Combining Distributed Ontology Repositories into a Global Service

Kim Viljanen, Jouni Tuominen, Eetu Mäkelä, and Eero Hyvönen

Semantic Computing Research Group (SeCo)
Aalto University, School of Science, and University of Helsinki
http://www.seco.tkk.fi/, firstname.lastname@aalto.fi

Abstract. Ontologies and vocabularies are a key resource for creating interoperable metadata on the Semantic Web. To make finding and using ontologies easier, the idea of Ontology Repositories has been introduced with current implementations including e.g. the NCBO Bioportal, ONKI and Cupboard. There is a genuine need for different kinds of Ontology Repositories, each focusing on different kinds specific user-needs, different ontologies and different organizational requirements which cannot be addressed by a single general implementation. However, at the moment each Ontology Repository is a separate island with its own user interfaces and APIs. They also use varying ontology languages such as OWL, SKOS, and RDF Schema. Due to this, global search, browsing, and inference over the repositories is difficult and generally not done which means that, for example, finding and reusing existing ontologies becomes difficult. To address the problems, we have developed a loosely coupled Network of Ontology Repositories (NOR) architecture that makes the repositories globally interoperable while maintaining their unique functionalities and strengths. To participate in the network, each ontology repository is required to implement a shared API. As a proof-of-concept evaluation, we present three case implementations demonstrating different aspects of the NOR approach: 1) internal distributed architecture of ONKI, 2) global search of ONKI and NCBO Bioportal, 3) publishing non-ontological concept collections as NOR endpoints, demonstrated with the semantic portal CultureSampo and the metadata editor SAHA.

1 Introduction

Ontologies and Ontology Repositories have been considered to be a key resource for building a global infrastructure to enable the vision of the Semantic Web [1, 2]. Many Ontology Repository systems exist for publishing and sharing ontologies and vocabularies for content indexing, information retrieval, content integration, and other purposes, for example BioPortal [3], ONKI [4], Cupboard [5], and the forthcoming Open Ontology Repository [2]. In addition to Ontology Repositories there are a vast amount of other kinds of concept collections with various degrees of formality but which could also be relevant as identifiers for the Semantic Web. One example of such concept collection is the Wikipedia, which has been made compatible with the Semantic Web by the DBPedia project [6]. The Wikipedia

(DBPedia) page identifiers are used for example by the BBC for interlinking content [7]. Other examples of non-ontological concept collections include registries maintained by libraries such as books and people (e.g. ULAN), and other collections of identifiers such as locations (e.g. GeoNames) or categories and other content maintained in website content management systems. Therefore, in the following, we are considering such non-ontological concept collections as simple Ontology Repositories even though any formal ontological analysis and representation of the concepts and their relations has not been defined (yet).

Currently, Ontology Repositories can not be accessed in a uniform way which creates a hinder for the widespreading of using ontologies. As a solution to the problem, we propose a Network of Ontology Repositories (NOR) architecture that defines a unified way to access repositories globally. We define a normalized presentation for ontology concepts, a shared API, and shared practices for creating an interoperable network of Ontology Repositories. The approach has been implemented to combine in-use ontology repositories on the Semantic Web.

We restrict our focus on the Semantic Web and RDF compatible ontology languages but the ideas presented in this work can be of use also for developing Ontology Repositories for non-RDF ontology languages such as the Common Logic[1]. In addition, we address the problem of how to represent non-ontological concept collections as a first class citizens in the NOR network. The work is partially based on our previous work on the national ontology library ONKI [4, 8] and is related to the open ontology repository (OOR)[2] initiative which aims at developing an interoperability infrastructure for ontologies [2].

In the following, we first argue why a single Ontology Repository is not a viable solution for addressing all Ontology Repository needs. Then we present the proposed NOR architecture. After this, three implementations of the NOR approach are described. Finally, related work is discussed and contributions of the paper summarized.

2 One Size Does Not Fit All

There is a genuine need for different kinds of Ontology Repositories, each focusing on different kinds specific user-needs, different ontologies and different organizational requirements which cannot be addressed by a single general implementation. However, at the moment each Ontology Repository is a separate island with its own user interfaces and APIs. They also use varying ontology languages such as OWL, SKOS, and RDF Schema.

Due to this, global search, browsing, or inference over the repositories cannot be done, which creates a hindrance for using the available Ontology Repositories, since the user has to know in advance which repository addresses her needs. It is also not possible to extend existing general vocabularies with organization specific non-disclosed Ontology Repositories combining a public Ontology Repository service with an internal service. On the other hand, general search

[1] http://common-logic.org/
[2] http://ontolog.cim3.net/cgi-bin/wiki.pl?OpenOntologyRepository

engines such as Google[3] or general semantic search engines such as Swoogle[4] [9] are not focused on the specific Ontology Repository tasks such as finding the correct concept for annotation purposes, and may not index all relevant ontologies since they are not publicly available for business or other reasons.

Since the user may not find the correct ontology or concept for one's needs the following issues rise: 1) The quality of annotations may decrease if the optimal concept is not found. 2) Redundant new ontologies and concepts are created if the existing ontologies are not found. 3) Interlinking of data decreases due to creating redundant ontologies. 4) Merging data for Semantic Web applications becomes more difficult due to the need for ontology matching.

For the repository maintainers and ontology publishers the missing global access to Ontology Repositories means that: 1) High quality ontologies or high quality Ontology Repositories might not be used as much as they should since they are unknown to some of the users. 2) High quality non-ontological concept collections might not be considered to be used as identifiers because they are not published using standard Semantic Web formats. 3) Repositories do not benefit from ontologies available in other repositories. For example, (automatic) linking to relevant concepts in other ontologies could help the users to find the best ontologies and concepts for each need. 4) Ontology Repositories are not acting as model citizens of the Semantic Web, since their ontological content is not interlinked as much as it could be. 5) Same ontologies or concepts have to be maintained at different locations, which leads to redudant work.

Together the issues described above reflect the underlying problem of how to publish ontologies on the Semantic Web. There seems to be a lack of shared practices as discussed also in [2].

One solution for creating an interoperable network of ontology repositories would be to create a single application that would address all the needs of ontology publishers and users. The application could either be used for providing a single, centralized global service or installed distributedly by organizations that want to use and publish ontologies. However, we argue that restricting the ontology network to a single centralized or distributed Ontology Repository application is not a viable solution due to the following reasons:

Different ontologies and user needs require different functionalities. For example, the ONKI Ontology Repository supports different types of ontologies and implements different visualizations as depicted in Fig. 1. This is done to address the different needs of using e.g. general ontologies versus geographical ontologies. For example, the BioPortal has been designed originally to address the needs of the biomedical domain. Also different ontology languages require different technical implementations to maximize the benefits of the given formalism.

Due to license or other business reasons, some ontologies are not available as files but only as services. This means that such ontologies are not available to be uploaded to some general Ontology Repository, but only accessible via the given API or user-interface.

[3] http://www.google.com
[4] http://swoogle.umbc.edu

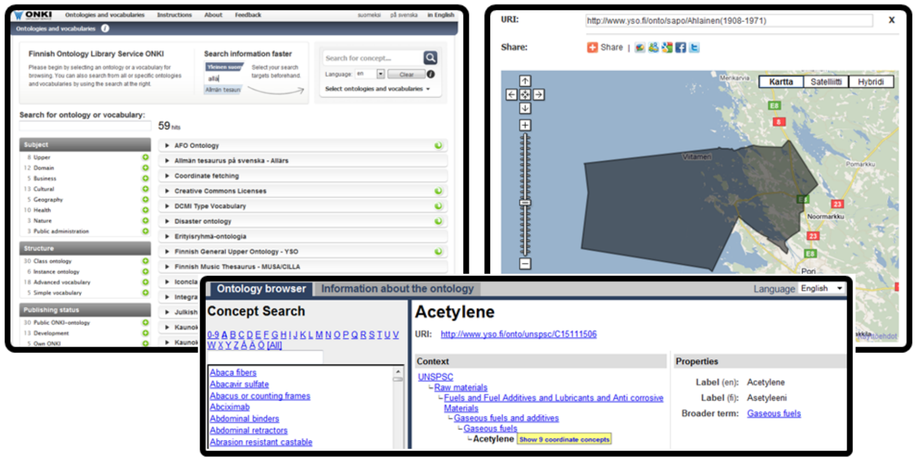

Fig. 1. Some of the user interfaces of the ONKI repository, including ontology listing, map visualization for geographical ontologies, and concept hierarchy visualization.

All concept collections are not ontology repositories. Many existing systems for publishing ontologies, thesauri or other kinds of concept registries exist, including e.g. content management systems such as Drupal has light-weight capabilities for maintaining and using thesauri and ontologies. Other examples include DBPedia [6], semantic metadata registries such as the SAHA [10] and CultureSampo [11]. A single ontology repository system most probably will not replace all of the different systems that are used for maintaining ontologies and vocabularies of various degrees of formality.

Security or other business reasons may require using internal ontology repositories. For example, when using ontologies for military purposes security reasons may require that selected ontologies are only available for internal use and that the server logs of who checked what ontological concepts remain confidental. Such requirements can be addressed with private, internal ontology repositories that are fully controlled by the organization itself. However, organizations using internal ontologies may (should) be using public ontologies for fostering interlinking when possible.

3 The Network of Ontology Repositories

We propose an architecture consisting of the following API: 1) A normalized presentation for ontology concepts, making thus the different ontology language schemas interoperable. 2) A normalized concept search for finding concepts from the ontology repositories in a uniform way. 3) Metadata about ontologies contained in an ontology repository.

Fig. 2 depicts an overview of the proposed architecture where multiple client applications can access the underlying heterogeneous Ontology Repositories via the common NOR API. In addition, the underlying ontology repositories can be accessed directly to benefit from the full power of the native ontology languages and repository specific functionalities. In the following, we discuss the details of the proposed architecture.

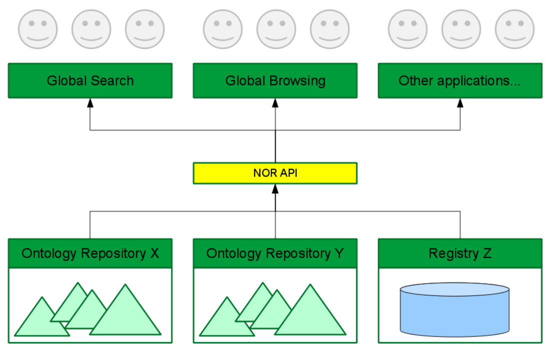

Fig. 2. The general architecture of the Network of Ontology Repositories.

3.1 A Normalized Representation of Ontological Concepts

Ontology repositories contain ontologies represented using different ontology languages such as OWL, RDFS, SKOS and – in the case of non-ontological concept collections – internal formats which have been chosen based on the modeling and inference needs. However, from the interoperability point of view, this creates a problem because the relations between different properties in different languages are not always known.

To avoid complicated mappings and inference of hierarchical and other relations, we propose that each ontology repository should provide a normalized, dumbed down presentation of the ontology concepts in addition to the native format of the ontology. As the normalization language we suggest using the RDF based Simple Knowledge Organization System (SKOS)[5], which is intended for presenting thesauri, classification schemes, subject heading systems and taxonomies within the framework of the Semantic Web.

Hiding ontological details makes it easier for the applications using the NOR for e.g. displaying the concepts originating from different Ontology Repositories

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

in a uniform way to the user. After finding an interesting concept, the user can be directed to the specific Ontology Repository with its full functionality for using the specific ontology. Our intention is to make it easier to access the basic information of ontological concepts in an unified way, not to restrict the user from using the original, full-blown ontology languages and functionalities of the underlying Ontology Repositories for specific needs.

In practice, each ontology repository participating in the NOR must provide a concept lookup method:

– *concept?uri=[concept identifier]*

The method returns the normalized SKOS version of the given concept, identified by the concept URI. For example, in the case of the ONKI Ontology Repository, to get the normalized concept representation of *yso:p907*, the lookup request URL looks like:

```
http://onki.fi/nor/concept?uri=http%3A%2F%2Fwww.yso.fi%2F
onto%2Fyso%2Fp907
```

which returns the following SKOS representation[6] of the given concept followed by the (optional) native representation:

```
# Namespace declarations (omitted)
# Normalized SKOS representation begins
<http://onki.fi/nor/concept?uri=http%3A%2F%2Fwww.yso.fi%2F
onto%2Fyso%2Fp907>
   a skos:Concept;
   skos:prefLabel "fish"@en, "kala"@fi;
   skos:broader
     <http://onki.fi/nor/concept?uri=http%3A%2F%2Fwww.yso.fi
%2Fonto%2Fyso%2Fp6580>;
   #...additional properties about the concept (omitted)
   nor:describes yso:p907  # link to the native concept format
.
# Native representation begins (optional)
yso:p907
   a ysometa:Concept;
   ysometa:prefLabel "fish"@en, "kala"@fi;
   #...additional properties (omitted)
.
```

The SKOS presentation above describes key information about the given concept (*yso:p907*) such as the labels (in English "fish", in Finnish "kala"), and the URL to the normalized broader concept *yso:p6580* (foods). In addition, the native representation *yso:p907* is also presented as part of the normalized concept lookup response.

To avoid cluttering the native presentation, that is, adding any RDF triplets to it, the native and normalized formats are kept apart from each other, by defining the following RDF property[7]:

[6] presented using the RDF Turtle syntax
[7] nor namespace: http://purl.org/finnonto/schema/nor

– *nor:describes*

for refering to the native version of the normalized SKOS representation. To avoid making unintended conclusions, we did not use for example the *owl:sameAs* property which would have meant that the normalized and the native presentations would refer to the same thing, which may not be true or in the intention of some Ontology Repository maintainers.

Finally, in some cases the Ontology Repository publisher has decided to use SKOS as the native representation for the concepts. If so, the normalized representation is at the same time also the native representation and therefore the *nor:describes* relation and the (duplicated) native representation are omitted.

3.2 Normalized Concept Search

In addition to the lookup of a single concept described above, an ontology repository participating in the NOR network must implement a search method for querying for concepts in the repository. The method is defined as follows:

– *search?q=[query]&l=[language]*

The search method is used for finding concepts matching the given query string and given language. The method returns a list of matching concepts. We propose using a JSON based response format, but other formats may be considered in the future.

For example, a search for "fish" to the ONKI Ontology Repository is done with the following URL:

```
http://onki.fi/nor/search?q=fish&l=en
```

The search returns the following result:


```
{"results" : [
  {"ontology-abbreviation" : "yso",
   "ontology-label" : "Finnish General Upper Ontology",
   "ontology-label-language" : "en",
   "ontology-uri" : "http://www.yso.fi/onto/yso",
   "concept-label" : "fish",
   "concept-label-language" : "en",
   "concept" : "http://www.yso.fi/onto/yso/p907",
   "normalized-concept" : "http://onki.fi/nor/concept?
   uri=http%3A%2F%2Fwww.yso.fi%2Fonto%2Fyso%2Fp907",
   "native-concept" : "http://www.yso.fi/onto/yso/p907"
  }
  ...
 ],
 "metadata" : {"containingHitsAmount" : 50, "moreHitsAmount" : 1467}
}
```

In the result, *concept* is the URI of the concept, *normalized-concept* is the URL of the normalized representation of the given concept, and *native-concept* is the URL to the native representation of the concept.

Later, additional parameters may be added to the search method, such as restricting the query to a specific ontology, to a specific part of an ontology or to a specific concept type.

We argue, that a simple HTTP API is easy to implement both for the Ontology Repository developers and for the developers that want to access the NOR compatible Ontology Repositories. In addition, a simple API is easy to implement even if the underlying Ontology Repository is not based on RDF but is e.g. a relation database based registry of people which would be highly relevant to publish as a NOR endpoint. Thus, compared to e.g. using the RDF query language SPARQL, the simple API approach makes it easier for both publishers and users to benefit of the NOR network.

This does not limit however the underlying Ontology Repositories from implementing in addition, for example, a SPARQL endpoint since a key idea behind NOR is that native functionalities of the underlying ontology repositories are available for users that need more functionalities than what the simple NOR API and normalized presentation can provide.

3.3 Normalized Ontology Repository Metadata

To find NOR compatible Ontology Repositories, a list of repositories that conform to the NOR principles would be helpful. However, to avoid the problems of centralized systems, we do not require Ontology Repositories to publish information about themselves to any specific registries.

From the practical point of view, we hope that lists of NOR compatible repositories emerge to spread the information to potential NOR users. For example, in Finland the Finnish Ontology Library Service ONKI[8] contains a ontology metadata registry of ontologies selected with a Finnish perspective for various domains.

To help finding suitable repositories for one's need, we suggest that the NOR Ontology Repositories should publish metadata about the ontologies that are available in the repository using the following method:

- *ontologies*

which returns metadata about the ontologies in the given repository.

Since ontologies on the Semantic web are typically presented as RDF and they may (should) be interlinked, we propose using the Vocabulary of Interlinked Datasets (voiD)[9] for describing the ontologies available in the ontology repositories.

[8] http://onki.fi/en/browser
[9] http://rdfs.org/ns/void#

Additional information about the ontology such as the title and description may be expressed using e.g. the Dublin Core metadata schema, the Ontology Metadata Vocabulary (OMV)[10] and the upcoming Catalogue Vocabulary (dcat)[11].

We also define the following RDF property for describing the NOR API base URL of a NOR compatible endpoint (repository):

– *nor:endpoint*

For example, in the case of ONKI Ontology Repository, the ontology metadata would be available at:

```
http://onki.fi/nor/ontologies
```

The ONKI Ontology Repository ontology metadata is presented as follows:

```
# Namespace declarations (omitted)
<http://onki.fi>
  dc:title "ONKI Ontology Repository"@en;
  nor:endpoint <http://onki.fi/nor/> .

<http://www.yso.fi/onto/yso>
  a void:Dataset ;
  dc:title "Finnish General Upper Ontology"@en ;
  dc:creator <http://www.yso.fi/onki-ns/onki/Finnonto> ;
  dc:license <http://creativecommons.org/licenses/by/3.0/> ;
  foaf:homepage <http://www.seco.tkk.fi/ontologies/yso> .
...
```

The metadata can be used for creating for example a catalogue of NOR compatible ontology repositories and concept collections.

4 Case Studies and Evaluation

To analyze the idea of NOR, we have implemented many proof-of-concept prototypes which will be presented and discussed in the following. The NOR approach generalizes and unifies experiences gained from these test applications.

4.1 NOR as an Internal Architecture: ONKI Ontology Repository

The ONKI SKOS ontology server [12] has been used for publishing over 70 ontologies in the Finnish Ontology Library Service ONKI [4], which has been running as a pilot service from September 2008. The system is in living lab use with over 10 000 unique human visitors monthly[12], and there are over 200 registered users of the APIs and widgets. Even though ONKI SKOS supports

[10] http://omv2.sourceforge.net
[11] http://www.w3.org/egov/wiki/Data_Catalog_Vocabulary
[12] Measured with Google Analytics.

especially vocabularies presented in SKOS, the server can be used for publishing ontologies presented in other languages too, such as OWL and RDF Schema. To access the different ontology servers, ONKI contains a front-end service that does metasearches to the ONKI SKOS back-ends using a HTTP API (see Fig. 3).

Searching for concepts using an ONKI SKOS server is done with its HTTP API method *search*, which returns concepts matching to the query string in a specified language. The *getFullPresentation* method returns all information about a given concept, such as the preferred and alternative labels, the transitive parent concept tree and related concepts. Independently of the language each ontology is presented in, each concept is always returned in a uniform SKOS inspired JSON format which describe the normalized basic information of the given concept.

Building on this underlying distributed architecture, three clients have been implemented. The ONKI3 Browser[13] is a global search and browsing user interface for accessing the ONKI SKOS back-end servers in a uniform way. For example, making a global query to all ontology servers can be done. Also, a directory listing of the ontologies in the ONKI Ontology Repository is provided based on the metadata about the published ontologies. The ONKI3 user interface was mostly implemented using PHP[14].

Another client is the JavaScript-based ONKI Selector widget [13] for adding ontological concept search to HTML forms which also is using the metasearch for finding the matching concepts. A third client is a simple URI resolver for dereferencing the end-user's ontology concept URI requests to a suitable representation provided via the ontology repository network, such as HTML or RDF.

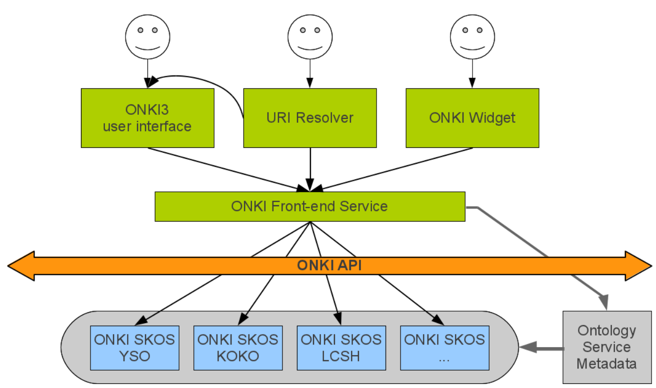

Fig. 3. The ONKI architecture.

[13] http://onki.fi/en/browser
[14] http://www.php.net/

Although the ONKI API is not directly compatible with the NOR API, both APIs are based on the same idea – they provide simple methods for searching concepts and getting relevant information about them in a uniform way. This makes it easy to implement a user-interface for accessing all the underlying Ontology Repositories.

The loosely coupled ONKI architecture has turned out to be a flexible and modularized approach for implementing an Ontology Repository consisting of multiple back-end ontology servers. Making multiple HTTP requests to back-end servers may be slow in the worst case, but in our test implementation this lag has not been a problem.

4.2 NOR Network Search: BioPortal and ONKI Implementation

To test the NOR approach in a distributed setting of multiple independent ontology repositories, we implemented a proof-of-concept prototype which provides a global metasearch to the ONKI SKOS [12] servers and the NCBO BioPortal [3]. This allows the user to find the relevant concepts from all participating ontology repositories, without having to know in advance which repository to make the search to.

NCBO BioPortal is an open repository of biomedical ontologies which has been used for publishing over 200 ontologies[15] [3]. It provides functionalities such as concept and ontology search and browsing, peer reviewing the ontologies, and support for creating and viewing mappings between ontologies.

The ONKI SKOS and BioPortal provide APIs for accessing the ontologies, but the APIs act differently and return different kinds of responses. BioPortal provides a HTTP REST API[16], which we used for concept search.

Since the ONKI front-end [4] was already designed to function as a front-end for back-ends which are (mostly) ONKI SKOS servers, the prototype of a Ontology Repository Network was implemented by creating a wrapper for BioPortal which implements the ONKI API's search and *getFullPresentation* methods. When calling the wrapper, it makes requests to BioPortal, parses BioPortal's XML message and transforms them to the ONKI JSON format. As, to our knowledge, an equivalent method to *getFullPresentation* of the ONKI API does not exist in the BioPortal, multiple HTTP REST requests have to be made to get all the needed information about a concept.

Fig. 4 presents the ONKI user interface displaying the result of an example metasearch query for "fish product" to the BioPortal and ONKI SKOS back-ends. The displayed result contains in total 22 hits which are found from the BioPortal and the ONKI SKOS back-ends. For demonstration purposes, the BioPortal hits are presented in the user inferace by the name "Bioportal" but for actual use, this should be replaced with the name of the respective ontologies.

As shown with the proof-of-concept implementation, the NOR approach is a working solution for a global search service to ontologies published in different ontology repository implementations, by using a common API (or wrappers).

[15] http://bioportal.bioontology.org

[16] http://www.bioontology.org/wiki/index.php/BioPortal_REST_services

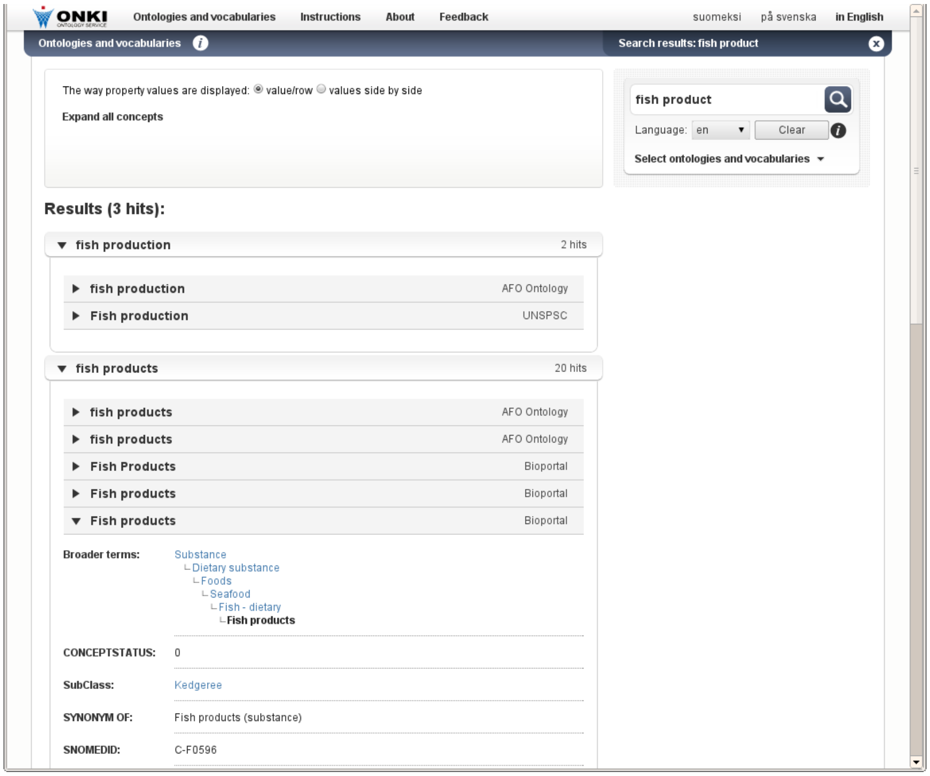

Fig. 4. The proof-of-concept implementation: results from ONKI SKOS and BioPortal presented using the ONKI3 metasearch interface.

4.3 NOR for Non-ontological Concept Collections

Besides Ontology Repositories, applications often need to refer also to non-ontological concept collections, such as authority or place databases. However, the functionalities required for such data sources are usually very similar to those required from Ontology Repositories. For example, in an editor environment, similar semantic autocompletion search functionalities are used for both ontological and non-ontological concept collections, along with the same functionality for describing and visualizing the possible choices returned from such a search. Non-ontological concept collections also often change more rapidly than their ontological counterparts, so it makes even more sense to access them centrally through programmatic APIs. In order to test how the NOR approach fared in the context of such non-ontological concept collections, the ONKI API was implemented in two applications: the semantic portal CultureSampo [14] and the SAHA metadata editor [10].

For CultureSampo, the ONKI API was actually implemented to benefit those using SAHA to edit data. This was because the CultureSampo database contains for example a large number of places, people and organizations that are useful to people indexing new content. For added freedom, the CultureSampo ONKI API was parametrized, so that the types of objects that search operations return can be specified dynamically. This way, one can for example say that they want an autocompletion facility for example of all the organizations, all the places or all the historical events in CultureSampo.

While SAHA was already a client to the ONKI API of the ONKI Ontology Repository and CultureSampo, the API was implemented also into SAHA itself. This was done to make possible the creation of a network of dynamically updated, collaboratively curated concept collections. Even currently, the multiple projects using the SAHA editor to index content often need to add new places, organizations or people to their list of reference values. However, until now, these have all resided in the private data spaces of the projects. Now, the intention is to move these created concepts by type into SAHA projects of their own, so that one SAHA project will hold collaboratively curated place database, while another contains a database of organizations and people. These can then be linked through the ONKI APIs to each other, as well as the primary indexing projects. In this way, the various projects can start to directly benefit each other.

5 Related Work

Compared to more general methods of accessing RDF data, such as SPARQL[17] and Linked Data [15], the NOR approach focuses on ontologies. For example, when querying for the concepts with the NOR search API, one does not need to know what RDF properties are used in the data to express the relevant labels. In addition, the ontology repositories can be optimized to respond quickly to specific API queries. A normalized presentation of ontological concepts (SKOS) could however also be beneficial for querying the data via SPARQL and browsing the ontology repositories as linked data. For example, one does not have to know which specific hierarchical relation (e.g. *rdfs:subClassOf* or *skos:broader*) has been used, because the normalized hierarchical relation is constant.

APIs for accessing ontologies and vocabularies presented previously include the SKOS API[18] and the OWL API[19]. Compared to them, the NOR approach provides a higher abstraction, independent from specific ontology languages, and a lightweight and simple API. Compared to the APIs of BioPortal [3], Swoogle[20] [9] and Watson[21] [16], the goal of NOR is to create a network of ontology servers based on a shared API that is implemented by all services.

[17] http://www.w3.org/TR/rdf-sparql-query/
[18] http://www.w3.org/2001/sw/Europe/reports/thes/skosapi.html
[19] http://owlapi.sourceforge.net/
[20] http://swoogle.umbc.edu/
[21] http://watson.kmi.open.ac.uk/

Therefore, the NOR API focuses on a few basic methods that reflects the basic functionality of ontology repositories, e.g. concept search.

Ontology Repositories such as BioPortal and Cupboard support publishing interlinked ontologies, but the ontologies have to be uploaded into a centralized service for a global search. On the other hand, the OOR[2] initiative intends to design an Ontology Repository framework that addresses the needs of all users and includes an inter-repository content change protocol to keep the different OOR repositories up to date. In contrast to these, the NOR approach does not restrict the ontology publishers regarding where to publish the ontologies or what software to use. Instead, the ontologies can be published using an ontology service that is optimized for the specific ontology and the user's needs. If the organization wants to promote and make their ontologies available to the NOR users, they can implement the NOR API to make their repository compatible with other NOR repositories. If needed, the NOR API of a repository can be restricted to selected users or made publicly available for anybody.

Compared to agent communications languages FIPA-ACL[22] and KQML[23], our approach focuses on accessing ontologies and other concept collections, whereas the scope of FIPA-ACL is more general, standardizing communication protocols of software agents and knowledge-based systems. In the context of FIPA-ACL, the FIPA Ontology Service supports ontology-based agent communication by providing services for discovering, maintaining, translating and matching ontologies. The Open Knowledge Base Connectivity (OKBC) specification[24] defines a general protocol for accessing knowledge bases with various knowledge representation systems. In contrast, the aim of the NOR is to provide a simple and focused API for finding ontologies, concepts and getting relevant information about them.

6 Discussion

This paper argues that Ontology Repositories should be made accessible using a shared API that would provide a simple but universal methods for accessing the ontology content in a uniform way. As a solution, we propose the NOR approach consisting of a normalized concept representation, normalized concept search and a normalized Ontology Repository metadata format.

The NOR approach has been evaluated with three case studies: The ONKI Ontology Repository case study demonstrates using the NOR approach for building an ontology service consisting of over 70 underlying back-ends with over 10 000 unique monthly users. The NCBO BioPortal and ONKI case study demonstrates using the NOR approach for creating a global search and browsing user-interface using the NOR approach for accessing independent distributed Ontology Repositories. Finally, the SAHA metadata editor and the CultureSampo

[22] http://www.fipa.org/repository/aclspecs.html
[23] http://www.cs.umbc.edu/csee/research/kqml/
[24] http://www.ai.sri.com/ okbc/spec.html

semantic portal case study demonstrates that the NOR approach can be used for accessing non-ontological concept collections.

The outcome of this work is that the NOR approach is feasible for providing a unified access to a multitude of Ontology Repositories which makes it possible to provide for example global search and global browsing functionalities to a collection of separate underlying Ontology Repositories. At the same time, the NOR does not restrict the individual Ontology Repository providers from creating advanced ontology, business and user specific implementations because the relation between the normalized representation and the native representation is kept intact. The NOR helps the ontology user to find relevant concepts and Ontology Repositories after which the user may access the underlying Ontology Repository for full-blown functionalities.

Future work includes doing additional evaluation on the approach, for example statistical analysis on how well the concepts returned from different NOR compatible Ontology Repositories match each other. In addition, the API and the normalized representations could be developed further by adding additional functionalities. The normalized concept representations would benefit from a Linked Data approach, where the normalized concept representation would contain links between ontologies originating from other repositories. An interesting research problem would be how to support automatical concept linking as part of the NOR system.

Acknowledgements This work is part of the National Semantic Web Ontology project in Finland[25] (FinnONTO, 2003-2012), funded mainly by the National Technology and Innovation Agency (Tekes) and a consortium of 38 organizations. We thank Osma Suominen and other members of the Semantic Computing Research Group and the OOR network for fruitful discussions.

References

1. Hyvönen, E., Viljanen, K., Tuominen, J., Seppälä, K.: Building a national semantic web ontology and ontology service infrastructure—the FinnONTO approach. In: Proceedings of the ESWC 2008, Tenerife, Spain, Springer–Verlag (2008)
2. Baclawski, K., Schneider, T.: The open ontology repository initiative: Requirements and research challenges. In: Proceedings of Workshop on Collaborative Construction, Management and Linking of Structured Knowledge at the ISWC 2009, Washington DC., USA (October 2009)
3. Noy, N.F., Shah, N.H., Whetzel, P.L., Dai, B., Dorf, M., Griffith, N., Jonquet, C., Rubin, D.L., Storey, M.A., Chute, C.G., Musen, M.A.: BioPortal: ontologies and integrated data resources at the click of a mouse. Nucleic Acids Research **37**(Web Server issue) (2009) 170–173
4. Viljanen, K., Tuominen, J., Hyvönen, E.: Ontology libraries for production use: The Finnish ontology library service ONKI. In: Proceedings of the ESWC 2009, Heraklion, Greece, Springer–Verlag (2009)

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

5. d'Aquin, M., Lewen, H.: Cupboard - a place to expose your ontologies to applications and the community. In: Proceedings of the ESWC 2009, Heraklion, Greece, Springer–Verlag (June 2009) 913–918
6. Bizer, C., Lehmann, J., Kobilarov, G., Auer, S., Becker, C., Cyganiak, R., Hellmann, S.: Dbpedia a crystallization point for the web of data. Journal of Web Semantics: Science, Services and Agents on the World Wide Web (7) (2009) 154165
7. Kobilarov, G., Scott, T., Raimond, Y., Oliver, S., Sizemore, C., Smethurst, M., Bizer, C., Lee, R.: Media meets semantic web – how the bbc uses dbpedia and linked data to make connections. In: Proceedings of the ESWC 2009, Heraklion, Greece, Springer–Verlag (2009)
8. Viljanen, K., Tuominen, J., Salonoja, M., Hyvönen, E.: Linked open ontology services. In: Workshop on Ontology Repositories and Editors for the Semantic Web (ORES 2010), the Extended Semantic Web Conference ESWC 2010, CEUR Workshop Proceedings, http://ceur-ws.org/ (June 2010)
9. Ding, L., Finin, T., Joshi, A., Pan, R., Cost, R.S., Peng, Y., Reddivari, P., Doshi, V.C., Sachs, J.: Swoogle: A Search and Metadata Engine for the Semantic Web. In: Proceedings of the Thirteenth ACM Conference on Information and Knowledge Management, ACM Press (November 2004)
10. Kurki, J., Hyvönen, E.: Collaborative metadata editor integrated with ontology services and faceted portals. In: Workshop on Ontology Repositories and Editors for the Semantic Web (ORES 2010), the Extended Semantic Web Conference ESWC 2010, Heraklion, Greece, CEUR Workshop Proceedings, http://ceur-ws.org/ (June 2010)
11. Hyvönen, E., Mäkelä, E., Kauppinen, T., Alm, O., Kurki, J., Ruotsalo, T., Seppälä, K., Takala, J., Puputti, K., Kuittinen, H., Viljanen, K., Tuominen, J., Palonen, T., Frosterus, M., Sinkkilä, R., Paakkarinen, P., Laitio, J., Nyberg, K.: Culturesampo – a national publication system of cultural heritage on the semantic web 2.0. In: Proceedings of the 6th European Semantic Web Conference (ESWC2009), Heraklion, Greece. (May 31 - June 4 2009) Springer-Verlag.
12. Tuominen, J., Frosterus, M., Viljanen, K., Hyvönen, E.: ONKI SKOS server for publishing and utilizing SKOS vocabularies and ontologies as services. In: Proceedings of the ESWC 2009, Heraklion, Greece, Springer–Verlag (2009)
13. Viljanen, K., Tuominen, J., Hyvönen, E.: Publishing and using ontologies as mashup services. In: Proceedings of the 4th Workshop on Scripting for the Semantic Web (SFSW2008), 5th European Semantic Web Conference 2008 (ESWC 2008). (June 1-5 2008)
14. Hyvönen, E., Mäkelä, E., Kauppinen, T., Alm, O., Kurki, J., Ruotsalo, T., Seppälä, K., Takala, J., Puputti, K., Kuittinen, H., Viljanen, K., Tuominen, J., Palonen, T., Frosterus, M., Sinkkilä, R., Paakkarinen, P., Laitio, J., Nyberg, K.: CultureSampo – Finnish culture on the semantic web 2.0. Thematic perspectives for the end-user. In: Proceedings, Museums and the Web 2009, Indianapolis, USA. (April 15-18 2009)
15. Bizer, C., Cyganiak, R., Heath, T.: How to publish linked data on the web. http://www4.wiwiss.fu-berlin.de/bizer/pub/LinkedDataTutorial/ (July 27 2007)
16. d'Aquin, M., Sabou, M., Motta, E., Angeletou, S., Gridinoc, L., Lopez, V., Zablith, F.: What can be done with the semantic web? an overview of watson-based applications. In: 5th Workshop on Semantic Web Applications and Perspectives, SWAP 2008. (2008)