

## Reuse of library thesaurus data as ontologies for the public sector

### Mikko Lappalainen

Research Library, National Library of Finland, Helsinki, Finland.  
mikko.lappalainen@helsinki.fi

### Matias Frosterus

Library Network Services, National Library of Finland, Helsinki, Finland.  
matias.frosterus@helsinki.fi

### Susanna Nykyri

Research Library, National Library of Finland, Helsinki, Finland.  
susanna.nykyri@helsinki.fi



Copyright © 2014 by Mikko Lappalainen, Matias Frosterus and Susanna Nykyri. This work is made available under the terms of the Creative Commons Attribution 3.0 Unported License:

<http://creativecommons.org/licenses/by/3.0/>

---

### Abstract:

*In the spring of 2013 the National Library of Finland took on a mission to build a national level ontology infrastructure across the public sector. Central to this endeavor is the open source ontology service Finto and the General Finnish Ontology YSO. The ontology is based on the General Finnish Thesaurus, originally developed in the 1980s in the National Library mainly for book indexing. Moving from index terms used in the thesaurus to ontological concepts of YSO has made it feasible to map the ontology to other resources. Similar transformations have been employed to thesauri from other domains. By linking the various ontologized thesauri together, metadata harmonization has been achievable across databases. YSO and Finto are currently in pilot use in several organizations. Today the National Library of Finland is adopting a new role in steering the development and future practices of indexing in the world of linked data.*

**Keywords:** Thesauri, Ontologies, Metadata, Content Management, Open Linked Data.

---

# Reuse of library thesaurus data as ontologies for the public sector

## 1 INTRODUCTION

With the pervasive presence of the Internet and its emphasis on free text search, it briefly seemed that the role of annotations would diminish. However, quite quickly it was noticed that due to the ever increasing amount of data, annotations are more essential than ever but in a new, machine readable format allowing for better automation to help human users to apprehend the data. Here libraries with their extensive expertise as curators of knowledge are an ideal candidate to answer the call.

The building of a semantic web infrastructure for Finland was originally initiated in the FinnONTO research project<sup>1</sup> (2003-2012) by the Aalto University and the University of Helsinki. The project realized several semantic portals, an ontology service, a number of tools for using ontologies in annotations and also transformed several key in-use thesauri from various domains into ontologies.

Based on the results of the FinnONTO project, the National Library of Finland began its current ONKI project<sup>2</sup> in collaboration with the Ministry of Finance and the Ministry of Education and Culture in 2013 with the aim of building a national-level ontology service and a set of vocabularies and ontologies for various indexing and interlinking needs. Finto ontology service<sup>3</sup> was launched in January 2014 providing a service for the publication and utilization of ontologies and thesauri alike. Furthermore, the project develops the General Finnish Ontology YSO<sup>4</sup> which aims to provide a general set of concepts common to all domains.

In comparison to the research focus of the FinnONTO project, ONKI project puts more emphasis on reliability, ease of use and the perspective of an actual information seeker. Furthermore, the underlying principle behind the ontologies has been shifted from trying to accurately reflect the world as it is into providing as useful a tool as possible for annotation and information retrieval. The shift in emphasis was necessary in that stability and usability became prime concerns for the service in order for it to be a reliable “building block” for other systems. The information retrieval focus was a result of attaining a more definite understanding on how to better meet the requirements of the actual end-user applications.

With the move from thesauri to ontologies and from terms to concepts, new possibilities of interoperability and reuse have opened up. Terms carry different meanings in different domains but by penetrating deeper into the actual concepts behind the terms and giving the concepts unique identifiers and defined relations to other concepts, it becomes feasible to map the various vocabularies from different domains together. A total of fifteen in-use thesauri from varied domains have been transformed into ontologies and mapped together

---

<sup>1</sup> See <http://www.seco.tkk.fi/projects/finnonto/>

<sup>2</sup> See <https://wiki.helsinki.fi/display/ONKI/>

<sup>3</sup> See <http://finto.fi/en/>

<sup>4</sup> See <http://finto.fi/yso/en/>

into a harmonized cloud of interlinked ontologies. Currently the interlinked ontology cloud and the Finto ontology service are in pilot use in multiple organizations in the public sector.

In the process of moving from thesauri to ontologies, our approach emphasizes the importance of maintaining a clear link to the annotations made with pre-existing vocabularies. Another key aspect is to distribute the ontology development into expert organizations that have been responsible for the original thesauri. Yet, instead of simply mapping together independently-designed ontologies, we wish to cultivate processes of co-operation and stronger interlinking guaranteeing the quality of the resulting hierarchies.

In this paper we first discuss the transformation of the widely used General Finnish Thesaurus YSA into General Finnish Ontology YSO. We also describe how the terms were explicated into precise concepts with machine-readable relations and a complete hierarchy and how this was a necessary step in order to allow for the interlinking of data across different organizations. In the third section we give an overview of the national ontology service that has been developed in the ONKI project and detail the linked ontology cloud. Finally, in the fourth section, we ponder on the new role that has opened for national libraries in the era of the linked data.

## 2 FROM FUZZY TERMS TO LINKED CONCEPTS

At the heart of the ONKI project and Finto service lies the General Finnish Ontology YSO, which is based on General Finnish Thesaurus YSA, originally developed in the 1980s in the National Library of Finland primarily to be used for string based indexing of printed library materials.

YSO was originally built up in the FinnONTO research project in a process where YSA index terms were transformed into concepts by defining the underlying units of thought behind them. The new concepts were given unique identifiers (URIs) to identify them independently of specific language terms. In order to make the ontology bilingual, General Swedish Thesaurus Allårs<sup>5</sup> was used to provide the Swedish labels for the concepts. The new concepts were also given English labels by translating the Finnish terms into English.

A new top-level ontology was founded on existing models like DOLCE and BFO<sup>6</sup> and all concepts were placed into a single tree-like hierarchical structure using standardized relationships between the concepts. The resulting ontology (YSO) was presented in a machine readable form so it could be easily used in various applications in accordance with the Linked Data principles.<sup>7</sup>

The initial “ontologization” of YSA was, however, executed in a manner that emphasized the research aspect at the expense of actual usability of the resulting ontology. Thus the first versions of YSO were criticized by the library sector as obscure and unusable, and YSO ever replacing YSA was considered unlikely. Users complained that the top-level of YSO was confusing and incoherent and that the new concepts that were created to complete the

---

<sup>5</sup> Allårs was already mapped to YSA.

<sup>6</sup> Masolo et al. 2003.

<sup>7</sup> See <http://www.w3.org/DesignIssues/LinkedData.html>

hierarchy were badly chosen. This resulted in losing the overall intuitiveness of the ontology. It appeared that in the process of making YSA machine readable, it was made machine readable only.

With the ONKI-project special attention has been paid in making YSO intuitive and user-friendly without losing the benefits of machine readable semantic data. Below we discuss the central challenges in achieving this goal.

## 2.1 Transforming a thesaurus into an ontology: key challenges related to content

YSA is a general thesaurus in Finnish with content covering all fields at a general level. Its primary purpose is to be used as a general tool in indexing in libraries, archives and museums in Finland, though in practice its use outside the library sector has been limited. The thesaurus is maintained by the National Library of Finland. YSA has been used for indexing Finnish publications since 1987 and it also has a Swedish translation, Allärs, which has been in use since 1996.

YSA has been developed and maintained primarily from the perspective of traditional library practices and needs, and therefore it is not suitable as such to serve the more heterogeneous users of the semantic web<sup>8</sup>, and their differing needs. Semantic web includes different kinds of data, publications, actors and discourses, which should all be linked together in order to be operable across different different databases and organizations.

In practice, the transfer from a thesaurus into an ontology means two major changes: firstly, the emphasis moves beyond the term level into concept level ('unit of thought'<sup>9</sup>), and, secondly, the hierarchical structure is complete and consistent. In thesauri, the terms do not necessarily represent isolated and well-defined concepts, but can contain underlying ambiguity. Furthermore, the relations between thesaurus terms do not necessarily explicitly differentiate between e.g. is-a and part-of relations. Thesaurus (as YSA) allows for fuzzy practices, if the hierarchical structure is not carefully designed and complete. The user reads the thesaurus from a certain perspective, and does not necessarily come to think of other ways of understanding the term, which ontology structure makes visible.

Especially since YSA is a general thesaurus, fuzziness was a common problem on corpus level. For instance, a preferred term "aura" has in YSA no broader or narrower terms, but as a thematic group it has both 48 Medicine<sup>10</sup> and 50 Folklore<sup>11</sup>. If the users notice the thematic groups and are familiar with indexing principles and practices, it may be obvious for them, that the term has two different meanings and it thus actually represents two different

---

<sup>8</sup> According to the W3C, "The Semantic Web is about two things. It is about common formats for integration and combination of data drawn from diverse sources, where on the original Web mainly concentrated on the interchange of documents. It is also about language for recording how the data relates to real world objects. That allows a person, or a machine, to start off in one database, and then move through an unending set of databases which are connected not by wires but by being about the same thing." See <http://www.w3.org/2001/sw/>

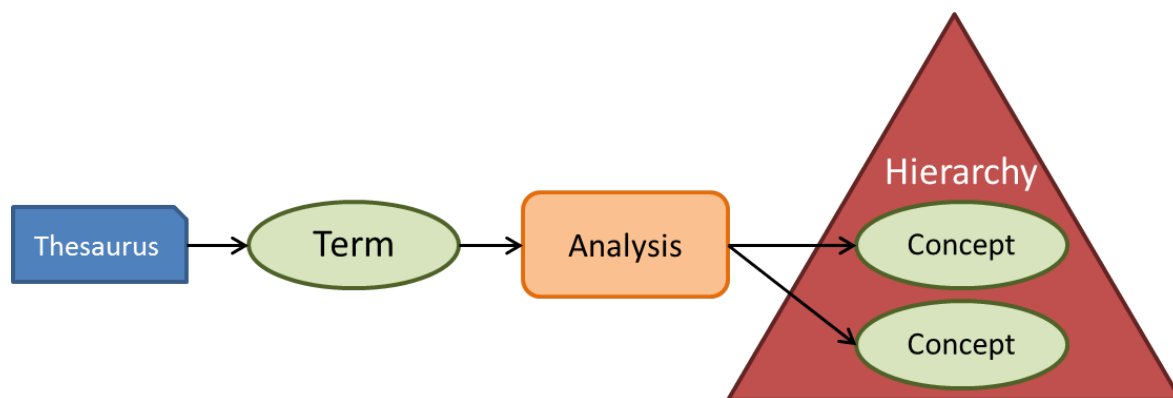
<sup>9</sup> According to ISO (ISO 5963-1985, 1) a concept is a unit of thought and "The semantic content of a concept can be re-expressed by a combination of other and different concepts, which may vary from one language or culture to another." (Ibid.) ISO -standard 25964-1:2011 (p. 3) clarifies that concepts exist in the mind as abstract entities independent of terms used to express them.

<sup>10</sup> [In Finnish: Lääketiede. Anatomia. Fysiologia. Patologia. Tautioppi. Psykiatria. Hammaslääketiede. Eläinlääketiede. Farmasia. Kauneudenhoito. Hoitomenetelmät]

<sup>11</sup> [In Finnish: Kansanperinne. Perinnetieteet. Kulttuuriantropologia. Okkultismi]

concepts. If they are familiar with the practices, the variety in search results does not surprise them, and/or they know how to narrow down the results by adding other relevant concepts to the search strategy.<sup>12</sup>

In the YSO ontology, the thesaurus term “aura” has been divided into two concepts and thus into two different term entries – “aura (paranormal phenomena) (en)” and “aura (symptoms) (en)” (the last-mentioned referring to perceptual disturbance typically in connection with migraine). (See below figure 1.) Thus they have different broader concepts, and both also include scope notes to specify their meaning in the ontology. This “one concept, one term” model allows indexers to make the conceptual difference during the indexing process allowing the the information seeker to avoid fuzziness and unwanted variety in search results.



**Figure 1:** The ontologization process

Ontologies, as well as thesauri, are controlled vocabularies. Standard ISO 5964-1985 and the more recent ISO 25964-1:2011 represents two principal means for achieving vocabulary control in thesauri: a) Concepts and terms are deliberately restricted in scope to selected meanings, and b) when the same concept can be expressed by two or more synonyms or quasi-synonyms in the same language, one of these terms is usually selected as a preferred term.<sup>13</sup>

Although the aim in FinnONTO-project was to operate on concept level, it was not fully achieved. Firstly, various terms in the YSA corpus were fuzzy in meaning, and secondly, the perspective of the YSO construction was not functional, but rather terminological. This led to solutions, where the central semantic components of concepts were not selected to conform to the actual use of YSO, which was to be used as a tool in indexing and information retrieval.

<sup>12</sup> Melinda, the Union Catalogue of Finnish Libraries, had ”aura” in four different meanings (thus representing four different concepts) in the first five records resulted in the subject search “aura”.

<sup>13</sup> “Terms are deliberately restricted in scope to selected meanings. Unlike the terms in a dictionary, which may be accompanied by a number of different definitions reflecting common usage, each term in a thesaurus is generally restricted to whichever single meaning best serves the needs of an indexing system. The structure of a thesaurus, notably its display of hierarchical relationships, frequently indicates the intended meaning of a term. If this technique is not sufficiently explicit, a definition or scope note should be appended to the term. This should state the chosen meaning, and may also indicate other meanings which are recognized in natural language but which have been deliberately excluded for indexing purposes; when the same concept can be expressed by two or more synonyms, one of these terms is usually selected as the preferred term (---) which is then used consistently in indexing. Reference to the preferred term should be made from any synonym which might also function as a user’s access point.” (ISO 5964,1985, 5)

In the FinnONTO-version of YSO terms were sometimes founded not based on actual usage (Literary Warrant) but rather with the aim of comprehensively covering all possible meanings the terms may embody. Because of this practice many concepts were divided into several concepts, and as a result YSO included also terms resulting no hits in common database searches (i.e. their existence in YSO is not based on actual indexing needs). In the current ONKI-project the same corpus has been considered more from the perspective of its function – to act as a tool in information retrieval and indexing – and therefore a step backward (towards its thesaurus model) has been taken. In practice it means, that polysemous terms do not automatically have separate term entries for each meaning but only for the ones needed in indexing. It also means that concept analysis pays more attention to the usability of YSO as a tool for indexing and information retrieval. The function of the ontology determines the perspective from which content analysis is carried out, ultimately determining how the concept is represented in the hierarchy.

The shift from thesauri to ontologies also means a shift into a more language independent understanding and representation of concepts. Its corpus, General Finnish Thesaurus YSA, is in Finnish, but it also has Swedish language version, Allårs, which has been constructed as a separate vocabulary.<sup>14</sup> The transfer into an ontology has made it easier to include equivalents in several languages, since the focus is now clearly on concepts, not on terms. In the corpus development, Finnish and Swedish have equal status, but English has a status of translation language. The English translations have made it feasible for YSO to be mapped with the Library of Congress Subject Headings (LCSH). In the first phase the Finnish concepts were given English labels (i.e. terms) and in the second phase equivalents from LCSH were detected.

Words and meanings may be translatable, but they may still represent or carry different worldviews and values in different cultures, and thus the meanings of words may be open for interpretation.<sup>15</sup> Although multilingual, YSO is still monocultural, and related to Finnish context and perspective, and reflects its society and culture (e.g. education system, idea of what is meant by ‘family’).

In addition, a central challenge in ontologization is time aspect. As new possibilities arose, the old practices needed to be taken care of, too. Since YSA has a long history in indexing, special attention had to be given to the indexing done with the YSA - it is absolutely vital that old material remains accessible and is not lost due to a new indexing tool, YSO.

## 2.2 Technical choices and structural details

One of the key tenets of linked data is machine-readability. In essence, this means that the properties linking the different resources have explicit semantics defining what extra relations a machine reasoner is allowed to deduce from the existing relations. Originally the ontologies were described in OWL ontology language which has later been changed to SKOS.<sup>16</sup> However, strict is-a type semantics have been retained for the hierarchical relations even though a more lenient skos:broader property has been employed to denote the hierarchies.

---

<sup>14</sup> Finland has two national languages, Finnish and Swedish. Finnish is the mother language of 90 % and Swedish of 5 % of Finnish population.

<sup>15</sup> Katan 2004, see more in Nykyri 2010

<sup>16</sup> See <http://www.w3.org/2001/sw/wiki/OWL> and <http://www.w3.org/2004/02/skos/>

This means that the hierarchy in YSO can be reliably used transitively as a subclass hierarchy.

The transitivity requires a well-thought and functional upper ontology. One of the first tasks of the ONKI project was to revise the top level of YSO so that it would work as intended. Taking in notion the criticism that YSO had had in the past, it was also considered important that while the upper ontology served machine readability and reasoning, it had to be also intuitive to humans. Since YSO functions as an upper ontology to other special domain ontologies (see section 3.2.), the uppermost classes had to be developed in a way that their usability would be guaranteed across different fields of knowledge.

A complete hierarchical structure also guarantees automatic query expansion and enables sophisticated methods<sup>17</sup> for improving precision and recall of queries. Since the focus of the ONKI project has been on building ontologies that can be effectively used in information retrieval, the hierarchical choices of YSO were considered crucial.

In the beginning of the ONKI project many available upper ontologies were scrutinized and also general metaphysics categorizations were studied in order to be able to evaluate the different choices made in different models. It was decided that a new top level had to fulfil three requirements: it had to be coherent all the way to leaf concepts so that automatic deduction would provide useful results; it had to be based on used and approved models so that the mapping of YSO to other resources would be easier; and lastly that the top level concepts needed to be defined in a way that not only machines, but also humans could understand them.

The new top-level of YSO was founded on the *Descriptive Ontology for Linguistic and Cognitive Engineering* (DOLCE)<sup>18</sup> mainly because its aim is not to “model the world” but to clarify the hidden ontological assumptions in existing lexical resources. The fact that DOLCE was already mapped to some major lexical resources like WordNet was also considered a plus.

Another important consideration is the use of permanent unique identifiers to denote the annotation concepts as opposed to term labels. YSO uses dereferenceable URIs that, when accessed with a browser serve the relevant concept page in Finto, and RDF data when a third-party application makes use of Finto’s machine interfaces. The URI scheme chosen does not contain human-readable information in order to keep them language independent and also to avoid pressure for changing an ID based on a change in the preferred label of a concept. For example, the URI of the concept of ‘cat’ in YSO is <http://www.yso.fi/onto/yso/p193> and therefore shows no preference to a given language nor is under pressure to change even if the preferred label for the concept of cat were to change.

### 3 ENABLING REUSE

In the previous section we detailed the transformation of a general thesaurus into a general ontology. However, in order to make use of this general ontology outside of the purview of the National Library, two things were needed. First of all, an ontology service allowing for

---

<sup>17</sup> See for example Agrawal et al. 2009.

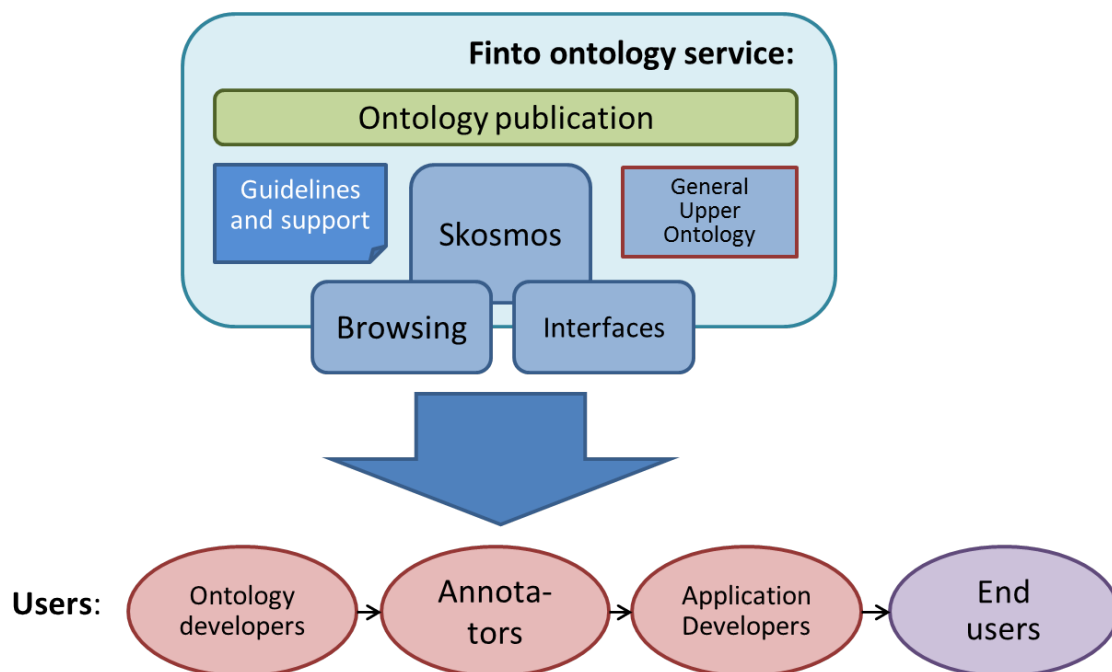
<sup>18</sup> Masolo et al. 2003, 13-41.

the publication and utilization of ontologies. The service needs to be stable and offer modern interfaces in order to make its integration into various applications as easy as possible. Second, the thesauri used in other organizations need to be harmonized and linked with the general ontology. This results in a cloud of interlinked ontologies defining the relations between the concepts used in annotations and thus allowing for searches across database boundaries.

### 3.1 Finto Thesaurus and Ontology Service

FinnONTO project originally realized an ontology service<sup>19</sup> meaning a centralized repository of ontologies which can be accessed using a web browser or various interfaces allowing for integration into other systems.<sup>20</sup> When SKOS was developed it was adopted for publication of light-weight ontologies and in 2012, a new and simplified version of the ontology service was built from the ground up working on top of SPARQL.<sup>21</sup> When the National Library of Finland began the ONKI project in 2013 this latest incarnation, was chosen as the basis for the new, production version of the service.

In order to differentiate clearly from the previous versions, which were meant for research use, the new service was named Finto, Finnish Thesaurus and Ontology Service. The underlying software, Skosmos, is being developed openly in GitHub.<sup>22</sup> Finto is a specific implementation of Skosmos - a thesaurus and vocabulary browser using SKOS and SPARQL - but it can be used to set up an ontology service anywhere.



**Figure 2:** Finto ontology service and its user groups

The components of the Finto service are depicted in Figure 2. The most important functionality is the publication of ontologies providing a centralized place for finding and

<sup>19</sup> See <http://onki.fi/>

<sup>20</sup> Viljanen et al. 2009

<sup>21</sup> Suominen et al. 2012

<sup>22</sup> See <https://github.com/natlibfi/skosmos>



using ontologies as well as common methods and interfaces for accessing and utilizing them. This makes integrating different ontologies into the same system as simple as possible making it feasible to, e.g., set up an annotation system where different fields are linked to different ontologies.

Aside from the technical platform, Finto also offers general guidelines and support for transforming thesauri into light-weight SKOS ontologies as well as for maintaining and updating them. The National Library of Finland also published the general Finnish ontology YSO, which contains annotation concepts needed in several domains.

Figure 1 also includes the user groups of Finto. First are the ontology developers who publish ontologies and can use the browser interface for examining other ontologies while working on their own. Second are the annotators who use Finto integrated into their annotation environment but who also have the possibility to browse the ontological hierarchies as needed. The third user group are application developers wishing to integrate ontologies and ontological annotation into their applications. Finally, end-users likely will not use Finto directly but rather through the applications made by the third user group. Note that the National Library of Finland includes users from all of these groups as Finto is in pilot use in-house.

### **3.2 The Linked Ontology Cloud KOKO**

In Section 2 we described how the thesaurus YSA was changed into the ontology YSO. Following the same principles, FinnONTO project also realized the transformation of fifteen specialized, in-use thesauri from various domains into domain ontologies. The transformation was performed in co-operation with the domain experts developing the underlying thesauri and the resulting ontologies were mapped into YSO. Thus YSO was used as a general upper ontology providing a shared hierarchy, which was then expanded by each of the domain ontologies. When all of these are combined, the end result is a linked ontology cloud, called KOKO.<sup>23</sup>

A central challenge was encountered in the differing amounts of resources available to different expert organizations maintaining the new domain ontologies. Since the domain expertise needed for the maintenance of the domain ontologies is high, the co-operation of the organizations responsible for the original thesauri was of paramount importance. This, however, resulted in widely differing update cycles for the ontologies.

If all the domain ontologies had been linked to one another, a change in one would affect potentially fourteen others. This was an untenable proposition due to the resource discrepancies. The solution was to use YSO as the central hub or glue relating all the domain ontologies to one another while minimizing the number of direct links between them. This way, the domain ontology developers needed to only concern themselves with changes to one ontology, YSO.

The work on this harmonized cloud of interlinked ontologies was begun during FinnONTO and has now been continued under the ONKI project. The ultimate goal is to use KOKO to relate the annotations in the datasets of the various organizations facilitating interoperability and breaking down silos.

---

<sup>23</sup> Frosterus et al. 2013

<b>Ontology</b>	<b>Domain</b>	<b>Number of concepts</b>
YSO	General upper ontology	24,800
MAO	Museum artifacts	6,800
MUSO	Music	1,000
TAO	Design	3,000
TERO	Health	6,500
VALO	Photography	2,000
AFO	Agriculture	7,000
JUHO	Government	6,300
KAUNO	Literature	5,000
KTO	Linguistics	900
KITO	Literary research	850
KULO	Cultural research	1,500
LIITO	Economics	3,000
MERO	Seafaring	1,300
PUHO	Military	2,000
TSR	Employment	5,100

**Table 1:** The component ontologies of KOKO

The different ontologies that comprise KOKO are shown in Table 1. As can be seen, they cover a wide range of different domains and they have all been built from in-use thesauri so each has, through the links to the thesauri, a vast amount of data that has already been used in annotation. The current version of KOKO is a harmonized global ontology of some 45,000 concepts aligned into a single hierarchy.

KOKO has been in pilot use as an annotation vocabulary in, e.g., various museums and in the National Broadcasting Company – i.e., in organizations that potentially deal with material from all possible domains.

#### **4 NEW ROLE FOR A NATIONAL LIBRARY: CONTROL GAINED AND CONTROL DIVIDED**

When the ONKI-project started in early 2013 it was evident that the work initiated in a technical research project could prove truly useful to a national library, despite of the strong opposition from library sector that had tempered the project in the past. This was, despite of the incompleteness, an opportunity that should not be lost.

The adopting of the project meant adopting a new role in the controlling of indexing at a national level. World had changed in the three decades that had passed since YSA was originally constructed in the National Library. Indexing in general had become something that was available for everyone: people tagged their own blog posts, hashtagged twitter updates and annotated their photos for social media. The ONKI project with its linked open data inclination is in a way a herald of this type of world and with new possibilities new responsibilities arise. The new role is a role on the much larger stage of linked data and indexing in general, way beyond the limits of library sector.<sup>24</sup>

In a way it was natural that a library, particularly a national library, should become the leader of a project that was going to build a service and tools for the indexing needs of the whole public sector. After all it was libraries that had the decades of experience in building thesauri and other controlled vocabularies and utilizing them in indexing. The central challenge was to find a way to combine the know-how of traditional book indexing to the best linked open data ideas and practices and thus take a role that was both time-honoured and new.

Although YSA had been in use outside the National Library, it was still essentially a closed system in the sense that there was no consideration over how YSA was used outside the library sector and how it should be interoperable with other indexing resources. With ONKI project, the responsibility grew on two levels: first YSO had to be built in a way that it was usable and reusable across the whole public sector, and secondly, new processes had to be developed in order to control the ways in which other organisations within the Finto network build their own ontologies.

The new network included new kinds of players who wanted to utilize already made ontologies available through Finto, but also to link their own existing indexing resources to others through the ontology cloud KOKO. These new players included museums, archives, media houses and governmental agencies, some with whom the National Library had had only little collaboration in the past. In order to guide the new kind of collaboration with new partners it was decided that the National Library should form a special interest group for guiding the indexing work of this large network.

As of spring 2014 the assembling of this group was still in the making and the final arrangements and procedures were still to be determined. The first principles have, however, been laid out: the intention is that in the future, the National Library would provide the service (Finto) for accessing and using the indexing resources, and that it would form the principles for building and maintaining of all of the interlinked ontologies. The objective was that the special expert organisations would still be responsible for the actual content of special domain vocabularies.

Through this project the National Library of Finland has, in a way, gained a much bigger control on indexing on a national level. This was achieved by adopting the practices of linked open data, which is founded on the notion of people and organizations working together in order to benefit each other. Although the National Library was the agent that built the service for the common use of ontologies and formed the principles through which the national ontologies were to be built, it was clear that in the final form of the new infrastructure the National Library of Finland would be just one entity among others. In essence the

---

<sup>24</sup> About the benefits of library data in linked open form see Baker et al. 2011.

organizational aspects of the ONKI project were about reformation and networking, first gathering the parties and then getting them to work together as equals.

Libraries' recognized expertise in content management and knowledge organization provide a natural basis for cross-domain interoperability and harmonization of metadata that the ONKI project aims to further. Ultimately this should lead to a more precise and comprehensive information retrieval for the end user.

## 5 REFERENCES

**Agrawal, Rakesh; Gollapudi, Sreenivas; Halverson, Alan; Jeong, Samuel 2009:** "Diversifying Search Results".

<http://research.microsoft.com/apps/pubs/default.aspx?id=73931>

**Baker, Thomas; Bermès, Emmanuelle; Coyle, Karen; Dunsire, Gordon; Isaac, Antoine; Murray, Peter; Panzer, Michael; Schneider, Jodi; Singer, Ross; Summers, Ed; Waites, William; Young, Jeff; Zeng, Marcia 2011:** "Library Linked Data Incubator Group Final Report. W3C Incubator Group Report 25 October 2011."

<http://www.w3.org/2005/Incubator/lld/XGR-lld/>

**Frosterus, Matias; Tuominen, Jouni; Pessala, Sini; Seppälä, Katri; and Hyvönen, Eero 2013:** "Linked Open Ontology Cloud KOKO--Managing a System of Cross-domain Lightweight Ontologies." *The Semantic Web: ESWC 2013 Satellite Events*, pp. 296-297, Springer-Verlag, Berlin Heidelberg, Montpellier, France, May 26-30, 2013.

ISO 5964-1985 = International Standards Organization (ISO). *Guidelines for the Establishment and Development of Multilingual Thesauri*. 1st ed. Geneva: ISO, 1985 (ISO 5964-1985-02-15).

ISO 25964-1:2011 = International Standards Organization (ISO). *Information and documentation -- Thesauri and interoperability with other vocabularies -- Part 1: Thesauri for information retrieval*. 1st ed. Geneva: ISO, 2011 (ISO 25964-1- 2011-08-05).

**Katan, David 2004 (1999):** *Translating Cultures: An introduction for Translators, Interpreters and Mediators*. 2nd. ed. Manchester, UK: St. Jemore Publishing.

**Masolo, Claudio; Borgo, Stefano; Gangemi, Aldo; Guarino, Nicola; Oltramari, Alessandro 2003:** "WonderWeb Deliverable D18. Ontology Library (final)."

**Nykyri, Susanna 2010:** *Equivalence and translation strategies in multilingual thesaurus construction*. Åbo Akademi University Press. <http://urn.fi/URN:ISBN:978-951-765-521-7>. 413 pp.

**Suominen, Osma; Johansson, Alex; Ylikotila, Henri; Tuominen, Jouni; Hyvönen, Eero 2012:** "Vocabulary Services Based on SPARQL Endpoints: ONKI Light on SPARQL." Poster proceedings of the 18th International Conference on Knowledge Engineering and Knowledge Management (EKAW 2012), Galway, Ireland, October, 2012.

**Tuominen, Jouni; Frosterus, Matias; Viljanen, Kim; Hyvönen, Eero 2009:** “ONKI SKOS Server for Publishing and Utilizing SKOS Vocabularies and Ontologies as Services.” Proceedings of the 6th European Semantic Web Conference (ESWC 2009), Heraklion, Greece, May 31 - June 4, 2009. Springer-Verlag.

**Viljanen, Kim; Tuominen, Jouni; Hyvönen, Eero 2009:** “Ontology Libraries for Production Use: The Finnish Ontology Library Service ONKI.” Proceedings of the 6th European Semantic Web Conference (ESWC 2009), Heraklion, Greece, May 31 - June 4, 2009. Springer-Verlag.