Extending the Linked Data Cloud with Multilingual Lexical Linked Data

Ernesto William De Luca

University of Applied Sciences Potsdam,
Friedrich-Ebert-Str. 4, 14467 Potsdam, Germany,
<deluca@fh-potsdam.de>

ABSTRACT: A lot of information that is already available on the Web, or retrieved from local information systems and social networks, is structured in data silos that are not semantically related. Semantic technologies make it apparent that the use of typed links that directly express their relations are an advantage for every application that can reuse the incorporated knowledge about the data. For this reason, data integration, through reengineering (e.g., triplify) or querying (e.g., D2R), is an important task in order to make information available for everyone. Thus, in order to build a semantic map of the data, we need knowledge about data items itself and the relation between heterogeneous data items. Here we present our work of providing Lexical Linked Data (LLD) through a meta-model that contains all the resources and gives the possibility to retrieve and navigate them from different perspectives. After giving the definition of Lexical Linked Data, we describe the existing datasets we collected and the new datasets we included. Here we describe their format and show some use cases where we link lexical data, and show how to reuse and inference semantic data derived from lexical data. Different lexical resources (MultiWordNet, EuroWordNet, MEMODATA Lexicon, the Hamburg Metaphor Database) are connected to each other towards an Integrated Vocabulary for LLD that we evaluate and present.

1.0 Introduction

With the advent of Linked Open Data (LOD, http://linkeddata.org/), more and more different resources are interconnected and shared on the Web. The idea of Linked Open Data is to connect and share data, information, or knowledge following semantic web principles like using URIs and RDF descriptions. While most linked data concentrate on linking facts, like music, movies, and geo- or demographic information, we believe that one important task is to connect language resources in order to support the process of language engineering. We also believe that natural language processing plays an important role in order to achieve this goal. Language engineering involves the development and application of software systems that perform tasks concerning the processing of human natural language (Cunningham 1999). Different tools have been designed, constructed, and are used for tasks like translation, language teaching, information extraction, and indexing.

Other, more intangible “language engineering tools” are language resources. Language resources are essential components of language engineering, containing a wide range of linguistic information with different degrees of complexity. These linguistic resources are sets of language data and descriptions in machine readable form, used for building, improving, or evaluating natural language and speech systems or algorithms. (Cole et al. 1997) give a brief overview about the various types of language resources, i.e., written and spoken language corpora, lexicons, and terminological databases.

1.1 Lexical resources

In the following, we concentrate on lexical resources that provide linguistic information about words. This information can be represented in very diverse data structures, from simple lists to complex repositories with many types of linguistic information and relations attached to each
1.2 Lexical linked data

Because the Web is evolving from a global information space of linked documents to one where both documents and data are linked, we agree that a set of best practices for publishing and connecting structured data on the web, known as linked data, is needed. The Linked Open Data (LOD) project (Bizer et al. 2009) is bootstrapping the Web of Data by converting it into RDF and publishing existing available “open datasets.” In addition, LOD datasets often contain natural language texts, which are important to link and explore data not only in a broad LOD cloud vision, but also in localized applications within large organizations that make use of linked data (Baldassarre et al. 2010; Nuzzolese et al. 2011).

The combination of natural language processing and semantic web techniques has become important in order to exploit lexical resources directly represented as linked data. One of the major examples is the WordNet RDF dataset (van Assem et al. 2006), which provides concepts (called synsets), each representing the sense of a set of synonymous words (Gangemi et al. 2003). It has a low level of concept linking, because synsets are linked mostly by means of taxonomic relations, while LOD data are mostly linked by means of domain relations, such as parts of things, ways of participating in events or socially interacting, topics of documents, temporal and spatial references, etc. (Nuzzolese et al. 2011). An example of interlinking lexical resources like Framenet (http://framenet.isi.berkeley.edu/) (Baker et al. 1998) to the LOD Cloud is given in Gangemi and Presutti (2010). They create a LOD dataset that provides new possibilities to the lexical grounding of semantic knowledge (Gangemi and Presutti 2010), and boosts the “lexical linked data” section of LOD, by linking, for example, FrameNet to other LOD datasets such as WordNet RDF (van Assem et al. 2006).

1.3 Knowledge organization and linked data

In recent years, the future of knowledge organization on the web has been presented and discussed with the special focus on linked data at ISKO UK 2010, among other venues. This was the most successful ISKO UK event to date. Several speakers referred to Tim Berners-Lee’s thoughts on linked data principles, which he first wrote about in July 2006. The idea is simple: you can add value to your information by linking it to that of others (http://www.w3.org/2006/02/14-rdf-schema/). Already at ISKO UK 2009, David Crystal (2009) gave a keynote on “Semantic targeting past, present and future,” the evolution of the linguistic approach to content analysis which he has been developing over the past 20 years. It began with the knowledge management taxonomy used for the Cambridge family of general encyclopedias and followed its transformation into an Internet taxonomy, with applications in automatic document classification, search engine assistance, ecommerce, online advertising, and Internet security. The recent developments focus on advertising, a field which has seen ideas develop from simple keyword analysis to contextual advertising and now to semantic targeting. These notions also included the ways of handling site sensitivity, sentiment, intention, and cultural localization. Other work has been done in this area in a multilingual context, comparing two indexing vocabularies for image retrieval (Ménard 2010) or tagging Medline abstracts semantically for enhancing information access (Ibekwe-SanJuan 2009). The exploitation of data in the cloud has been also addressed by another presentation given by Paul Miller in his talk “Exploiting data in the cloud,” where the increase of the quantity of data has been discussed and the need of standardization (e.g., through linked data) for a better structuring process has been recognized. Knowledge organization systems are needed in order to better access information and knowledge using the new technologies like RDF representations (Scott and Smethurst 2009), social tagging (Matthews et al. 2010), relational databases (Sarinder et al. 2010), or faceted classification approaches (Ménard 2010).

2.0 RDF/OWL/WordNet and RDF/OWLEuroWordNet

The WordNet Princeton (Fellbaum 1998) has already been converted into an OWL format as described in W3C (2004) using the OWL-DL sublanguage. This representation in RDF/OWL is based on the WordNet data model shown in Figure 1.

When we compare the original Princeton WordNet synset (having only word senses) with the OWL representation, we can see that the RDF/OWL schema (in its full
In order to avoid redundancy, only relations in one transitive direction (e.g., hyponymOf and not hypernymOf) are listed, the others can be retrieved with the owl:inverseOf property implemented in the RDF Schema (see the Figures 4 and 5). Altogether there are 27 relations implemented in the RDF/OWL representation of WordNet. The instances of all classes and properties are separated in several data files, one for the synsets, one for the WordSenses and Words, and one for each relation. Although the RDF Schema is used to describe most class and property definitions, there are several OWL statements integrated in the schema to provide better semantical descriptions, like checking the correctness of the data or defining inverse relations. For these statements software have to support the OWL DL standard in order to store and query the data.

3.0 RDF/OWL EuroWordNet

Because of the different problems related to WordNet and its variants, we decided to convert it into an RDF/OWL representation (see below), in order to enable the development of more flexible revision methods. In EuroWordNet, one synset contains all related word senses, synonyms and relations to other synsets and to the Inter-Lingual-Index. This information had to be prepared for inclusion in the appropriate RDF Schema and reorganized for a new data representation. The decision of converting EuroWordNet was also based on the need of extending it (because not all meanings are covered) with other resources. Furthermore, since most domain-specific ontologies are represented in OWL and a WordNet monolingual RDF/OWL representation has already been implemented, a EuroWordNet conversion would add multilingual capabilities to these resources. Therefore, we converted EuroWordNet into an RDF/OWL representation based on the work presented in van Assem et al. (2004). Since EuroWordNet has several relations and a structure that are different from the Princeton WordNet, several steps were required to adapt the data to the RDF/OWL Schema of WordNet and to extend this RDF Schema with the new relations. We first analyzed the requirements for EuroWordNet and adapted the WordNet RDF Schema to a multilingual representation of EuroWordNet. Then, we converted the EuroWordNet relations into OWL properties and extended the ontology with two domain ontologies (De Luca et al. 2007).

In the following, we describe the steps of this conversion and the problems that arose in more detail. An example on how to add ontologies is later given in Section 4.1 based on the OWL pizza and travel ontologies. The RDF/OWL-EuroWordNet Representation has also been extended with linguistic data included in the Hamburg Metaphor Database (HMD) in Section 4.
3.1 Conversion of EuroWordNet in RDF/OWL

The steps required to convert EuroWordNet in RDF/OWL can be subdivided into:

– Analysis of the requirements for EuroWordNet
– Adaptation of The WordNet RDF-Schema to EuroWordNet
– Multilinguality
– OWL Property Conversion
– OWL Domain Extension

van Assem et al. (2004) distinguish Word and WordSense in their datamodel for two reasons. First of all, several relations are defined for word senses and synsets, and WordNet uses this distinction in its database. Secondly, for the sake of ontological clarity, they assume that synsets include word senses, in order to partition the logical space of the lexicon (words as forms or meanings, and synsets as clusters of word senses by abstracting their distributional context). Agreeing with their model, we adapted their schema to convert EuroWordNet, applying this assumptions also for a multilingual task. An example of an OWL-EuroWordNet synset is given in Figure 2. Here the word sense “bank” is shown within its synset (and synsetId), WordSense, Word, and synonyms (containsWordSense).

Analyzing the structures of EuroWordNet and of the OWL representation of WordNet, we could recognize that some relations are supported in both versions (see Figure 4). Since EuroWordNet contains relations and properties that are not supported in the WordNet OWL representation, we had to adapt the RDF-OWL Schema to our needs in order to cover these gaps. Therefore, we created and stored new RDF structures, containing these new relations and thus extended the WordNet OWL implementation (see Figure 5). Some other properties that are covered in the WordNet OWL representation (e.g., the property tagCount used in the WordSense OWL declaration) are not available in EuroWordNet, so that we could not consider them.

Because we also tried to avoid redundancy as discussed in van Assem et al. (2004), we decided to delete the relations in EuroWordNet, which have an inverse form. We compared them updating their inverse relation property, where necessary (see the Figures 4 and 5). This means, if an instance of an inverse relation of a distinct instance is present, the instance can be deleted. If no inverse instance is available, the instance is added as inverse relation instance. An example is the hyperonym-hyponym relation resulting in a simple hyponymOf relation. Here the hyponym relation was available in both representations, but its name had to be changed from has hyperonym EuroWordNet format to the hyponymOf OWLWordNet format (see Figure 3). Another point to consider was that EuroWordNet also contains different relations belonging to the same “upper relation description” (e.g., ROLE AGENT, ROLE INSTRUMENT, ROLE LOCATION, etc. belonging to ROLE), because of their similar functionality. We decided in this case to merge them all into the same “Upper Relation” RDF file. A similar decision was done in the Princeton Conversion within the pertainsto relation. The complete mapping between EuroWordNet, OWL-WordNet and OWL-EuroWordNet relations is given in Figure 4. The relations that are only available in EuroWordNet and have been included as new RDF/OWL-EuroWordNet are shown in Figure 5.

![Figure 2. RDF/OWL-EuroWordNet synset Example.](https://doi.org/10.5771/0943-7444-2013-5-320)

![Figure 3. RDF/OWL-EuroWordnet hyponymOf Example for the word “bank”](https://doi.org/10.5771/0943-7444-2013-5-320)
Figure 4. EuroWordNet, RDF/OWL-EuroWordNet and RDF/OWL-WordNet relations. The crossed out relations have been removed, since they are inverse relations to the existing ones.
Since EuroWordNet is a multilingual resource (and not a monolingual like WordNet), we had to create for every language a unique set of files (containing the language-dependent synsets and relations). Every file of this set was additionally tagged with the name of the corresponding language (e.g., for English: eurowordnet-english-synset.rdf, eurowordnet-english-wordsensesandwords.rdf, eurowordnet-english-hyponymOf.rdf, etc.). We used the available EuroWordNet Inter-Lingual-Index that contains synsets, having the same identifier (synsetId) for all word meanings in all languages and an illustrative gloss. Because of the redundancy problem already described above, we decided to maintain only the gloss information included in the Inter-Lingual-Index deleting the word senses and synsetIds (already included in the english conversion). Therefore, the gloss entries were extracted and stored in a separate RDF file. Depending on this decision, the synsets of all languages are connected to another through the same identifier (synsetId) describing the same concept in different languages, instead of the Inter-Lingual-Index entries.

4.0 Interlinking RDF/OWL-EuroWordNet with other Lexical Resources

4.1 Interlinking RDF/OWL-EuroWordNet with the pizza.owl and travel.owl domain-ontologies

After a first conversion of EuroWordNet in an OWL representation, we decided to try to interlink it with two OWL ontologies (pizza.owl and travel.owl). For integrating them in our EuroWordnet OWL representation, we analyzed their hierarchy of classes that were also built in OWL DL. Every term is declared as a class (owl:Class), and every underlying term as a subclass (rdfs:subClassOf). There are additional restrictions, e.g., owl:disjointWith or owl:someValuesFrom statements. But there are no additional properties (except the OWL DL statements defined).

In order to extend EuroWordNet with these ontologies, we used a two steps approach:
1. Conversion from the OWL format to the EuroWordNet OWL format
2. Integration of the converted data in the EuroWordNet OWL hierarchy

First of all, we converted all classes (owl:class) into RDF/OWL synset classes (e.g., ewn20Schema:NounSynset), so that they were easier to add into the OWL/EuroWordNet hierarchy. Using some of the merging methods (De Luca and Nürnberger 2006a), we started a query for finding the synset we wanted to use as hyperonym of the domain ontology to be added (in this example “pizza”). We first explored the hierarchy with the LexiRes RDF/OWL tool. Then we tried to disambiguate the word senses of the searched word (“pizza”) in order to find the correct synset to be extended. After having found the correct synset, we manually merged the complete converted domain ontology under the appropriate hyperonym (synset). In this case, we could enlarge the EuroWordNet coverage with domain-specific terms. The extension can be easily recognized (and deleted if needed), because we added to the synsetId the name of the domain ontology followed by the number of word senses (see Figure 6). The pizza.owl ontology, for example, also has a language description for every class (e.g., xml:lang="en"), so that we could add it to the correct set of language files, if available. The same procedure was applied to the travel.owl ontology. A more detailed description of this extension work is given in De Luca et al. (2007).

4.2 Interlinking RDF/OWL-EuroWordNet with the Hamburg Metaphor Database

The RDF/OWL-EuroWordNet representation has also been interlinked with data included in the Hamburg Metaphor Database (HMD), a relational database of French and German corpus attestations containing metaphorical expressions (Lönneker-Rodman 2008). In the HMD, each metaphor is manually analyzed and annotated at several levels; among other lexical features, HMD provides references to EuroWordNet synsets. In addition, conceptual information is indicated in terms of domain labels from the Berkeley Master Metaphor List (Lakoff et al. 1991).

To provide an RDF/OWL representation of HMD data, we started by defining a new relation between the different synsets, the conceptual relation extMetaphorOf (“extension by metaphor of ... ”). This conceptual relation holds between a synset with a metaphorical meaning and a synset with a literal meaning of at least one of the contained word senses. The relation as such is defined by an RDF schema. We then populated the extMetaphorOf relation by deriving 107 instances from the HMD data for French. This was done by converting the data concerning attested metaphorical mappings between EWN synsets from the HMD relational database into RDF. The 107 instances of the extMetaphorOf-relation thus represent cases where both the literal and the metaphorical synset were already contained in the original version of EuroWordNet. As with each relation in RDF/OWL EuroWordNet, the resulting information is stored in a separate RDF-file (extMetaphorOf.rdf) and can be distributed as such. A detailed description about the integration of the Hamburg Metaphor Database into the RDF/OWL-EuroWordNet format can be found in De Luca and Lönneker-Rodman (2008).

4.3 Interlinking RDF/OWL-EuroWordNet with the Basic Multilingual Lexicon MEMODATA (BMD)

Because of the well known WordNet problem of conceptual coverage (De Luca and Nürnberger 2006b), we decided to interlink RDF/OWL-EuroWordNet with the Ba-

Figure 6. OWL-EuroWordnet (Merged) “vegetarian pizza” synset Example.
sic Multilingual Lexicon MEMODATA (BMD) (http://
catalog.elra.info/product info.php?products id=100). This
resource includes different multilingual concepts that can
be used for enriching the already available RDF/OWL-
EuroWordNet structure. First of all, before converting the
BMD, we analyzed the different structures of the resources
in order to find out similarities and differences. The BMD
contains words associated by the meaning in 5 languages:
English, French, German, Italian, Spanish. The lexical
categories included are: nouns (5 * 18 000), verbs (5 * 8
000), adjectives (5 * 6 000), and adverbs (5 * 1 500). Sixteen
parts of speech (POS) are distinguished and grammatical
information is also contained. This resource is divided into
five files, one for each language. Every file has different
lines, each of them representing one word. Each line (e.g.,
19223;E;Guyanesis; masc plur) includes

1. id number - this number links the word to the respec-
tive word represented in other languages
2. language code – it represents the language of the word
   (E=English, F=French, G=German, I=Italian, S=Spanish)
3. word - word or a word group
4. POS - part of speech or grammatical information (as,
   e.g., “s masc plur” for male plural nouns).

separated by semicolon. In addition to the five language
files, there is a meta file including the glosses of the words
(but only in French), each sorted at line level including the
id number, the language code A and one (or more) descrip-
tions, all separated by semicolon. The descriptions are
categorized by tags and can be used to disambiguate word
senses in combination with the complete database, The Integral
Dictionary (http://www.springerlink.com/content/4wc
0lb70m57km7eu/).

In order to interlink the BMD to RDF/OWL-Euro
WordNet, we analyzed the RDF/OWL-EuroWordNet
and BMD classes and relations and decided to merge
them with the same procedure we already applied in (De
Luca et al. 2007). We converted the BMD classes and
adapted them to the RDF/OWL WordNet classes (Laske
and De Luca, 2010). The sixteen parts of speech (POS)
of the BMD have been introduced as additional classes to
WordSense and SynSet. The properties that connect the
main classes (containsWordSense and word), the lexical-
Form —the string representation of a word—and the
synsetId have been extended: each synset contained in the
BMD having several POS, have been split into more syn-
sets with only one POS, like in WordNet. In order to be
sure that each new synsets has a unique synsetId, we al-
tered the synsetId from a number into a complex string
with the pattern:

%original synsetId in the BMD%-BMD-%part of
speech%

In this way, two synsets belonging to the same source syn-
setID, will differ in our extension by their part of speech
(POS). Some new relations that exist in the BMD had to
be added to the RDF/OWL EuroWordNet representa-
tion. For the grammatical POS information included in
some BMD entries, we added a grammaticalForm as a rela-
tion fromWord to xsd:string. Furthermore, the URIs of
the Words had to be changed from wn20instances: +
word- + %label% to bmd20instances: + word- + %la-
be% + (%grammatical form%). This avoided conflicts
that a Word could have more than one grammatical form
(e.g., transitive and intransive form for a verb).

Figure 7 shows the adaptation of the RDF/OWL
BMD classes that give a more precise description of the
Synsets (concepts) included in this resource. Here we ex-
tend the WordSense and SynSet classes with the additional
subclasses included in the RDF/OWL BMD. Figure 8
presents the comparison between the RDF/OWL-Euro
WordNet and the RDF/OWL BMD relations. The
crossed out relations have been removed, since they are
inverse relations to the existing ones. The gloss relation
has been extended (doing it more specific) due to the dif-
ferent glosses included in the BMD. The ‘grammatical-
Form’ and the ‘derivationallyRelated’ relation of a word
have been added to the already existing relations. Finally,
the interlinkage of BMD with RDF/OWL EuroWordNet
is useful for multilingual information retrieval or for ex-
fanding the coverage of the EuroWordNet synsets, as,
7.0 Conclusions

The web of data, with its canonical datasets (DBpedia,
geographical and biological data, social network data, bib-
ligraphical, musical, and multimedia data, etc.), and the
data emerging from the use of RDFa, Microformats, etc.,
has eventually provided an empirical basis to the semantic
web, and indirectly to knowledge engineering. In recent
years, on one hand, much attention has been paid to the
representation of lexical meaning and the development of
lexical-semantic resources on the other. Wealth of data
has implications. Firstly, a lot of data that are produced
have problems in their structure; secondly, large datasets
are difficult to describe in ways that enable their consump-
tion: what is typically described by those data? How are
data characteristically organized? Vocabularies do not help
much, since they provide a set of predicates and axioms,
Figure 7. Comparison of the RDF/OWL-EuroWordNet and the new RDF/OWL BMD classes.
### OWL EuroWordNet relations

- containsWordSense
- word
- lexicalForm
- synsetID
- tagCount
- gloss

### OWL Basic Multilingual Dictionary relations

- containsWordSense
- word
- lexicalForm
- grammaticalForm
- synsetID
- descriptionGloss
- domainGloss
- physicalGloss
- encyclopedicalGloss
- topicGloss
- glossHyperonymGloss
- coHyperonymOf
- characterisedBy
- grammaticalGloss
- derivationallyRelated

**Figure 8.** Comparison of the RDF/OWL-EuroWordNet and RDF/OWL BMD relations.
which is not tailored to the size and shape of data in the large; size and shape can only be empirically discovered.

We addressed these issues and showed how hybridization research can be done by porting different lexical resource to the LOD cloud (see Figure 9). Our contribution here is twofold: 1) the production and publishing of Lexical LOD datasets; and, 2) the description of a method to produce a common lexical linked data knowledge repository. In this article, we listed the already available datasets and how they can be linked to another through a metamodel we developed.

References

Baker, Collin F., Fillmore, Charles J. and Lowe, John B. 1998. The Berkeley FrameNet project. In COLING ’98 Proceedings of the 17th International Conference on Computational Linguistics. Stroudsburg, PA: Association for Computational Linguistics, pp. 86-90.

Baldassarre, Claudio, Daga, Enrico, Gangemi, Aldo, Gliozzo, Alfio Massimiliano, Salvati, Alberto and Troiani, Gianluca. 2010. Semantic scout: Making sense of organizational knowledge. In Cimiano, Philipp and Pinto, Helena Sofia, eds., EKAW. Lecture notes in computer science 6317. Springer, pp. 272-86.

Bizer, Christian, Heath, Tom and Berners-Lee, Tim. 2009. Linked data - The story so far. International journal on semantic web and information systems 5: 1-22.

Cole, Ronald A., Mariani, Joseph, Uszkoreit, Hans, Zae- nen, Annie and Zue, Victor. 1997. Survey of the state in human language technology. Pittsburgh, PA: Carnegie Mellon University.

Crystal, David. 2010. Semantic targeting: past, present, and future. Aslib proceedings 62. pp. 355-65.

Cunningham, Hamish. 1999. A definition and short history of language engineering. Natural Language Engineering 5: 1-16.

De Luca, Ernesto William. 2010. A corpus for evaluating semantic multilingual web retrieval systems: The sense folder corpus. In Calzolari, Nicoletta, Choukri, Khalid, Maegaard, Bente, Mariani, Joseph, Odijk, Jan, Piperidis, Stelios, Rosner, Mike and Tapias, Daniel, eds., Proceedings of the Seventh International Conference on Language Resources and Evaluation (LREC’10), Malta, Valletta, May 19-
De Luca, Ernesto William and Nürnberger, Andreas. 2008. Rebuilding lexical resources for information retrieval using sense folder detection and merging methods. In Proceedings of the 5th International Conference on Language Resources and Evaluation (LREC 2006), Genova, Italy. Paris: European Language Resources Association (ELRA), pp. 99-102.

De Luca, Ernesto William and Nürnberger, Andreas. 2006a. Rebuilding lexical resources for information retrieval using sense folder detection and merging methods. In Proceedings of the 5th International Conference on Language Resources and Evaluation (LREC 2006), Genova, Italy. Paris: European Language Resources Association (ELRA), pp. 99-102.

De Luca, Ernesto William and Nürnberger, Andreas. 2007. Converting EuroWordNet in OWL and extending it with domain ontologies. In: Proceedings of the Workshop on Lexical-Semantic and Ontological Resources. In conjunction with the GLDV-Frühjahrstagung (GLDV 2007), Tübingen, 2007.

Fallbaum, Christiane. 1998. WordNet, an electronic lexical database. Cambridge, MA: MIT Press.

Gangemi, Aldo and Presutti, Valentia. 2010. Towards a pattern language for the Semantic Web. Semantic Web Journal 1: 61-8.

Gangemi, Aldo, Navigli, Roberto and Velardi, Paola. 2003. The OntoWordNet project: Extension and axiomatization of conceptual relations in WordNet. In Meersman, Robert, Tari, Zahir and Schmidt, Douglas C., eds., Proceedings of the Sixth International Conference on Knowledge Resources and Evaluation (LREC 2004), Marrakech, Morocco, May 28-30. Paris: European Language Resources Association (ELRA), pp. 422-8.

De Luca, Ernesto William and Nürnberger, Andreas. 2006b. Using clustering methods to improve ontolgy-based query term disambiguation. International journal of intelligent systems 21: 693-709.

De Luca, Ernesto William, Eul, Martin and Nürnberger, Andreas. 2007. Converting EuroWordNet in OWL and extending it with domain ontologies. In: Proceedings of the Workshop on Lexical-Semantic and Ontological Resources. In conjunction with the GLDV-Frühjahrstagung (GLDV 2007), Tübingen, 2007.

Fellbaum, Christiane. 1998. WordNet, an electronic lexical database. Cambridge, MA: MIT Press.

Gangemi, Aldo and Presutti, Valentia. 2010. Towards a pattern language for the Semantic Web. Semantic Web Journal 1: 61-8.

Gangemi, Aldo, Navigli, Roberto and Velardi, Paola. 2003. The OntoWordNet project: Extension and axiomatization of conceptual relations in WordNet. In Meersman, Robert, Tari, Zahir and Schmidt, Douglas C., eds., Proceedings of the Sixth International Conference on Knowledge Resources and Evaluation (LREC 2004), Marrakech, Morocco, May 28-30. Paris: European Language Resources Association (ELRA), pp. 422-8.

Laske, Anja and De Luca, Ernesto William. 2010. Conversion of the basic multilingual dictionary into the RDF/OWL format. Technical Report. Magdeburg: Otto-von-Guericke University of Magdeburg.

Lönneker-Rodman, Birte. 2008. The hamburg metaphor database project: issues in resource creation. Language resources and evaluation 42: 293-318.

Matthews, Brian, Jones, Catherine, Puzon, Bartlomiej, Moon, Jim, Tudhope, Douglas, Golub, Koraljka and Nielsen, Marianne Lykke. 2010. An evaluation of enhancing social tagging with a knowledge organization system. Aslib proceedings 62: 447-65.

Ménard, Elaine. 2010. Ordinary image retrieval in a multilingual context: a comparison of two indexing vocabularies. Aslib proceedings 62: 476-88.

Nuzzolese, Andrea Giovanni, Gangemi, Aldo and Presutti, Valentia. 2011. Gathering lexical linked data and knowledge patterns from FrameNet. K-CAP '11 Proceedings of the Sixth International Conference on Knowledge Capture in cooperation with the AAAI, 25 - 29 June Banff, Alberta, Canada. New York: ACM, pp. 41-8.

Pianta, Emanuele, Bentivogli, Luisa and Girardi, Christian. 2002. MultiWordNet: Developing an aligned multilingual database. Proceedings of the 1st International WordNet Conference, January 21-25, Mysore, India, pp. 293-302. Available http://multiwordnet.fbk.eu/paper/MWN-India-published.pdf

Sarinder, K.K.S., Lim, L.H.S., Merican, A.F. and Dimyati, K. 2010. Biodiversity information retrieval across networked data sets. Aslib proceeding 62: 514-22.

Scott, Tom and Smethurst, Michael. 2009. Building coherence at bbc.co.uk. In Content Architecture, exploiting and managing diverse resources. Presented at Content Architecture: Exploiting and Managing Diverse Resources ISKO UK Conference, London 22-23 June.

van Assem, Mark, Gangemi, Aldo and Schreiber, Guus. 2006. RDF/OWL Representation of WordNet. W3C Working Draft 19 June 2006. Available http://www.w3.org/TR/2006/WD-wordnet-rdf-20060619/.

Vossen, Piek. 1999. EuroWordNet General Document, Version 3, Final. Available www.illc.uva.nl/EuroWordNet/docs/GenericDocPS.zip.

W3C. 2004. Wordnet in RDF and OWL: Editor’s Draft, Available http://www.w3.org/2001/sw/BestPractices/WNET/wordnet-sw-20040713.html