been developed [29, 41, 28]. Semantic portals are a central application type in the FinnONTO project.

6. *Pilot applications*. The framework is being evaluated by implementing a number of practical applications in the domains of eCulture [18], eHealth [13, 35], eGovernment [33], eLearning [25], and eCommerce [26].

We envision that provision of infrastructure components and tools under open source licensing and as ready-to-use public web services is needed in order to share contents, enforce common practices, and to support organizations in utilizing the new technologies. The project "eats its own dog food" by creating a number semantic portals and applications using the technology developed. These systems provide a public test bed for evaluating the usefulness of the infrastructure. The results are published and demonstrations are available through the home page of SeCo.

In the following the goals of FinnONTO and results obtained are discussed in more detail.

2 Metadata Standards

Metadata standards typically specify what properties to use for content descriptions. For example, Dublin Core[5] lists 15 core elements such as Title, Creator, and Subject. Content interoperability across different application domains is obtained by using commonly agreed elements. In FinnONTO, metadata standards are being developed and adapted in several application fields.

In the eCulture domain, a metadata scheme for representing museum artifact collection metadata was developed and is in use in the semantic portal "MuseumFinland—Finnish Museums on the Semantic Web"[6] [16]. The scheme is being developed further in the follow-up system "CultureSampo—Finnish Culture on the Semantic" [18, 20] that addresses the problem of semantic interoperability of different kinds of cultural contents and metadata schemas. In geoinformatics, the project has participated in creating a national metadata recommendation for spatial information [1]. In eHealth, metadata for health promotion [36] and services are in focus, and in eLearning, the project contributes in developing the national FinnMeta metadata schema for representing learning materials based on the IEEE Learning Object Metadata (LOM)[7] and Dublin Core standards.

Metadata standards are essential for syntactic interoperability on the semantic web but not enough for semantic interoperability. Here the problem is standardization of the *values* used in the commonly agreed metadata elements. For semantic interoperability on the web, large shared reference ontologies are needed. For example, in the MuseumFinland system, the values of the ArtifactType, Material, Creator, PlaceOfCreation, and other elements of a collection artifact are taken from a set of seven ontologies [15]. They contain some 10,000 resources that define the meaning of individual persons, organizations, artifact types, locations, actions etc. Their meaning is shared between the different museums providing the collection metadata content, which enables semantic interoperability of contents. This idea is similar to the Open Directory Project[8], where the shared reference ontology used in annotations contains over 590,000 categories.

3 From Thesauri to Ontologies

The traditional approach for harmonizing content indexing is to use keyword terms taken from shared vocabularies or thesauri [6, 2]. FinnONTO encourages organizations to start transforming thesauri into ontologies, an idea also suggested in [39] and by the SKOS initiative[9]. However, we stress that although a syntactic transformation into SKOS is useful, it is not enough from a semantic viewpoint. The fundamental problem with a traditional thesaurus [6, 2], such as YSA[10], MASA [27], or Agriforest[11], is that its semantic relations have been constructed mainly to help the indexer in finding indexing terms, and understanding the relations needs implicit human knowledge. Unless the meaning of the semantic relations of a thesaurus is made more explicit and accurate for the computer to interpret, the SKOS version is equally confusing to the computer as the original thesaurus, even if semantic web standards are used for representing it.

The idea of using ontologies is to define the meaning of indexing terms and concepts explicitly and accurately enough for the machine to use. This is essential in many application areas, such as semantic search, information retrieval, semantic linking of contents, automatic indexing, and in making contents semantically interoperable. Even

[5] http://dublincore.org/

[6] Operational at http://www.museosuomi.fi/ with an English tutorial.

[7] http://ltsc.ieee.org/wg12/

[8] http://www.dmoz.org/

[9] http://www.w3.org/2004/02/skos/

[10] http://vesa.lib.helsinki.fi/

[11] http://www-db.helsinki.fi/triphome/agri/agrisanasto/Welcomeng.html

with little extra work, e.g. by just systematically organizing concepts along subclass hierarchies and partonymies, substantial benefits can be obtained.

For example, consider the following term entries of the YSA thesaurus, where BT indicates the "broader term" relation used in thesauri.

```
Halley's comet BT Comet
Comet BT solar system
```

We can easily understand its meaning but the machine is confused: Is Halley's Comet an individual or a class of them, such as Comet? Can there be many Halley's comets or only one? Is a comet a kind of solar system or a part of a solar system? Is it a part as a concept or are all individual comets a part of some solar system? Do comets have the properties of solar systems, e.g. own planets, based on the BT relation. Using the BT relations for term expansion, a search for "solar systems" would retrieve comets although comets are not solar systems.

In our work, the central ontology developed is the upper ontology YSO[12] [19]. YSO is based on the general Finnish keyword thesaurus YSA[13] that contains some 23,000 terms divided into 61 domain groups, such as Physics, History etc. YSA is maintained by the National Library of Finland. Since YSA is widely used in Finnish organizations, YSO is an important step in solving semantic interoperability problems in Finland. The ontology is trilingual. Swedish translations of the YSO ontology labels were acquired from the Allärs thesaurus[14], and a translation of the terms into English was produced by the City Library of Helsinki. This makes it possible in the future to align YSO with international English ontologies of the semantic web.

The thesaurus-to-ontology transformation of YSO was not a syntactic one, but was done by refining and enriching the semantic structures of YSA. Our goal was to reorganize and complete the YSA structures into a single simple taxonomic ontology based on the rdfs:subClassOf relation. In standard thesauri this involves the following major problems:

1. *Missing links in the subclass-of hierarchy.* The BT relations do not structure the terms into a full-blown hierarchy but into a forest of separate smaller subhierarchies. In the case of YSA there were thousands of terms without any broader term. Many interesting relations between terms are missing in thesauri, especially concerning general terminology, where BT relations are not commonly specified in practice. In YSO, the concepts were divided into three upper classes (Abstract, Endurant, and Perdurant) in the same spirit as in DOLCE [7]. A central structuring principle in constructing the hierarchies was to avoid multiple inheritance across major upper ontology categories.

2. *Ambiguity of the BT relation.* The semantics of the BT relation is ambiguous: it may mean either subclass-of-relation, part-of relation (of different kinds, cf. [4]), or instance-of relation. This severely hinders the usage of the structure for reasoning [8]. For example, the BT relation cannot by used for property inheritance because this requires that the machine knows that BT means subclass-of and not e.g. part-of relation.

3. *Non-transitivity of the BT relation.* The transitivity of the BT relation chains is not guaranteed from the instance-class-relation point of view. If x is an instance of class A whose broader term is B, then it is not necessarily the case that x is an instance of B, although this a basic assumption in RDFS and OWL semantics [3]. For example, assume that x is a "make-up mirror", whose broader term is "mirror", and that its broader term is "furniture". When searching with the concept "furniture" one would expect that instances of furniture are retrieved, but in this case the result would include x and other make-up mirrors, if transitivity is assumed. This means e.g. that term expansion in querying cannot be used effectively based on the BT relation.

4. *Ambiguity of concept meanings.* Lots of terms in our thesauri are ambiguous and cannot be related properly with each other in the hierarchy using the subclass-of relation. For example, in YSA there is the indexing term "child". This term has several meanings such as "a certain period of human life" or "a family relation". For example, George W. Bush is not a child anymore in terms of age but is still a child of his mother, Barbara Bush. The computer cannot understand this and is confused, unless the meanings of "child" are separated and represented as different concepts (with different URIs) in different parts of the ontology.

In ontologizing the YSA thesaurus lots of terms turned out to be ambiguous, i.e., they could not be placed in one place in the hierarchy. In such cases the term had to be split into several concepts in YSO. However, a lesson learned in our work was that also the general ambiguous concept encompassing several meanings, say "child", can be useful for indexing purposes and should be available in YSO. For example, assume a painting depicting playing children in a park with their mothers watching. When selecting keywords (concepts) describing the subject, it would be tedious to the indexer to consider all the meaning variants of "childness" in YSO, while the single ambiguous indexing term "child" of YSA would encompass them properly in this

[12] http://www.yso.fi/onto/yso/
[13] http://vesa.lib.helsinki.fi/
[14] http://vesa.lib.helsinki.fi/allars/index.html

case. We therefore included some useful ambiguous concepts, such as “child”, in YSO as special *aggregate indexing concepts*. They lay outside of the subclass-hierarchies but can be defined, e.g., in terms of them by using Boolean class expressions as in OWL[15].

Another principle in transforming YSA was that each YSA term should have a counterpart in YSO. This makes it possible to use YSO for more accurate reasoning about content annotated using YSA. Since the original term meanings in YSA change when the term is connected into an ontology, the original YSA terms had to be preserved in the YSO ontology as they are. YSO therefore consists of the following major parts: 1) a meaning preserving SKOS version of the original YSA, 2) an ontology of concepts corresponding to YSA terms, and 3) a mapping between the two structures.

The mapping makes it possible to explicitly tell the relation between YSO concepts and YSA terms. In our mapping schema, the relation between two concepts A and B is defined in terms of extensional overlap that can be expressed as two numerical values in the range (0,1]: 1) how much A overlaps B proportionally and 2) how much B overlaps A. This model is an adaptation of [22] where geographical overlap in area is considered. For example, if A is a subclass of B, then B overlaps A in meaning by 1, and A overlaps B by some value in the range (0,1]; equality means two overlap values 1, and partial overlaps can be expressed by selecting other values. In the first version of YSO equality of YSA and YSO concept is used by default.

4 A System of Mutually Aligned Ontologies

Thesauri are widely used for harmonizing content indexing. Different domain fields have thesauri of their own. The thesauri are typically developed by domain specific expert groups without much systematic collaboration with other fields. When using such thesauri in cross-domain environments, such as the web, semantic problems arise, e.g., due to ambiguity of literal word expressions. For example, in the finance domain the term “bank” has an obvious meaning as an institution, but when considering the nature or musical instrument domains, there are other meanings. In semantic web ontologies, the ambiguity problem is solved by dealing with unambiguous resources identified by URIs instead of literal words. However, support is needed for sharing the URIs across domains and users. If one needs to define the notion of “river bank”, one should be aware of not to mix this concept with “money bank”. On the other hand, if one is defining the notion of “blood bank”, it is possible to use the more general notion of “bank” and modify it, thus sharing this common notion with other kind of banks considered in other ontologies.

[15]http://www.w3.org/TR/owl-ref/

In focused domains and applications it may be possible to agree upon common ontological concepts, but on a larger cross-domain setting, this usually becomes more difficult. Different domains and applications may need different ontological representations even for the same real world objects, and different parties tend to have different philosophical opinions and needs on how to model the world. As a result, there is the danger that the global semantic web will not emerge but there will rather arise a set of isolated, mutually incompatible semantic web islands.

There are various complementary approaches for making semantic web ontologies interoperable. First, ontology mapping and alignment [9] can be used for mapping concepts with each other. Second, ontologies can share and be based on common foundational logical principles, like in DOLCE. This easily leads to complicated logical systems that may not scale up either epistemically or computationally to real word situations and practical usage. Third, horizontal top ontologies, such as the IEEE SUMO[16] can be created for aligning the concepts between vertical domain ontologies. Fourth, ontology engineering support systems for creating ontologies in the first place as interoperable as possible can be created.

In FinnONTO, we share the vision of the IEEE SUO: a shared top ontology is useful for enhancing semantic interoperability between various domain ontologies. In Finland the YSA thesaurus is widely used for content indexing in libraries, museums, and archives of various kinds both in public and in the industry. Since the terms of YSA are used in various vertical domain ontologies, YSA can be considered as a kind of semantic terminological “glue” between many other Finnish thesauri.

Once the structure of the top ontology is defined, the same choices of hierarchical structures can be reused in many cases in the vertical ontologies that typically share lots of concepts with the top ontology. For example, when we created the cultural ontology MAO, based on the Finnish Cultural thesaurus MASA [27], about 2,000 out of MASA’s 6,000 terms turned out to have a corresponding term in YSA. We now work e.g. on the Agriforest thesaurus[17], and on some other thesauri, where thousands of terms originate from YSA. A simple method and a tool has been created by which a Protégé-2000[18] project is created in which both YSO and concepts from a thesaurus are initially mapped to each other by comparing their labels. A human editor then checks the mappings by hand and aligns the two ontologies with each other.

Our goal is a system of mutually interlinked and aligned ontologies, as illustrated in figure 1. In our vision, vertical domain ontologies add semantic depth to the top ontol-

[16]http://suo.ieee.org/

[17]http://www-db.helsinki.fi/eviikki/Welcome_eng.html

[18]http://protege.stanford.edu/

ogy. Interoperability is obtained by aligning the ontologies with each other. Development of several ontologies is underway, including the Finnish geographical place ontology SUO, an actor ontology TOIMO of persons and organizations based on the Universal List of Artist Names vocabulary (ULAN)[19] and national sources, photography ontology VALO, a Finnish version of the iconographic vocabulary ICONCLASS[20], and others.

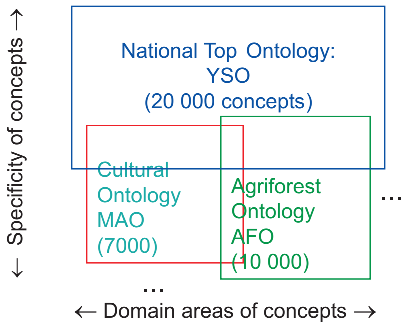

Figure 1. YSO and related vertical ontologies intersect each other and share hierarchical structures.

5 Public Ontology Services

FinnONTO ontologies are provided to end-users as ready-to-use services by the Ontology Library Service ONKI[21] [23]. It provides services for three user groups:

1. For ontology developers, ONKI provides a collaborative ontology development and versioning environment [24].

2. For content indexers, ONKI provides a web-based browser and an AJAX-based mash-up service based on semantic autocompletion [17] for finding desired concepts and for transporting the corresponding URIs from the ONKI server into external applications.

3. For information searchers, ONKI browser can be used for finding and disambiguating keyword meanings, and for transporting the corresponding URIs into search engines and other applications.

In our mash-up application scenario for using ONKI, an external legacy web application, say a cataloging system at a museum, is connected to an ONKI server by associating input fields of an HTML form with ONKI AJAX[22] services. By typing "ban..." in the input field of an HTML form, ONKI automatically tries to complete the string into the possible ontological meanings the user is aiming at, and shows them to her (e.g. "river bank"). After this the right intended meaning can be selected by clicking on it. As a result, the corresponding URI is read into the application and can be used for indexing or for searching in applications such as MuseumFinland supporting ontology-based information retrieval. A demonstration illustrating ONKI's indexing services is available online[23].

Another instance of ONKI services is the ONKI-Paikka[24], an ontology service for geographical names and coordinate data. The main function of ONKI-Paikka is to store and provide services related to the SUO place ontology, yet another YSO-based ontology developed in FinnONTO. SUO ontology has been populated, at the moment, with place information from the Geographic Names Register (GNR) from the National Land Survey of Finland, and with data from the GEOnet Names Server (GNS)[25] maintained by the National Geospatial-Intelligence Agency (NGA) and the U.S. Board on Geographic Names (US BGN). The GNR contains about 800,000 names of natural and man-made features in Finland including information like place or feature type and coordinates. The GNS register contains similar information of about 4,100,000 places around the world excluding places in the United States.

A related FinnONTO effort in the geography domain is to model historical changes of Finnish cities, municipalities, and counties [22], and include the resulting historical ontology in ONKI-Paikka. The dataset used is based on a database created originally at the Geological Survey of Finland (GTK).

6 Open Source Tools

The project has developed the OntoViews tool [29] including the semantic search engine Ontogator [28] and the recommendation server Ontodella [41] for creating semantic portals. In our work, we have generalized the multi-facet search paradigm into using semantic web ontologies, reasoning, and standards [14]. The first application demonstrating the usability of our method and tool set was the MuseumFinland portal[26]. Since then, the tool has been extended and applied in various other demonstrational portals

[19] http://www.getty.edu/research/conducting_research/vocabularies/ulan/
[20] http://www.iconclass.nl/
[21] http://www.seco.tkk.fi/services/onki/
[22] http://en.wikipedia.org/wiki/AJAX/
[23] http://www.seco.tkk.fi/applications/onki/
[24] http://www.seco.tkk.fi/services/onkipaikka/
[25] http://gnswww.nga.mil/geonames/GNS/index.jsp
[26] This application http://www.museosuomi.fi/ got the international Semantic Web Challenge Award (2. prize) in 2004 and a Prime Minister's acknowledgement price in Finland. The software and contents are available open source at http://www.seco.tkk.fi/projects/semweb/dist.php.

[30].

Creation of metadata is a central bottleneck in fulfilling the vision of the semantic web. In order to enable distributed metadata annotation by using centralized ontology servers, we have created a prototype of the SAHA annotation editor [38, 37].

The project also develops tools for semiautomatic content annotation, such as Terminator used for term extraction, and Annomobile for matching keywords with ontology URIs [15, 16]. Natural language processing techniques are being developed for creating tools for annotating Finnish text documents [12, 40].

7 Pilot Applications

FinnONTO technology has and is being applied to case studies in several application domains in order to test its usefulness:

1. *eCulture*. "CultureSampo—Finnish Culture on the Semantic Web" [18, 20, 32] is the next generation of the MuseumFinland portal [16]. The idea of the system is to investigate how cultural contents of different kinds could be made semantically interoperable and be published by a shared publication channel on the semantic web. Content types under study include photographs, fine art, videos, folk poetry, documents, such as manuscripts and biographies, web pages, and cultural process descriptions, such as handwork and farming. This portal also demonstrates new mash-up techniques for visualizing cultural ontological content using e.g. maps [21]

2. *eHealth*. HealthFinland (TerveSuomi)[27] [13, 35] is the pilot version of a national health promotion portal that is being developed in a larger project managed by the National Public Health Institute (KTL)[28]. The contents of the portal will be created in a distributed environment by a variety of Finnish health organizations facilitated by FinnONTO tools, such as ONKI ontology services and the SAHA annotation editor [37].

   An additional research and development topic here is to study how content from other content repositories and semantic portals could be integrated automatically with HealthFinland, and especially how other ordinary web portals, such as the eGovernment portal Suomi.fi (cf. below) could re-use the content of HealthFinland as mash-up Web 2.0 services. Reusing contents would eliminate unnecessary duplication of content work in various governmental organizations, and would enrich the services from the end-user's viewpoint [33]. A first demonstration of the idea of interportal semantic linking is the Orava portal [25] combining materials of an eLearning portal with those in MuseumFinland.

   We also do research on uncertainty in ontologies [11] and fuzzy ontology mapping in order to map end-user vocabularies with medical ontologies used by professionals [10].

3. *eGovernment*. Suomi.fi[29] is the official citizen's information portal provided and maintained by the Ministry of Finance and a number of public organizations in Finland. We created a demonstration of a semantic version of Suomi.fi [33]. The main idea here was to show, that by using shared ontologies, metadata of the web contents and services of different organizations could be integrated automatically into a seamless repository. Different orthogonal views to the content can be provided for the end-user, and related data aggregated along the views and categories in them. For example, the end-user may search the contents based on a classification view of major events in human life, such as "Baby is born". Information related to this event can automatically be gathered from a variety of heterogeneous organizations of health care, social support, legal administration, church etc.

4. *eLearning*. Opintie is an extension of the semantic Orava portal [25] created for the video and learning object materials of the Klaffi portal[30] of the national Finnish Broadcasting Company YLE. Our goal is to create a demonstration of an open publication channel for learning materials combined with the cultural content of the MuseumFinland and CultureSampo portals. As in HealthFinland, the content will be provided by a variety of parties using the ONKI ontology library, its services, and tools provided by FinnONTO. The content focus is initially on high school materials and (Finnish) history. The contents will be automatically integrated with relevant semantic content created in other parts of FinnONTO, especially in CultureSampo.

   Another demonstration system called Opas is a demonstration of how ontology techniques can be used to support help-desk services, both in indexing and in information retrieval [40]. The system is based on an ontologized version of HKLJ, the Finnish library subject heading taxonomy used by the Helsinki City Library, combined with the YSO ontology. The dataset is a collection of over 20,000 indexed question-answer pairs of the national public "Ask the librarian" -service[31], where librarians answer to email questions

[27] http://www.seco.tkk.fi/applications/tervesuomi/
[28] http://www.ktl.fi/
[29] http://www.suomi.fi/
[30] http://www.yle.fi/klaffi/
[31] http://www.kirjastot.fi/tietopalvelu/

of their customers. The service are connected with various online services in order to enrich the answers automatically with links to related material on the semantic web.

5. *eCommerce*. Veturi is a demonstration of semantic yellow pages [31]. Here the idea was to model services and products offered by companies and the public sector with ontologies, and to provide the user with semantic searching facilities in order to ease the problem of finding matching offerings for customer needs. The dataset used was based on the yellow page data of Fonecta Ltd (over 200,000 entries) used in the commercial 020202.fi[32] yellow pages service, and a database of public health services. An innovation of this development was the idea of semantic autocompletion [17] where user input keywords are completed automatically into full search categories and concepts based on the underlying ontologies and reasoning.

8 Discussion

The main idea of the FinnONTO project is to try to build a national infrastructural foundation for the coming semantic web by establishing a large research consortium representing universities, public organizations, and companies, and by working collaboratively together on a national level. Feasibility of the technology developed is tested by several practical application case studies. The international standards and research are guiding the work, but due to language barriers and various national conventions, adaptation and application is needed.

We emphasize the idea that the technology and solutions should be transferred from the universities to the participating organizations and companies. For this purpose, many researchers from the funding organizations have been working with the research team at the university for creating the ontologies and the semantic content of the applications.

Our work focuses on the applied side of the semantic web. In our view, demonstrations of good practical semantic web applications are still largely missing but needed in order to convince public organizations and companies to strive for the era of the semantic web.

Some 30 researchers working in the FinnONTO team at various periods of time have contributed to the work discussed above. FinnONTO project is currently funded by the National Technology and Innovation Agency Tekes and the following public and private organizations (in alphabetical order and in Finnish): AAC Global Oy, Antikvariagroup, CelAmanzi Oy, Connexor Oy, Eduskunta, Elisa Oyj, Espoon kaupunginmuseo, Geologian tutkimuskeskus GTK, Grip Studios Interactive Oy, Helsingin kaupunginkirjasto, Helsingingin yliopiston Viikin tiedekirjasto, Kansalliskirjasto, Kansanterveyslaitos, Kustannus Oy Duodecim, Maa- ja metsätalousministeriö, Maanmittauslaitos, M-Cult ry, Museovirasto, Opetushallitus, Opintoluotsi-projekti/OPM, Osuuspankkikeskus OSK, Sanoma Data Oy, Sanastokeskus TSK ry, Sininen Meteoriitti Oy, Sosiaaliporttiprojekti/Stakes, Stakes (luokituskeskus), Suomalaisen Kirjallisuuden Seura, Suomen maatalousmuseosäätiö, Suomen valokuvataiteen museo, Taideteollisen korkeakoulun kirjasto, Terveyden edistämisen keskus Tekry, TietoEnator Oyj, Valtion taidemuseo, Valtiovarainministeriö, Veljekset Karhumäki Oy, and Yleisradio Oy.

[32] http://020202.fi/

References

[1] JHS 158 paikkatiedon metatiedot (public recommendation for geographic metadata (in Finnish)). ministry of internal affairs, 2006. http://www.jhs-suositukset.fi/suomi/jhs158.

[2] J. Aitchison, A. Gilchrist, and D. Bawden. *Thesaurus construction and use: a practical manual*. Europa Publications, London, 2000.

[3] G. Antoniou and F. van Harmelen, editors. *A semantic web primer*. The MIT Press, Cambridge, Massachusetts, 2004.

[4] C. Fellbaum, editor. *WordNet. An electronic lexical database.* The MIT Press, Cambridge, Massachusetts, 2001.

[5] D. Fensel. *Ontologies: Silver bullet for knowledge management and electronic commerce (2nd Edition)*. Springer-Verlag, 2004.

[6] D. J. Foskett. Thesaurus. In *Encyclopaedia of Library and Information Science, Volume 30*, pages 416–462. Marcel Dekker, New York, 1980.

[7] A. Gangemi, N. Guarino, C. Masolo, A. Oltramari, and L. Schneider. Sweetening ontologies with DOLCE. In *13th International Conference on Knowledge Engineering and Knowledge Management (EKAW02), Lecture Notes in Computer Science, vol 2473*. Springer–Verlag, Berlin, 2002.

[8] N. Guarino and C. Welty. Evaluating ontological decisions with ONTOCLEAN. *Comm. of the ACM*, 45(2), 2001.

[9] A. Hameed, A. Preese, and D. Sleeman. Ontology reconciliation. [34], pages 231–250.

[10] M. Holi and E. Hyvönen. Fuzzy view-based semantic search. In *Proceedings of the 1st Asian Semantic Web Conference (ASWC2006), Beijing, China*. Springer-Verlag, September 3-7 2006.

[11] M. Holi and E. Hyvönen. Modeling uncertainty in semantic web taxonomies. In Z. Ma, editor, *Soft Computing in Ontologies and Semantic Web*. Springer–Verlag, Berlin, 2006.

[12] M. Holi, E. Hyvönen, and P. Lindgren. Integrating tf-idf weighting with fuzzy view-based search. In *Proceedings of the ECAI Workshop on Text-Based Information Retrieval (TIR-06)*, Aug 2006.

[13] M. Holi, P. Lindgren, O. Suominen, K. Viljanen, and E. Hyvönen. Tervesuomi.fi—a semantic health portal for citizens, June 20 2006. Poster paper, 1st Asian Semantic Web Conference (ASWC 2006).

[14] E. Hyvönen, S. Saarela, and K. Viljanen. Application of ontology based techniques to view-based semantic search and browsing, 2004.

[15] E. Hyvönen, M. Salminen, S. Kettula, and M. Junnila. A content creation process for the Semantic Web, 2004. Proceeding of OntoLex 2004: Ontologies and Lexical Resources in Distributed Environments, Lisbon, Portugal.

[16] E. Hyvönen, E. Mäkela, M. Salminen, A. Valo, K. Viljanen, S. Saarela, M. Junnila, and S. Kettula. MuseumFinland—Finnish museums on the semantic web. *Journal of Web Semantics*, 3(2):224–241, 2005.

[17] E. Hyvönen and E. Mäkelä. Semantic autocompletion. In *Proceedings of the first Asian Semantic Web Conference (ASWC 2006), Beijing*. Springer-Verlag, New York, August 4-9 2006.

[18] E. Hyvönen, T. Ruotsalo, T. Häggström, M. Salminen, M. Junnila, M. Virkkilä, M. Haaramo, T. Kauppinen, E. Mäkelä, and K. Viljanen. CultureSampo—Finnish culture on the semantic web. The vision and first results. In *Semantic Web at Work. Proceedings of the 12th Finnish Artificial Intelligence Conference STeP 2006, Vol. 1.*, 2006.

[19] E. Hyvönen, A. Valo, V. Komulainen, K. Seppälä, T. Kauppinen, T. Ruotsalo, M. Salminen, and A. Ylisalmi. Finnish national ontologies for the semantic web—towards a content and service infrastructure. In *Proceedings of International Conference on Dublin Core and Metadata Applications (DC 2005)*, Nov 2005.

[20] M. Junnila, E. Hyvönen, and M. Salminen. Describing and linking cultural semantic content by using situations and actions. In *Semantic Web at Work. Proceedings of the 12th Finnish Artificial Intelligence Conference STeP 2006, Vol. 1*, 2006.

[21] T. Kauppinen, R. Henriksson, J. Väätäinen, C. Deichstetter, and E. Hyvönen. Ontology-based modeling and visualization of cultural spatio-temporal knowledge. In *Developments in Artificial Intelligence and the Semantic Web - Proceedings of the 12th Finnish AI Conference STeP 2006*, 2006.

[22] T. Kauppinen and E. Hyvönen. Modeling and reasoning about changes in ontology time series. In R. Kishore, R. Ramesh, and R. Sharman, editors, *Ontologies in the Context of Information Systems*. Springer–Verlag, Berlin, 2007.

[23] V. Komulainen. Public services for ontology library systems. Master's thesis, University of Helsinki, Department of Computer Science, January 2007.

[24] V. Komulainen, A. Valo, and E. Hyvönen. A tool for collaborative ontology development for the semantic web. In *Proceedings of International Conference on Dublin Core and Metadata Applications (DC 2005)*, Nov 2005.

[25] T. Känsälä and E. Hyvönen. A semantic view-based portal utilizing Learning Object Metadata, August 2006. 1st Asian Semantic Web Conference (ASWC2006), Semantic Web Applications and Tools Workshop.

[26] M. Laukkanen, K. Viljanen, M. Apiola, P. Lindgren, E. Mäkelä, S. Saarela, and E. Hyvönen. Towards semantic web-based yellow page directory services. In *Proceeedings of the Third International Semantic Web Conference (ISWC2004), poster papers*, Nov 2004.

[27] R. L. Leskinen, editor. *Museoalan asiasanasto*. Museovirasto, Helsinki, Finland, 1997.

[28] E. Mäkelä, E. Hyvönen, and S. Saarela. Ontogator—a semantic view-based search engine service for web applications. In *Proceedings of the 5th International Semantic Web Conference (ISWC 2006)*, Nov 2006.

[29] E. Mäkelä, E. Hyvönen, S. Saarela, and K. Viljanen. OntoViews—a tool for creating semantic web portals. In *Proceedings of 3rd International Semantic Web Conference (ISWC 2004), Hiroshima, Japan*, November 2004.

[30] E. Mäkelä, E. Hyvönen, and T. Sidoroff. View-based user interfaces for information retrieval on the semantic web. In *Proceedings of the ISWC-2005 Workshop End User Semantic Web Interaction*, Nov 2005.

[31] E. Mäkelä, K. Viljanen, P. Lindgren, M. Laukkanen, and E. Hyvönen. Semantic yellow page service discovery: The Veturi portal. In *Poster paper, 4th International Semantic Web Conference*, Nov 2005.

[32] T. Ruotsalo and E. Hyvönen. A method for determining ontology-based semantic relevance. In *Proceedings of the 18th International Conference on Database and Expert Systems Applications*. Springer–Verlag, Berlin, 2007.

[33] T. Sidoroff and E. Hyvönen. Semantic e-goverment portals—a case study. In *Proceedings of the ISWC-2005 Workshop Semantic Web Case Studies and Best Practices for eBusiness SWCASE05*, Nov 2005.

[34] S. Staab and R. Studer. *Handbook on ontologies*. Springer–Verlag, Berlin, 2004.

[35] O. Suominen, K. Viljanen, and E. Hyvönen. User-centric faceted search for semantic portals. In *Proceedings of the European Semantic Web Conference ESWC 2007, Innsbruck, Austria*. Springer–Verlag, Berlin, June 4-5 2007.

[36] O. Suominen, K. Viljanen, E. Hyvönen, M. Holi, and P. Lindgren. *TerveSuomi.fi:n metatietomäärittely (Metadata schema for TerveSuomi.fi), Ver. 1.0 (26.1.2007)*, 2007. http://www.seco.tkk.fi/publications/.

[37] O. Valkeapää, O. Alm, and E. Hyvönen. Efficient content creation on the semantic web using metadata schemas with domain ontology services (System description). In *Proceedings of the European Semantic Web Conference ESWC 2007, Innsbruck, Austria*. Springer–Verlag, Berlin, June 4–5 2007.

[38] O. Valkeapää and E. Hyvönen. A browser-based tool for collaborative distributed annotation for the semantic web. September 26 2006. 5th International Semantic Web Conference, Semantic Authoring and Annotation Workshop, November, 2006.

[39] M. van Assem, M. R. Menken, G. Schreiber, J. Wielemaker, and B. Wielinga. A method for converting thesauri to RDF/OWL. Springer–Verlag, Berlin, October 2004.

[40] A. Vehviläinen, E. Hyvönen, and O. Alm. A semi-automatic semantic annotation and authoring tool for a library help desk service. In *Proceedings of the first Semantic Authoring and Annotation Workshop, ISWC-2006, Athens, GA, USA*, November 2006.

[41] K. Viljanen, T. Känsälä, E. Hyvönen, and E. Mäkelä. Ontodella - a projection and linking service for semantic web applications. In *Proceedings of the 17th International Conference on Database and Expert Systems Applications (DEXA 2006), Krakow, Poland*. IEEE, September 4–8 2006.