Linked Open Ontology Cloud KOKO—Managing a System of Cross-Domain Lightweight Ontologies

Matias Frosterus*, Jouni Tuominen, Sini Pessala*, Katri Seppälä, and Eero Hyvönen

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

1 Introduction

The Linked Data movement has focused on building cross-domain interoperability by creating and using (typically) `owl:sameAs` mappings between the datasets in the Linked Data Cloud (LOD). However, when linking data, ontologies would allow for deeper interoperability. Because different ontologies have been used when annotating different datasets, we argue that the LOD cloud needs to be complemented by developing a lightweight "Linked Open Ontology Cloud" (LOO). Aligning the ontologies requires more refined techniques than mapping data instances for the LOD.

In a LOO, special care must be taken to ensure that the (subclass) reasoning remains valid when performed over mappings between ontology boundaries. Furthermore, when an ontology is changed, reasoning over ontologies mapped to it may become invalid and therefore coordination and collaboration is needed when developing the ontologies.

We have created a LOO cloud called KOKO made up of a single general upper ontology and 15 domain ontologies (e.g., health, cultural heritage, agriculture, government, defense; with more being integrated into the system) as part of the national FinnONTO project (2003–2012). The system is based on transforming a set of 15 legacy thesauri in use into lightweight SKOS ontologies and interlinking them with each other through the general upper ontology. The result is KOKO, a harmonized global ontology cloud of some 45,000 concepts aligned into a single hierarchy. From the end-user's viewpoint, KOKO cloud is seen and used like a single ontology without boundaries. For the developers, the cloud is seen as a collection of interlinked ontologies in order to easily divide responsibilities based on domain expertise. KOKO has been published and is in use as a service in the national ontology library system ONKI[1] [2].

2 A Model for Managing a Linked Open Ontology Cloud

Based on our work with KOKO, we created and propose the following model depicted in Fig. 1 for developing a LOO cloud in a coordinated and collaborative manner.[2] The model with related tools is in field testing at the moment.

* Also affiliated with The National Library of Finland, `http://www.nationallibrary.fi/`

[1] online browsing: `http://onki.fi/en/browser/overview/koko-beta`
download: `http://tinyurl.com/koko-download`

[2] This work is part of the FinnONTO and Linked Data Finland projects, funded mainly by Tekes.

P. Cimiano et al. (Eds.): ESWC 2013, LNCS 7955, pp. 296–297, 2013.
© Springer-Verlag Berlin Heidelberg 2013

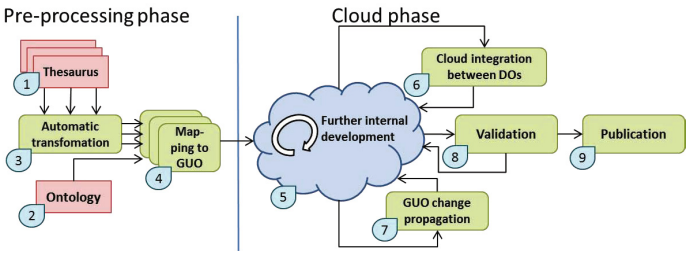

Fig. 1. The model of Linked Open Ontology Cloud formation and management

In the *Pre-processing Phase*, thesauri (**1**) and ontologies (**2**) are selected for building blocks of the ontology cloud. A thesaurus is converted into RDF format using a shared ontology schema (**3**) and aligned with a general upper ontology (GUO) (**4**). Aligning domain ontologies with a GUO forms a basis for interoperability by providing a complete hierarchy and is much easier to maintain than direct mappings between domain ontologies [1]. A similar approach was used by the IEEE Standard Upper Merged Ontology SUMO[3] working group. In our case, a natural basis for the GUO was the General Finnish Thesaurus YSA that was transformed into a General Finnish Ontology YSO[4].

The *Cloud Phase* begins after the domain ontologies have been aligned with the GUO (**5**). Since the alignments were made between domain ontologies and the GUO only, there may be mutually overlapping parts between the domain ontologies. To facilitate the integration of domain ontologies (DO) (**6**), processes and tools for discovering overlapping parts of the ontologies are needed. Based on the analysis, it is possible to eliminate redundant development work by deciding between domain ontology developers which ontologies maintain the overlapping parts.

Changes in the GUO create pressure for the domain ontologies to be updated accordingly to ensure the consistency of the cloud. For example, in our system, some 2,000 changes are made annually to the upper ontology YSO. These should be taken into account in the development process of the domain ontologies (**7**). Fortunately, only changes to concepts linked to the domain ontology via equivalency or subclass-of relations are likely to be relevant to the domain ontology. After editing the ontology, its logical consistency and other quality aspects should be validated (**8**) and fixed as necessary. Finally, the ontology cloud can be published as services for humans and machines, e.g., as user interfaces, APIs, and downloadable files (**9**).

References

1. Hameed, A., Preece, A., Sleeman, D.: Ontology reconciliation. In: Staab, S., Studer, R. (eds.) Handbook on Ontologies, pp. 231–250. Springer, Germany (2004)
2. 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)

[3] http://suo.ieee.org/SUO/SUMO/index.html
[4] http://www.seco.tkk.fi/ontologies/yso/