

Ontology Libraries for Production Use: The Finnish Ontology Library Service ONKI

Kim Viljanen, Jouni Tuominen, and Eero Hyvönen

Semantic Computing Research Group (SeCo)
Helsinki University of Technology (TKK) and University of Helsinki
`firstname.lastname@tkk.fi`
<http://www.seco.tkk.fi/>

Abstract. This paper discusses problems of creating and using ontology library services in production use. One approach to a solution is presented with an online implementation—the Finnish Ontology Library Service ONKI—that is in pilot use on a national level in Finland. ONKI contributes to previous research on ontology libraries in many ways: First, mashup and web service support with various tools is provided for cost-efficient utilization of ontologies in indexing and search applications. Second, services covering the different phases of the ontology life cycle are provided. Third, the services are provided and used in real world applications on a national scale. Fourth, the ontology framework is being developed by a collaborative effort by organizations representing different application domains, such as health, culture, and business.

1 Introduction

The Semantic Web¹ is based on ontologies [1,2,3]. With the help of ontologies, the content and services on the Web can be described with metadata in an explicit, machine “understandable” way which enables, for example, interoperability on a semantic level and intelligent semantic searching and browsing of heterogeneous distributed content in semantic portals [4,5,6]. Utilizing ontologies, including thesauri and other vocabularies, in new and existing applications requires efficient tools for finding, managing, searching and browsing ontologies. *Ontology library systems*² offer functions for managing, adapting and standardizing groups of ontologies, for indexing content with ontologies, and for utilizing ontologies in applications [6,7,8,9].

This paper discusses the requirements for ontology library systems from a practical viewpoint, and presents an approach for building such a service. As a concrete result of the research and a case study, the national Finnish Ontology Library Service ONKI³ is presented. ONKI is a major objective of the National

¹ <http://www.w3.org/2001/sw>

² Ontology library systems are referred to in the literature with terms “ontology servers” and “ontology services”, too. We use the term “ontology library systems” to encompass all these meanings.

³ <http://www.yso.fi/>

Semantic Web Ontology project FinnONTO (2003–2010)⁴ which aims at developing a Semantic Web ontology infrastructure on a national level in Finland [6]. The consortium behind the initiative represents a wide spectrum of functions of the society, including libraries, health organizations, cultural institutions, government, media, and education.

In the following, we first set requirements for an ontology library system in terms of services at different phases of ontology development and usage. After this, the ONKI system and its services are presented along the same phases, and the implementation is discussed. The system has been used in several case applications that are briefly surveyed next. In conclusion, related work is discussed and contributions of ONKI summarized.

2 Requirements for an Ontology Library Service

Requirements for ontology library services can be identified by analyzing the life cycle of an ontology. Based on literature [8,7] and our own work, we identify following parts to be typical in a life cycle of an ontology.

1) *Designing the ontologies.* The first step in the life cycle of an ontology is to design the structure and modelling principles of the ontology based on analysis of the subject domain and by identifying the business and application problems the ontology is intended to solve. The foundational classes, properties and instances of the ontology are created. Sometimes the ontology may also be based on existing vocabularies such as a thesaurus, which are then “ontologised”. The main actors of this phase are workgroups and individual ontologists. Ontology libraries should provide support e.g. for collaborative ontology editing, reuse and alignment [10]. 2) *Populating the ontologies.* Ontologies may consist of huge amounts of instances such as people, organisations and places. Populating and maintaining the information can either be a one-time effort or constantly continuing process even after the ontology has been published. Populating may be e.g. a community-based distributed effort or based on utilizing existing registries and sources as input. Ontology libraries should support such content collection and updating processes. 3) *Publishing the ontologies.* When an ontology has been created, methods for publishing and promoting it are needed to ensure that the ontology is actively used to achieve the benefits of creating the ontology in the first place. In addition to provide such publishing and promoting mechanisms, the ontology library should also have mechanisms - both manual and automatic - for ensuring the quality of the ontologies to be published. The main actors of this phase are the ontology owners. 4) *Finding, comparing and committing to ontologies.* When considering using ontologies for some purpose, the finding of suitable ontologies require support from the ontology library. Typical users of this phase are information architects. 5) *Ontology based semantic application creation.* Learning, evaluating and implementating ontology services to applications require functionalities from the ontology library service for making this

⁴ Our work is funded by the National Funding Agency for Technology and Innovation (Tekes) and a consortium of 38 companies and public organizations.

<http://www.seco.tkk.fi/projects/finnonto/>

process as fluent and easy as possible - and to support the wide usage of ontologies in applications [11,12]. Typical users of this phase are software architects.

6) *Ontology based semantic content creation.* Ontologies are mostly used for describing and indexing content semantically. Ontology libraries should support the work of content indexers by providing efficient tools and services for e.g. browsing ontologies, finding and fetching concepts for annotating purposes [9] or automatically indexing documents [13].

7) *Ontology-based end-user applications.* Ontological content search, semantic browsing, semantic portals are examples of typical ways to provide the end-user with benefits from using ontologies in an applications (see e.g. [4,5,6,14]). Ontology libraries should support creating such end-user applications by providing services for the application builders. Ontology libraries may also provide services directly to the end-users such as the possibility to learn about some domain with the help of ontologies.

3 Finnish Ontology Library Service ONKI

The Finnish Ontology Library Service ONKI is a pilot system for addressing the requirements of an ontology library service on a national scale, but with the special focus on ontology publishing and using them in content indexing, and information retrieval through both user and application interfaces [14,6]. ONKI contains currently over 40 ontologies from various domain areas (see Table 1). Most of the ontologies are freely available to anybody to test and use in their applications.

3.1 Designing the Ontologies, Populating the Ontologies

Ontologies developed within the ONKI framework are mostly created with the Protégé⁵ editor. Version management of the ontology files is done with Subversion⁶. In many cases, the ontologies are based on an existing thesaurus or other content which have first been transformed with an custom-made program to OWL and then been refined ontologically by the ontologist using Protégé and by aligning the ontologies with the Finnish Upper Ontology YSO [6]. Populating the ontologies have been done either with Protégé by the ontologist(s) or collaboratively with the browser-based annotation editor SAHA [15]. For many ontologies, populating have been done with custom-made programs.

Quality of the ontologies is controlled using three ways: gate keeping, quality requirements and training. Gate keeping is practised by selecting only trusted participants in the ontology creation and publishing which include both companies, governmental and non-governmental organizations. Typically the main author of any single ontology in ONKI is the leading authority in Finland of the respective domain. Quality requirements enforced in ONKI cover ontology presentation and ontology creation process issues. The ontology should be presented using some RDF-based ontology language, such as SKOS, OWL or RDF

⁵ <http://protege.stanford.edu>

⁶ <http://subversion.tigris.org/>

Table 1. A selection of ontologies currently available in ONKI (The amount of concepts consists classes and/or instances, depending on the ontology.)

Ontology	Concepts	Format	Public?
<i>Upper and Holistic Ontologies</i>			
Holistic Collaborative Finnish Ontology KOKO	ca. 30,000	OWL	yes
General Finnish Upper Ontology YSO	20,649	OWL	yes
General Finnish Thesaurus YSA	26,633	SKOS	yes
Wordnet	ca. 230,000	SKOS	yes
<i>Cultural Ontologies</i>			
Ontology for Museum Domain MAO	6,775	OWL	yes
Ontology of Applied Arts TAO	29,940	OWL	yes
Finnish Ontology of Photography VALO	22,596	OWL	yes
Ontology for music MUSO	21,650	OWL	yes
Art and Iconography classification Iconclass	26,636	SKOS	yes
Kaunokki thesaurus for fictive literature	4,373	SKOS	yes
Music thesaurus MUSA	931	SKOS	yes
Art & Architecture Thesaurus AAT	27,992	OWL	no
<i>Agriforest and Natural Science Ontologies</i>			
Agriforest Ontology AFO	26,612	OWL	yes
Ontology of Birds AVIO	11,161	SKOS	no
Ontology of Mammals MAMO	6,059	SKOS	no
<i>Health Ontologies</i>			
Medical Subject Headings MeSH	24,355	SKOS	yes
European Multilingual Thesaurus on Health Promotion HPMULTI	1,271	SKOS	yes
<i>Business Ontologies, Governmental Ontologies</i>			
Seafaring thesaurus MESA	1,448	SKOS	yes
United Nations Standard Products and Services Code UNSPSC	20,794	SKOS	no
Finnish Governmental Thesaurus VNAS	6,342	SKOS	yes
<i>Instance Ontologies</i>			
Finnish Geo-ontology SUO	ca. 800,000	OWL	yes
Finnish Time-Location Ontology SAPO	1,102	OWL	yes
Getty Thesaurus of Geographic Names TGN (exluding USA)	142,990	OWL	no
Getty Union List of Artist Names ULAN	ca. 100,000	OWL	no

Schema. The consistency of the ontologies is checked including syntax checking (valid RDF) and conceptual checking manually by the ontologists. One of the most important ways to enforce the quality is that the ontologies in ONKI library are created using a common development process and modelling idea which are supervised by the core YSO developer team [6]. This promotes using compatible development processes in all other ontologies also and thus provides a more compatible collection of ontologies as a result. If possible, ontologies are aligned with a common upper ontology, the Finnish Upper Ontology YSO. This alignment to YSO adds value to YSO, the ontology at hand and the ONKI Library as a whole, because each additional alignment adds new possibilities to find concept relations. To spread good practices and knowledge about the modelling methods used in YSO and other relevant ontologies, training is provided to ontology developers. To enforce the reuse of ontologies, the license of the published ontologies should allow publishing, using and redeveloping the ontology as

freely as possible. The default license used for ONKI ontologies is the Creative Commons license⁷.

3.2 Publishing the Ontologies

Publishing an ontology in ONKI typically contains the following phases: First, the ontology to be published is added to the Subversion repository and the needed configurations for the ontology are created [16]. Second, a URI normalization for the ontology to be published is done where the original URIs are transformed to persistent numeric URIs (PURIs). Instead of (typically) human readable URIs we propose that URIs should not contain any reference to human languages to avoid unnecessary needs for changing the concept URIs e.g. when translating the ontology to some other language⁸. For example, instead using the URI “myonto:semanticweb” we propose using the URI “myonto:p12345”. Third, if the ontology is published part of the KOKO ontology, the automatic updating of KOKO takes place. Finally, the ontology is added to ONKI and made available via different services such as human user interfaces and machine APIs.

If the ontology is maintained in some external system it can be published using ONKI by establishing a publishing pipeline from the external system to ONKI. This method has been used e.g. in publishing the General Finnish Thesaurus YSA, maintained by the National Library of Finland [16]. The thesaurus is fetched each night from the National Library’s server using the MARCXML format⁹. The content is then transformed to SKOS and finally published in ONKI. ONKI provides also an upload functionality “Your ONKI” for publishing SKOS (or other) ontologies in the library. When an ontology has been uploaded, it is moderated by the server administration and if the content is suitable, it will be added to the library. ONKI quality requirements presented earlier are recommended also for Your ONKI submissions.

3.3 Ontology Discovery, Ontology Library Service Evaluation

To support finding, evaluating and choosing an ontology for specific purposes each ontology is described with metadata including title, description, classification, version information and available access methods. Depicted in Figure 1 is the main user interface with the list of available ontologies, which also shows the available access methods for each ontology. Access methods are described later, but include e.g. the possibility to browse and search the ontologies. The ontologies can be described and documented in a wiki, part of the ONKI system.

The list of ontologies is available in RDF for machine usage. When allowed by the publishing license, current and previous versions of ontologies are available for downloading.

⁷ <http://creativecommons.org/>

⁸ The idea of stable URIs is also discussed in <http://www.w3.org/Provider/Style/URI>

⁹ <http://www.loc.gov/standards/marcxml/>

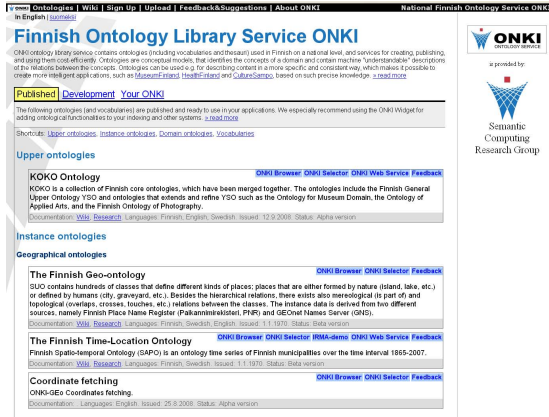


Fig. 1. List of ontologies in ONKI. Each ontology contains the links to the available access methods such as the ontology-specific browser.

3.4 Ontology-Based Semantic Content Creation

For users creating semantic content, e.g. by describing resources by using ontological concepts, ONKI service provides the ONKI Selector, depicted in Figure 2 [9]. With ONKI Selector content creators can find suitable concepts for their annotation tasks. When the ONKI Selector is integrated into a HTML input field, the field turns into a semantic autocompletion search interface. When typing search string into the input field, the matching concepts are returned as a hit list. Desired concepts can be selected from the list and added to the content creation application. Depending on the use case, concept’s URI, label or both of them can be fetched to the application.

In combination with the ONKI Selector, domain-specific ONKI browsers can be used to browse the ontologies when searching for suitable concepts. The browsers have a “Fetch Concept” button which returns the selected concept into the content creation application. ONKI SKOS Browser[16] is an ontology

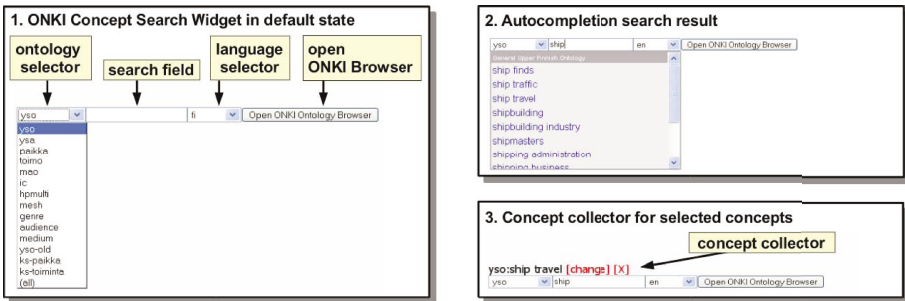


Fig. 2. ONKI Selector

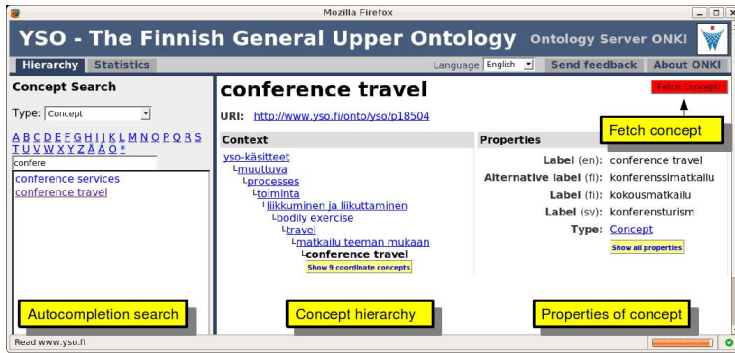


Fig. 3. ONKI SKOS Browser

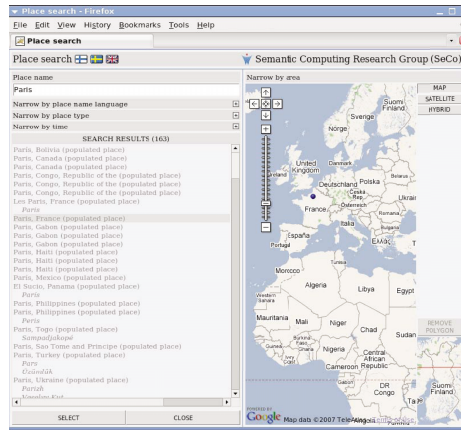


Fig. 4. ONKI Geo Browser

browser for thesaurus-like class ontologies. It supports visualizing and browsing of vocabularies conforming to SKOS recommendation, and also RDF(S) and OWL ontologies with additional configuration. ONKI SKOS Browser consists of three main components: 1) *concept search with semantic autocompletion*, 2) *concept hierarchy* and 3) *concept properties*, as depicted in Figure 3. ONKI Geo Browser [17] is used for accessing geographical instance data with a map interface, as depicted in Figure 4. It provides unambiguous place identifiers (URIs) and coordinates for arbitrary points or polygons to be used in content annotation. ONKI People [18] is used for browsing and searching ontologies of persons, organizations, and similar instance registries.

3.5 Ontology-Based End-User Applications

For ontology-based end-user applications ONKI service provides means for finding ontological concepts and using them, e.g., in information retrieval tasks.

Compared to a simple free text search field, the ONKI Selector aids user to find query concepts with autocompletion search and ontology browsers. The ONKI Selector is useful even if the application is not ontology-based. In that case the labels of the concepts can be used as query terms.

To increase the recall of the information retrieval tasks, ONKI Selector performs query expansion by ontological inference. The properties used for performing the query expansion can be configured separately for each ontology. In class ontologies, a concept is typically expanded to its subconcepts. The query expansion could also be based on partonomy, associative relationships or other relations between concepts. In geographical instance data, a place instance is expanded to places that have historically had overlapping regions with the place. Other possible query expansion methods include partonomy, places with shared regional borders etc.

3.6 Ontology-Based Semantic Application Creation

ONKI services may be integrated into semantic applications at the user interface level by using the ready-to-use user interface component ONKI Selector, domain-specific ONKI Browsers, and by using application programming interfaces (API). ONKI supports the software developer in using ONKI services with the help of helper applications such as the ONKI Selector Builder, depicted in Figure 5, which helps the developer to generate the JavaScript code needed for integrating ONKI Selector into web-based applications. When the desired configuration properties have been set in the ONKI Selector Builder, the resulting JavaScript code can be copied into the application being developed.

The screenshot shows the 'ONKI Selector for KOKO Ontology' configuration page. It features a header with navigation links like 'Ontologies', 'Web', 'Sign Up', 'Upload', 'Feedback & Suggestions', and 'About ONKI'. The main content area provides instructions on how to integrate the selector into a system, showing code snippets for the HEAD and BODY sections of an HTML page. Below this, there is a configuration table for various properties of the generated ONKI Selector.

Property	Value	Description
ontology	koko	The ontology to be searched
key	reset_year_key	Your access key to www.yso.fi services: generate access key
fieldid	lastsearch	The id of the input field used for inputting the search term
language	en	The language of the user interface of the ONKI Selector
addConceptFunction	<input checked="" type="checkbox"/> true <input type="checkbox"/> false	A javascript function defined by you to add concepts from search results to your application
autocompleterSearch	<input checked="" type="checkbox"/> true <input type="checkbox"/> false	Show the autocompletion search field
fieldid	en	The name of the hidden input fields used for storing the selected concepts
lang	en	Search is limited to the labels of this language
languageMenu	<input checked="" type="checkbox"/> true <input type="checkbox"/> false	Show the available languages of the selected ontology in a selection menu
maxHits	50	Maximum amount of search results (*: 0 = unlimited)
maxSelect	10	The maximum amount of concepts that can be selected from search results (*: 0 = unlimited)
onkimenue	<input checked="" type="checkbox"/> true <input type="checkbox"/> false	Show all the www.yso.fi ontologies in a selection menu
showOnkName	<input checked="" type="checkbox"/> true <input type="checkbox"/> false	If the ontology menu is disabled, show the name of the selected ontology instead
fieldid	en	Label to be shown on the left side of the ONKI Selector
openonkibutton	<input checked="" type="checkbox"/> true <input type="checkbox"/> false	Is the opening of the ONKI Browser allowed ("Open ONKI Browser" button)
openonkilabel	Open ONKI Browser	The label of the "Open ONKI Browser" button
parentRestriction	<input type="checkbox"/> true <input checked="" type="checkbox"/> false	Search is limited to the sub tree of this concept (LIR)
profile	<input checked="" type="checkbox"/> true <input type="checkbox"/> false	Is the appropriate profile of the concepts shown as part of the labels of concepts selected from search results
showCaptions	<input checked="" type="checkbox"/> true <input type="checkbox"/> false	Show captions for selector's components
termsInLis	<input type="checkbox"/> true <input checked="" type="checkbox"/> false	Fetch terms instead of URIs
typeRestriction	<input type="checkbox"/> true <input checked="" type="checkbox"/> false	Search is limited to the concepts of this type (LIR)

At the bottom of the configuration table, there is a button labeled 'Generate ONKI Selector Code'.

Fig. 5. ONKI Selector builder

ONKI API provides methods for accessing ontologies, e.g., for searching for concepts, getting metadata of an ontology and performing ontology-based query expansion. ONKI API is implemented as Web Service (SOAP) and JavaScript interface. ONKI API contains the following methods:

- *search(query, lang, maxHits, type, parent)* - for searching for the ontological concepts. Returns a list of matching concepts.
- *expandQuery(URI, lang, maxHits, type)* - for querying for the query expansion for a concept. Returns a list of concepts.
- *getLabel(URI, lang)* - for fetching a label for a given concept URI in a given language.
- *getAvailableLanguages()* - for querying for the supported languages of an ontology. Returns a list of language codes.
- *getAvailableTypeUris()* - for querying for the concept types (rdf:type relations) existing in the ontology. Returns a list of URIs.

Software developers can also utilize the RDF files of the ontologies published in the ONKI service. All concept and instance URIs are designed so that they function also as URLs. When the URI of a concept is accessed with a web browser, the relevant view is opened in the ONKI browser. This means that the URI itself acts as a functional link when added to a HTML page. In accordance to W3C¹⁰, if the URI is accessed with an RDF aware system, the machine readable RDF presentation of the content is returned instead of the ONKI browser's HTML presentation.

4 Implementation and Usage Statistics

The ONKI service is constituted of a loosely coupled set of independent applications such as the ONKI SKOS and ONKI Geo servers which are combined by using a lightweight facade service made with an Apache web server, Apache rewrite rules and PHP scripts. Back-end ONKI applications conform also to the ONKI API described in section 3.6. Technologies used for implementing the various back-end applications include Java, Semantic Web Framework Jena¹¹, MySQL database, Lucene index¹², Direct Web Remoting DWR for AJAX functionalities¹³, Varnish HTTP accelerator¹⁴ and shell scripts and Subversion version management system. The ontologies are presented internally in various RDF formats, typically in SKOS or OWL. With the help of ontology-specific configurations, the ontologies are served to the user in a uniform way.

The ONKI is running as a pilot service publicly available on the web. It was officially launched in September 2008¹⁵. During year 2008 ONKI had ca. 36,000

¹⁰ <http://www.w3.org/TR/swbp-vocab-pub/>

¹¹ <http://jena.sourceforge.net>

¹² <http://lucene.apache.org>

¹³ <http://www.directwebremoting.org>

¹⁴ <http://varnish.projects.linpro.no>

¹⁵ <http://www.youtube.com/watch?v=qG2YhK17ifs>

Table 2. ONKI usage statistics for November 2008

Service	Hits
<i>Human interfaces</i>	
ONKI-SKOS Browser	89,346
ONKI Selector Widget	54,384
Persistent URI redirects	4,415
ONKI Selector builder	1,103
Web Service builder	1,403
ONKI-IRMA	495
ONKI-Geo Browser	203
ONKI-Geo coordinates Browser	168
<i>Application interfaces</i>	
Web service calls	87,388
Javascript calls	18,816
Total	257,721

unique visitors and ca. 104,000 visits. 91 organizations outside the research group have registered an access key for using the JavaScript and web service interfaces which of 25 have actually implemented a test application using the components. For an overview, table 2 presents the usage statistics of the main ONKI functionalities for a representative month¹⁶. The most popular ontologies are the Medical Subject Headings, the Finnish Upper Ontology YSO, the General Finnish Thesaurus YSA and the Ontology for Museum Domain MAO where each ontology got over ten thousand hits during the month.

5 Case Applications Using ONKI

5.1 Content Creation: HealthFinland, CultureSampo and Tilkut

HealthFinland and CultureSampo are two major pilot applications of the FinnONTO project [6]. They demonstrate the usage of Semantic Web technologies in the contexts of health promotion and cultural heritage. Both systems uses ONKI as the ontology server for indexing content especially with the the browser-based annotation editor SAHA¹⁷ [15]. For example (Figure 6), the Finnish General Ontology YSO [6] has been added as a ONKI Selector component to SAHA for finding and fetching annotation concepts.

The web laboratory Owela¹⁸ of the VTT Technical Research Centre of Finland has implemented a service for collecting and sharing text and image clips from the Web¹⁹. In the service one can organize the clips into folders and tag them with different categories. The ONKI Selector is used for tagging the clips.

¹⁶ The hits represents user activities of the system. Outliers caused e.g. by web crawlers have been removed as far as possible.

¹⁷ <http://www.seco.tkk.fi/services/saha/>

¹⁸ <http://owela.vtt.fi/>

¹⁹ <http://owela.vtt.fi/tilkut>

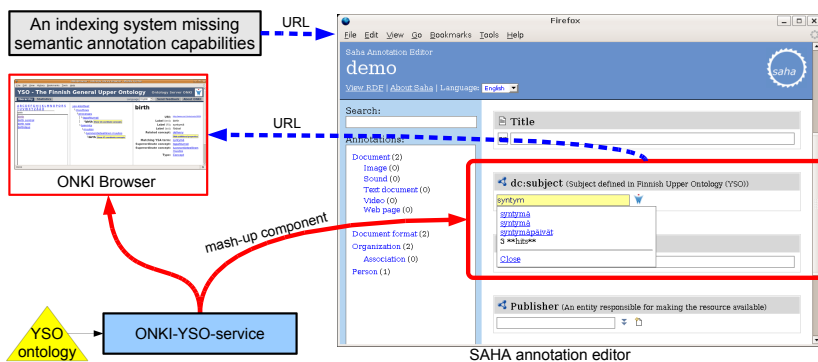


Fig. 6. The Finnish General Ontology connected to SAHA

5.2 Content Search: Kantapuu.fi and eViikki

Kantapuu.fi²⁰ is a web user interface for browsing and searching for collections of Finnish museums of forestry. The collection items are annotated with terms from General Finnish Thesaurus YSA, Thesaurus for Museum Domain MASA and Agriforest Thesaurus. Kantapuu.fi search page is a web form into which query strings are typed as free text. The query strings can be placed into specific fields, e.g. “keywords”, “place of use” or “time of use”. We have created a demonstration page containing a Kantapuu.fi’s search form with integrated ONKI Selectors which can be used for selecting query terms to be used in the Kantapuu.fi search²¹. The ONKI Selector is used for finding terms from vocabularies of the ONKI service. The used vocabularies are the same as those used in the annotation process of the items, or actually their ontologized versions. To find suitable query terms user can utilize the autocompletion search functionality or the ONKI Browser. Thus, the user does not need to be familiar with the vocabularies used in the annotations of the items, as in the case of free text search. The ONKI Selector performs query expansion based on the selected query terms. So, for example a query term “animals” would return items annotated with term “cats”. When the desired query terms are selected, the actual search to the Kantapuu.fi system can be executed.

The ONKI Selector Widget has also been integrated into the Viikki Science Library²² reference database system eViikki²³. eViikki is a search interface for the library’s collections, which consist of scientific literature on agriforestry. The ONKI Selector is used for populating the “keywords” field of the search form of eViikki. The fetched concept labels are used in the information retrieval task. Query expansion is not performed currently.

²⁰ <http://www.kantapuu.fi/>

²¹ <http://www.yso.fi/lusto>

²² <http://www.tiedekirjasto.helsinki.fi/english/>

²³ <http://www-db.helsinki.fi/eviikki/eviikkihaku.html>

6 Related Work

Based on reviews on ontology library systems [7,8], the main focus in existing systems tends to be in supporting ontology development and not the runtime usage of ontologies such as indexing and ontology-based end-user applications. Although ONKI provides support for the whole ontology life cycle, a major contribution of ONKI is the support for indexing and other runtime needs.

The DAML Ontology Library²⁴ is a classic implementation of an ontology library. The ontologies can be accessed via different categories, such as meta-data describing the ontologies such as keywords, Open Directory category²⁵ and submitting organization. Also information obtained from the ontologies such as names of classes and properties can be used for finding relevant ontologies. The main method for using the ontologies is to download them. In comparison, ONKI provides application support for e.g. adding ontologies as mash-up and web services to applications.

The Ontology Library Service ONKI provides methods for finding ontologies and concepts amongst the ontologies published in the centralized service, whereas Swoogle [19] and Watson [20] act as global Semantic Web search engines. They crawl the web and index the RDF files they find. Such search engines can be useful when searching for suitable ontologies to use in applications, providing an overview of ontologies of some domain published on the web. The ONKI Service is based on a different approach. It aims to be a community-based service that gathers together the users of ontologies providing them ready-to-use ontological functionalities which can be integrated into semantic applications.

Dameron et al. proposes that ontology services should be provided as Ontology Web Services (OWS) which could be used in applications for automatically find and use ontologies [12]. In ONKI we support the idea of providing application interfaces to the ontology library, but extend the idea to a higher abstraction level by providing also ready-to-use user interface components to avoid duplicated work by re-implementing user interface and visualization functionalities.

Faviki²⁶ is a semantic bookmarking service which uses Wikipedia's²⁷ term identifiers for tagging web pages. In comparison, ONKI is focusing on publishing ontologies and to support the creation of ontology-based indexing, content search and other applications.

Freebase²⁸ is a data repository on the web that aggregates information from many sources and provides a single topic and identifier for each logical entity, e.g. a person. One goal of ONKI is also to provide (optimally) single, shared identifiers for ontological concepts which can be used to aggregate distributed content repositories. Freebase is based on a bottom-up approach based on existing information e.g. in Wikipedia whereas ONKI ontologies are (typically) based on top-down analysis of a domain and its relevant concepts.

²⁴ <http://www.daml.org/ontologies/>

²⁵ <http://www.dmoz.org>

²⁶ <http://www.faviki.com>

²⁷ With the help of DBPedia, <http://dppedia.org>

²⁸ <http://www.freebase.com>

7 Discussion

This paper discussed the requirements of an ontology library system to support the different phases of an ontology life cycle and related user needs for creation, publishing, maintaining and using ontologies. The Finnish national ontology library service ONKI addresses all phases of the ontology life cycle and contributes especially in providing support for 1) collaborative ontology publishing, 2) content indexing, and 3) information searching. The ontology services can be used in external legacy and other applications as ready-to-use functionalities. The new idea here is to support mash-up usage of ontologies in a way similar to Google Maps and other similar services. Our approach of providing an integrable autocompletion widget for external systems is the same as in [21].

The ONKI system supports syntactically, structurally and semantically heterogeneous content. RDF-based content representations such as RDF Schema, SKOS and OWL can be easily published by the ONKI SKOS server. SKOS generators for especially thesauri presentation formats such as MARCXML, various database schemas and text files has been implemented. Also multilingual content is supported. ONKI has been built for and tested with real world data consisting of ontologies, well-known thesauri and registries. The geographical ontology SUO contains over 800,000 places in Finland, which is published using the ONKI Geo server. The typical size of ontologies and thesauri published using the ONKI SKOS server is tens of thousands concept, e.g., YSO contains 20,600 concepts. For testing the scalability of ONKI SKOS, the Wordnet with 230,000 concepts has been successfully presented with the system. The ONKI services have been tested in various ways: the ONKI Selector as part of the SAHA editor in creating content for e.g. HealthFinland and CultureSampo and by external parties in their indexing and search applications. The ONKI Browser has been tested by expert users from e.g. the National Library of Finland.

To conclude, this full scale national ontology library service ONKI is novel and has not been done before. There are currently thousands of individual users and hundreds of organizations from different domains testing the ONKI system and using it in pilot applications. The National Library's commitment to ONKI means that a substantial part of public and private organizations in Finland begin to use ONKI – and most promisingly also the KOKO ontology – for indexing and search, but also for publishing and accessing their ontologies and thesauri – and to join the Semantic Web.

Future work include continuing observing how the ontology development, publishing and using community continues to build up around ONKI. We are most interested in seeing what kind of new applications will emerge based on ONKI and how well the concept of interlinked ontologies works in practice. DBPedia could be an interesting repository to be published in ONKI. Research topics include developing further methods for supporting community-based ontology development, managing changes in ontologies and utilizing change history and change propagation in e.g. inference and searching.

References

1. Gruber, T.R.: A translation approach to portable ontology specifications. *Knowledge Acquisition* 5(2), 199–220 (1993)
2. Staab, S., Studer, R.: *Handbook on ontologies*. Springer, Heidelberg (2004)
3. Fensel, D.: *Ontologies: Silver bullet for knowledge management and electronic commerce*, 2nd edn. Springer, Heidelberg (2004)
4. Reynolds, D., Shabajee, P., Cayzer, S.: *Semantic Information Portals*. In: *Proceedings of the 13th International World Wide Web Conference on Alternate track papers & posters*. ACM Press, New York (2004)
5. Hyvönen, E., Mäkelä, E., Salminen, M., Valo, A., Viljanen, K., Saarela, S., Junnila, M., Kettula, S.: *MuseumFinland—Finnish museums on the semantic web*. *Journal of Web Semantics* 3(2), 224–241 (2005)
6. Hyvönen, E., Viljanen, K., Tuominen, J., Seppälä, K.: *Building a national semantic web ontology and ontology service infrastructure—the finnonto approach*. In: *Bechhofer, S., Hauswirth, M., Hoffmann, J., Koubarakis, M. (eds.) ESWC 2008*. LNCS, vol. 5021, pp. 95–109. Springer, Heidelberg (2008)
7. Ding, Y., Fensel, D.: *Ontology library systems: The key for successful ontology reuse*. In: *Cruz, I.F., Decker, S., Euzenat, J., McGuinness, D.L. (eds.) Proceedings of SWWS 2001, The first Semantic Web Working Symposium, Stanford University, California, USA, July 30 - August 1, 2001*, pp. 93–112 (2001)
8. Ahmad, M.N., Colomb, R.M.: *Managing ontologies: a comparative study of ontology servers*. In: *ADC 2007: Proceedings of the eighteenth conference on Australasian database*, pp. 13–22. Australian Computer Society, Inc., Australia (2007)
9. Viljanen, K., Tuominen, J., Hyvönen, E.: *Publishing and using ontologies as mash-up services*. In: *Bechhofer, S., Hauswirth, M., Hoffmann, J., Koubarakis, M. (eds.) ESWC 2008*. LNCS, vol. 5021. Springer, Heidelberg (2008)
10. Euzenat, J., Shvaiko, P.: *Ontology matching*. Springer, Heidelberg (2007)
11. Mäkelä, E., Viljanen, K., Alm, O., Tuominen, J., Valkeapää, O., Kauppinen, T., Kurki, J., Sinkkilä, R., Käänsälä, T., Lindroos, R., Suominen, O., Ruotsalo, T., Hyvänen, E.: *Enabling the semantic web with ready-to-use web widgets*. In: *Proceedings of the First Industrial Results of Semantic Technologies Workshop, ISWC 2007, November 11 (2007)*
12. Dameron, O., Noy, N.F., Knublauch, H., Musen, M.A.: *Accessing and manipulating ontologies using web services*. In: *McIlraith, S.A., Plexousakis, D., van Harmelen, F. (eds.) ISWC 2004*. LNCS, vol. 3298. Springer, Heidelberg (2004)
13. Vehviläinen, A., Hyvänen, E., Alm, O.: *A semi-automatic semantic annotation and authoring tool for a library help desk service*. In: *Emerging Technologies for Semantic Work Environments: Techniques, Methods, and Applications*. IGI Group, Hershey (2008)
14. Viljanen, K., Tuominen, J., Käänsälä, T., Hyvönen, E.: *Distributed semantic content creation and publication for cultural heritage legacy systems*. In: *Proceedings of the 2008 IEEE International Conference on Distributed Human-Machine Systems, Athens, Greece*. IEEE Press, Los Alamitos (2008)
15. Valkeapää, O., Hyvönen, E., Alm, O.: *A framework for ontology-based adaptable content creation on the semantic web*. *J. of Universal Computer Science* (2007)
16. Tuominen, J., Frosterus, M., Viljanen, K., Hyvönen, E.: *ONKI SKOS server for publishing and utilizing skos vocabularies and ontologies as services*. In: *Aroyo, L., et al. (eds.) ESWC 2009*. LNCS, vol. 5554, pp. 768–780. Springer, Heidelberg (2009)

17. Kauppinen, T., Henriksson, R., Sinkkilä, R., Lindroos, R., Väätäinen, J., Hyvönen, E.: Ontology-based disambiguation of spatiotemporal locations. In: Proceedings of the 1st international workshop on Identity and Reference on the Semantic Web (IRSW2008), Tenerife, Spain, June 1-5 (2008)
18. Kurki, J.: Finding people and organizations on the semantic web. In: AI and Machine Consciousness - Proceedings of the 13th Finnish Artificial Intelligence Conference STeP, August 20-22 (2008)
19. Ding, L., Finin, T., Joshi, A., Pan, R., Cost, R.S., Peng, Y., Reddivari, P., Doshi, V., Sachs, J.: Swoogle: a search and metadata engine for the semantic web. In: CIKM 2004: Proceedings of the thirteenth ACM international conference on Information and knowledge management, pp. 652–659. ACM Press, New York (2004)
20. d’Aquin, M., Baldassarre, C., Gridinoc, L., Sabou, M., Angeletou, S., Motta, E.: Watson: Supporting next generation semantic web applications. In: Proceedings of IADIS International Conference on WWW/Internet (ICWI 2007), Vila Real, Portugal (2007)
21. Hildebrand, M., van Ossenbruggen, J., Amin, A., Aroyo, L., Wielemaker, J., Hardman, L.: The design space of a configurable autocompletion component. Technical Report INS-E0708, CWI, Amsterdam (November 2007)