Semantic Web 0 (0) 1
IOS Press

lemonUby – a large, interlinked, syntactically-rich lexical resource for ontologies

Editor(s): Sebastian Hellmann, AKSW, University of Leipzig, Germany; Steven Moran, University of Zurich, Switzerland and University of Marburg, Germany; Martin Brümmer, AKSW, University of Leipzig, Germany
Solicited review(s): Tadej Štajner, Jožef Stefan Institute, Ljubljana, Slovenia;Vojtěch Svátek, University of Economics, Prague, Czech Republic; Fabian M. Suchanek, Télécom ParisTech University, Paris, France

Judith Eckle-Kohler, [a,*] John Philip M[c]Crae [b] and Christian Chiarcos [c]
[a] *Ubiquitous Knowledge Processing (UKP) Lab, Department of Computer Science, Technische Universität Darmstadt and German Institute for Educational Research (DIPF), Germany, http://www.ukp.tu-darmstadt.de*
[b] *Cognitive Interaction Technology (CITEC), Semantic Computing Group, Universität Bielefeld, Germany, http://www.sc.cit-ec.uni-bielefeld.de*
[c] *Applied Computational Linguistics (ACoLi), Department of Computer Science and Mathematics, Goethe-University Frankfurt am Main, Germany, http://acoli.cs.uni-frankfurt.de*

Abstract. We introduce *lemonUby*, a new lexical resource integrated in the Semantic Web which is the result of converting data extracted from the existing large-scale linked lexical resource UBY to the *lemon* lexicon model. The following data from UBY were converted: WordNet, FrameNet, VerbNet, English and German Wiktionary, the English and German entries of OmegaWiki, as well as links between pairs of these lexicons at the word sense level (links between VerbNet and FrameNet, VerbNet and WordNet, WordNet and FrameNet, WordNet and Wiktionary, WordNet and German OmegaWiki). We linked *lemonUby* to other lexical resources and linguistic terminology repositories in the Linguistic Linked Open Data cloud and outline possible applications of this new dataset.

Keywords: Lexicon model, lemon, UBY-LMF, UBY, OLiA, ISOcat, WordNet, VerbNet, FrameNet, Wiktionary, OmegaWiki

1. Introduction

Recently, the language resource community has begun to explore the opportunities offered by the Semantic Web, lead by the formation of the Linguistic Linked Open Data (LLOD) cloud and an increasing interest in making use of Linked Open Data principles in the context of Natural Language Processing (NLP) and Linguistics [7]. The use of RDF supports data integration and offers a large body of tools for accessing this data. Furthermore, the linked data approach gives rise to novel research questions in the context of language resources and their application.

For lexical resources, data integration has been in the focus of interest for many years, resulting in numerous mappings and linkings of lexica, as well as standards for representing lexical resources, such as the ISO 24613:2008 Lexical Markup Framework (LMF) [13]. In this context, the LLOD cloud can be considered as a new data integration platform, enabling linkings not only between lexical resources, but also between lexical resources and other language resources.

*Corresponding author, http://www.ukp.tu-darmstadt.de

1570-0844/0-1900/$27.50 © 0 – IOS Press and the authors. All rights reserved

We extend the LLOD cloud by a new lexical resource called *lemonUby*[1] which is the result of converting data extracted from the existing large-scale linked lexical resource UBY [14][2] to the *lemon* lexicon model. UBY has been developed independently from Semantic Web technology. It is LMF based and a subset of the LMF-compliant UBY lexicons is pairwise linked at the word sense level. The *lemon* lexicon model has been developed for lexical resource integration on the Semantic Web [19]. This lexicon model serves as a common interchange format for lexical resources on the Semantic Web and has been designed to represent and share lexical resources that are linked to ontologies, i.e., ontology lexica. Making use of a lexicon interchange format, such as *lemon* is not only important for data integration, but also for the reuse of lexicons.

While many lexical resources have already been included in the LLOD cloud, e.g., [3,21,22,18,10], the LLOD cloud is still missing a large-scale lexical resource rich in lexical information on verbs, including aspects such as syntactic behaviour and semantic roles of a verb's arguments. Such information is crucial for lexicalizing relational knowledge, e.g., the relation $like(Experiencer, Theme)$ can be lexicalized syntactically with a verb as in "NP likes NP".

The new resource *lemonUby* addresses this gap: Along with resources for word-level semantics (WordNet [12], English and German Wiktionary,[3] and the English and German entries of OmegaWiki,[4]) we converted two syntactically rich resources from UBY to the *lemon* format: FrameNet [2] and VerbNet [15]. For further data integration, we established links between *lemonUby* and other language resources in the LLOD cloud.

2. Representing lexical-semantic resources as Linked Data: *lemon*

There has been significant work towards integrating lexical resources using RDF and Semantic Web principles [6], and many resources are already available as Linked Data. Yet, representing lexical resources in RDF does not per se make them semantically interoperable. Consider, for instance, existing conversions of WordNet and FrameNet [27,24], where a simple mapping to RDF is provided, and augmented with OWL semantics so that reasoning could be applied to the structure of the resource. However, the formats chosen for the RDF versions of WordNet and FrameNet are specific to the underlying data models of WordNet and FrameNet. Although these lexicons are complementary resources [1], it is difficult (i) to link them in this form on the Semantic Web, and (ii) to use them as interchangeable modules in NLP applications.

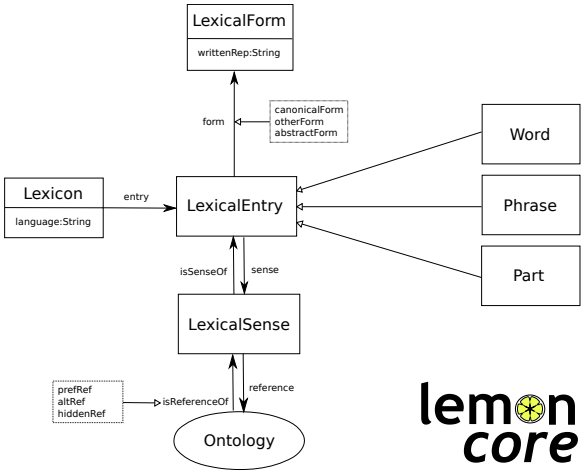

Fig. 1. The core of the *lemon* model

In order to overcome this difficulty, the *lemon* model [19] was proposed as a common interchange format for lexical resources on the Semantic Web. *lemon* has its historical roots in LMF and thus allows easy conversion from LMF-like, non-linked data resources. It links to data categories in annotation terminology repositories, and most of all, it realises a separation of lexicon and ontology layers, so that *lemon* lexica can be linked to existing ontologies in the linked data cloud.

This core model is illustrated in Fig. 1, which defines the basic elements used by all lexica published as linked data. In addition to this there are a number of modules used to model linguistic description, syntax, morphology and relationships between lexica.[5]

lemon has been used as a basis for integrating the data of the English Wiktionary with the RDF version of WordNet [18]. *lemon*'s similarity to the WordNet model made this conversion straight-forward, with only the need for a slight change in modelling to accommodate inflectional variants of lexical entries.

[1] http://www.lemon-model.net/lexica/uby/
[2] http://www.ukp.tu-darmstadt.de/uby/
[3] http://www.wiktionary.org
[4] http://www.omegawiki.org

[5] More detail of the model and descriptions of the modules can be found at http://lemon-model.net

3. Large-scale integration of lexical-semantic resources: UBY and UBY-LMF

UBY is both a network of interlinked lexical-semantic resources and a project on continuous integration and linking of lexical resources for NLP applications. It is motivated by the observation that an essential requirement in NLP is the availability of a wide range of lexical resources that can be used for many different NLP tasks. In a continuous process, such resources are integrated into UBY by means of (i) making them interoperable and (ii) linking them to other resources in UBY at the sense level.

In UBY, interoperability is achieved by standardizing lexical resources according to UBY-LMF [8,9], a lexicon model which is an instantiation of LMF, specifically designed for NLP. The lexicon model UBY-LMF has been developed to fully cover a wide range of heterogeneous lexical resources without information loss, which resulted in a fine-grained model of lexical information types (documented by data categories from ISOcat,[6] the implementation of the ISO 12620:2009 Data Category Registry) and was accompanied by an extension of the ISO standard LMF by a few elements. The extensibility of UBY-LMF was a primary design principle in order to enable the integration of further (in particular automatically acquired) lexical resources.

The mapping from UBY-LMF to *lemon* is motivated by an increase in interoperability with the Semantic Web and its resources, thereby making it available to a new group of potential users and novel applications. Beyond this, mapping UBY-LMF to *lemon* is an interesting task per se, because *lemon* links lexical resources and ontologies, whereas UBY-LMF is not related to any ontology. Another benefit is that LMF is not an open standard (in the sense that its specification is not freely available), while *lemon* fully complies with open data and open access principles.

4. Converting UBY lexica to *lemon*: *lemonUby*

To automatically convert data from UBY-LMF (extracted from UBY) to *lemon*, we performed a mapping of UBY-LMF elements to *lemon* concepts and properties. This involves two aspects: Formalizing and mapping UBY data structures in a Semantic-Web compliant way, and the actual conversion of UBY data.

[6] http://www.isocat.org/

lemonUby Resource	Triples	Links
WordNet	5,102,744	196,420
VerbNet	570,256	24,425
FrameNet	1,110,763	61,070
OmegaWiki English	6,173,515	113,930
OmegaWiki German	5,310,551	159,978
Wiktionary German	4,766,917	-
Wiktionary English	9,881,730	99,797
Total	32,916,476	327,810

Table 1
Number of triples and links for each resource.

4.1. Modeling Data Categories: ubyCat

In UBY-LMF, most of the terms used and many features (e.g., the feature "partOfSpeech" and the corresponding attribute values) are linked to community-maintained data categories in ISOcat.[7] As the mapping of UBY-LMF to *lemon* preserves this linking, *lemonUby* is linked to ISOcat as well. The content of ISOcat is also available as Linked Data [28], and therefore, provides a possible and direct way to interconnect *lemonUby* with other LLOD resources at the level of linguistic data categories.

However, ISOcat is not a formal ontology, but only a semistructured collection of terms, and while it serves as a repository of definitions, it does not provide a formal data model that can be applied to a resource: ISOcat contains doublets created by different data providers, and such superficially similar categories may actually have incompatible definitions, e.g., gerundive [DC-1294] is an "adjective formed from a verb" (excluding verbal nouns), whereas gerundive [DC-2243] is a "non-finite form (...) other than the infinitive" (including verbal nouns). Hierarchical relations between ISOcat terms are possible, but not obligatory, and when compared with a full-fledged ontology, ISOcat terms that represent superconcepts for a bundle of features do not distinguish relational and categorial aspects.

To render UBY/ISOcat data categories in a formal data model, we created the OWL/DL ontology *ubyCat*[8] which defines the semantics of linguistic terms used in *lemonUby*: With respect to *data structures*, it extends the *lemon* ontology,[9] concepts that are equivalent with *lemon* but have a different label are included, flagged as deprecated (and not used during the conver-

[7] http://www.isocat.org/rest/dcs/484
[8] http://purl.org/olia/ubyCat.owl
[9] http://www.monnet-project.eu/lemon

sion), but left in the ontology to help UBY-LMF users find the equivalent *lemon* categories.

With respect to *data categories*, *ubyCat* is linked to the Ontologies of Linguistic Annotations (OLiA) [4,5],[10] a modular architecture of OWL/DL ontologies that link resource-specific linguistic terminology (as for UBY) to an overarching 'Reference Model' which (a) provides a particular view on linguistic terminology, and (b) serves as an interface to multiple community-maintained terminology repositories such as ISOcat with which it is linked. Details of this linking are described in Sect. 5.1.

4.2. Converting the Data

The actual conversion of UBY data was achieved by means of an XML style sheet transform[11] that implements a mapping of the UBY-LMF model to the *lemon* model.

The following data from UBY were converted: WordNet, FrameNet, VerbNet, English and German Wiktionary, the English and German entries of OmegaWiki, as well as links between pairs of these lexicons at the word sense level:

- monolingual manual sense links for VerbNet–FrameNet,[12] VerbNet–WordNet,[13] and automatic sense links for WordNet–FrameNet [25,16], and WordNet–Wiktionary [23];
- cross-lingual automatic sense links between WordNet and German OmegaWiki [14], and manual inter-language links already given in OmegaWiki.

The resulting resource *lemonUby* has been published at `http://lemon-model.net/lexica/uby` under an open CC-BY-SA license. The choice of a share-alike license was due to UBY being published under the same license. Statistics for the *lemonUby* resources including the sense links within this dataset are given in table 1. An obvious gap is the German Wiktionary which is currently not linked to any other resource within *lemonUby*. Future work on increasing the link density within *lemonUby* should address this gap.

The XML style sheet mapping between the two lexicon models UBY-LMF and *lemon* revealed a number of differences between them. Most differences are due to the fact that *lemon* is a model for ontology lexica where the lexicon and ontology layers are kept separate. Thus, sense representations in *lemon* primarily consist of references to the associated ontology where a rich and domain-specific sense definition is provided. UBY-LMF, on the other hand, represents fine-grained sense information in the lexicon itself.

In line with previous authors [18], synsets in UBY-LMF (provided in WordNet and OmegaWiki) are considered to be SKOS concepts and referenced like ontology classes in *lemon*. Hypernym relationships stated in the lexicon are mapped directly to broader/narrower relationships in the ontology.

The "SemanticPredicate" class in UBY-LMF is used to represent semantic frames from FrameNet [9]; frames group senses which evoke the same kind of situation with participants taking over particular roles. Thus, senses in a FrameNet frame are semantically related, but not synonymous; e.g., the verbs *love* and *hate* are both in the same FrameNet frame. In *lemon* this relation is represented as subclass of the general lexical sense and linked to the rest of the lexicon by a *broader* (sense) relationship.

VerbNet verb classes and their hierarchical organization correspond to "Subcat[egorization]FrameSet" in UBY-LMF. VerbNet classes group verbs that share the same syntactic subcategorization frame, semantic roles, selectional restrictions, and semantic predicate. Although the resulting verb classes are semantically coherent, the semantic relatedness of verb senses in a VerbNet class is much more distant than in FrameNet, e.g., the verbs *believe*, *swear* and *doubt* are in the same class. Verb classes in *lemon* are represented by a broader sense as well, so *lemon* collapses the distinction between FrameNet-style and VerbNet-style sense groupings.

Translations to other human languages which are encoded as string-valued lexical items in the "Equivalent" class are mapped as a datatype property on the sense, in line with the UBY representation [17] rather than attempting to construct a cross-lingual linking for these terms as in the original resources (OmegaWiki and Wiktionary). Finally, provenance information is added to each resource at the root URL, including the source resource and version, time of creation and a link to the FOAF profile of the executor. Further examples

[10] `http://purl.org/olia`

[11] The XSL file can be downloaded at `https://raw.github.com/jmccrae/lemon.api/master/src/main/resources/xslt/ubylmf2lemon.xsl`

[12] `http://verbs.colorado.edu/semlink/`

[13] `http://verbs.colorado.edu/~mpalmer/projects/verbnet`

of usage and modelling are available from the website.[14]

5. Linking *lemonUby* to language resources

We linked *lemonUby* to other related resources in the LLOD cloud in order to make it accessible from existing datasets on the Semantic Web.

5.1. *Linking* ubyCat *with repositories of linguistic terminology*

We linked the data categories in *ubyCat* with the OLiA Reference Model, which provides formal definitions of annotation schemes for various linguistic phenomena in about 70 languages as OWL/DL ontologies, and which is further linked with other terminology repositories such as ISOcat and GOLD [11].

The *ubyCat* ontology plays an important role for the linking of UBY data categories to OLiA, because it defines grammatical concepts used in UBY-LMF in a formal data model. The linking between *ubyCat* and the OLiA Reference Model is implemented by subClassOf relationships between its concepts and the Reference Model. In OLiA, this mechanism is generally applied to physically separate resource-specific, interpretation-independent information (e.g., from an annotation scheme, hence the term 'OLiA Annotation Model') from resource-independent terminology (provided by the OLiA Reference Model) and its interpretation in terms of the latter ('Linking Model').

The separation of interpretation and interpretation-independent information by means of a declarative Linking Model[15] is necessary for reasons of transparency and reversibility, because *different interpretations* and thus, different linkings may be possible. Any interpretation requires to compare information provided by different communities (resource developer, terminology maintainer), it may be affected by differences in point of view or terminological traditions familiar to the ontology engineer. With the linking contained within a separate file, it is actually possible to provide alternative interpretations of the same Annotation Model.

Another advantage of an ontological formalization of linguistic terminology is that it is more expressive than a plain hierarchy or list of terms. An example for a complex linking is the verb form annotation 'ingForm', used for verb forms like *talking* in the resource. This category represents a language-specific merger of present participles (*he is speaking*, Old English *-inde*) and gerunds (*he began* [*the*] *speaking*, Old English *-inge*). The linking provides a language-neutral definition as PresentParticiple *or* Gerund, so that categories across different languages can be compared more easily.[16]

Through OLiA, grammatical information from *lemonUby* is interoperable with other LLOD resources linked to either OLiA or any of the terminology repositories it is linked with, including GOLD and ISOcat.

5.2. *Linking* lemonUby *to lexical resources*

As UBY is derived from existing lexical resources, the simplest links to create are those to other RDF versions of the resources that compose UBY. For WordNet, these links are simply created by mapping the data of UBY, which uses WordNet 3.0, to the linked data version of WordNet 3.0.[17] Here, we provided links at both the sense level and at the lexical entry level (lexical entries are "words" in WordNet 3.0). As can be expected, we found that this linking worked apart from 7 senses that did not map, which we believe is due to a bug in the WordNet API.

In addition, we provided links at the lexical entry level to two existing resources that are also widely used, i.e., RDF WordNet 2.0 [27] and an RDF instantiation of Wiktionary.[18] These links at the lexical entry level are created if two entries share the same lemma and part-of-speech information. We assume this linking is always correct given the lemma and part-of-speech are correct in the original resource.

For the RDF WordNet 2.0, we created such a simple linking at the lexical entry level, because a linking at the sense level based on the sense identifiers is not possible due to different WordNet versions using different sense identifiers.

[14] See, e.g., `http://lemon-model.net/lexica/uby/modelling.php` or `http://lemon-model.net/learn/learn.php`

[15] `http://purl.org/olia/ubyCat-link.rdf`

[16] It should be noted that such a complex linking requires the use of operators like *and* and *or* in the linking, to capture this information, OLiA ontologies employ OWL/DL. The direct mapping between annotations and reference concepts originally advocated for ISOcat and GOLD cannot represent this information.

[17] `http://semanticweb.cs.vu.nl/lod/wn30/`

[18] `http://wiktionary.dbpedia.org`

The linking of the Uby version of WordNet and the RDF export of Wiktionary at the lexical entry level is based on the assumption that corresponding word classes share the same part-of-speech category in both resources. While this is mostly true, we found a few word classes where we had to manually unify diverging part-of-speech categories; these word classes were initially not covered by the linking. For instance, for acronyms (e.g. IBM), Wiktionary assigns "Initialism" as a part-of-speech, whereas WordNet counts these as nouns. Statistics for all mappings including the coverage of the linkings are given in table 2.

6. Applications of *lemonUby*

As a web resource, *lemonUby* provides two main possible applications: firstly as a resource to allow better lexical description of ontology entities and secondly as a resource for NLP applications.

As a wide coverage resource containing lexical information, it would be possible to link existing ontologies and terminology vocabularies to *lemonUby* and then obtain rich information about the linguistic usage of a given term. As an example, consider the cross-lingual linking of verb senses across English and German available in *lemonUby*. Verb senses in the German OmegaWiki can be enriched by semantic role and selectional preference information from VerbNet and FrameNet via the cross-lingual linking between German OmegaWiki and WordNet, and via the WordNet–VerbNet and WordNet–FrameNet linking. Currently, no German lexica with these information types are freely available for research purposes.

As a resource for NLP applications, *lemonUby* can be easily deployed into existing applications [26]. The wide variety of lexical information types it offers, ranging from taxonomic relationships (e.g., hyponomy) and translations to fine-grained lexical-syntactic information, makes it an attractive resource for many different NLP tasks, such as Entity or Predicate Disambiguation, to name only two.

Recently, it has been shown that *lemonUby* is a valuable resource for machine translation [20]. By mining translations from *lemonUby* and combining them with the Moses statistical machine translation, the performance of the resulting machine translation system improved.

Furthermore, all the data from *lemonUby* is available via a SPARQL endpoint.[19] Thus, ad-hoc queries can be used to access and explore the resource, for example to find translational equivalents quickly and easily.

7. Conclusion and Outlook

We presented *lemonUby*, a new linked data resource which combines data from the standardized lexical resource UBY with the principled model, *lemon*, for representing lexical data on the web. This resource provides not only rich information about many lexical entries in two languages, it is also linked at the word sense level both within its component resources and to other language resources on the web.

We see two main directions for future research on *lemonUby*: first, the further linking of this dataset both within *lemonUby* and to other resources on the web, and second, the use of *lemonUby* in RDF-based NLP applications in order to gain a deeper understanding of the added value of linked lexical resources for NLP.

Acknowledgments

The work on *lemon* was developed in the context of the Monnet project, which is funded by the European Union FP7 program under grant number 248458, the CITEC excellence initiative funded by the DFG (Deutsche Forschungsgemeinschaft). The work on UBY has been supported by the Volkswagen Foundation as part of the Lichtenberg-Professorship Program under grant No. I/82806. The research of Christian Chiarcos has been partially supported by a German Academic Exchange Service (DAAD) funded PostDoc fellowship at the University of Southern California. We thank Silvana Hartmann, Sebastian Hellmann, Michael Matuschek, Christian M. Meyer, and Zijad Maksuti for their contributions to this work. Our special thanks go to the three reviewers Tadej Štajner, Vojtěch Svátek and Fabian M. Suchanek for their detailed and valuable feedback.

[19] `http://lemon-model.net/sparql.php`

lemonUby Resource	Target Resource	Links	Coverage of Uby	Coverage of target
WordNet	WordNet 3.0	206,773	99.9%	99.9%
WordNet	WordNet 2.0	84,416	40.8%	97.8%
WordNet	Wiktionary English	76,294	36.9%	18.4%

Table 2

Number of external links created between *lemonUby* and other resources in the LLOD cloud.

References

[1] Baker CF, Fellbaum C (2009) WordNet and FrameNet as Complementary Resources for Annotation. In: Proceedings of the Third Linguistic Annotation Workshop, Suntec, Singapore, Association for Computational Linguistics, Stroudsburg, PA, USA, pp 125–129

[2] Baker CF, Fillmore CJ, Lowe JB (1998) The Berkeley FrameNet project. In: Proceedings of the 36th Annual Meeting of the Association for Computational Linguistics and 17th International Conference on Computational Linguistics (COLING-ACL'98), Montreal, Canada, Association for Computational Linguistics, Stroudsburg, PA, USA, pp 86–90

[3] Bizer C, Lehmann J, Kobilarov G, Auer S, Becker C, Cyganiak R, Hellmann S (2009) DBpedia – A Crystallization Point for the Web of Data. Journal of Web Semantics: Science, Services and Agents on the World Wide Web 7(3):154–165

[4] Chiarcos C (2008) An ontology of linguistic annotations. LDV Forum 23(1):1–16

[5] Chiarcos C (2012) Ontologies of linguistic annotation: Survey and perspectives. In: Calzolari N, Choukri K, Declerck T, Doğan MU, Maegaard B, Mariani J, Moreno A, Odijk J, Piperidis S (eds) Proceedings of the Eighth International Conference on Language Resources and Evaluation (LREC 2012), Istanbul, Turkey, European Language Resources Association (ELRA), Paris, France, pp 303–310

[6] Chiarcos C, Hellmann S, Nordhoff S (2011) Towards a linguistic linked open data cloud: The open linguistics working group. Traitement automatique des langues 52(3):245–275

[7] Chiarcos C, Nordhoff S, Hellmann S (eds) (2012) Linked Data in Linguistics. Representing and Connecting Language Data and Language Metadata. Springer, Heidelberg

[8] Eckle-Kohler J, Gurevych I, Hartmann S, Matuschek M, Meyer CM (2012) UBY-LMF — A Uniform Format for Standardizing Heterogeneous Lexical-Semantic Resources in ISO-LMF. In: Calzolari N, Choukri K, Declerck T, Doğan MU, Maegaard B, Mariani J, Moreno A, Odijk J, Piperidis S (eds) Proceedings of the 8th International Conference on Language Resources and Evaluation (LREC 2012), Istanbul, Turkey, European Language Resources Association (ELRA), Paris, France, pp 275–282

[9] Eckle-Kohler J, Gurevych I, Hartmann S, Matuschek M, Meyer CM (2013) UBY-LMF - exploring the boundaries of language-independent lexicon models. In: Francopoulo G (ed) LMF: Lexical Markup Framework, theory and practice, ISTE - Wiley, London, UK, pp 145–156

[10] Ehrmann M, Cecconi F, Vannella D, Mccrae JP, Cimiano P, Navigli R (2014) Representing Multilingual Data as Linked Data: the Case of BabelNet 2.0. In: Calzolari N, Choukri K, Declerck T, Loftsson H, Maegaard B, Mariani J, Moreno A, Odijk J, Piperidis S (eds) Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC'14), Reykjavik, Iceland, European Language Resources Association (ELRA), Paris, France

[11] Farrar S, Langendoen DT (2003) Markup and the GOLD ontology. In: EMELD Workshop on Digitizing and Annotating Text and Field Recordings, Michigan State University, pp 1–19

[12] Fellbaum C (1998) WordNet. MIT Press, Cambridge, MA

[13] Francopoulo G, Bel N, George M, Calzolari N, Monachini M, Pet M, Soria C (2006) Lexical Markup Framework (LMF). In: Proceedings of the Fifth International Conference on Language Resources and Evaluation (LREC'06), Genoa, Italy, European Language Resources Association (ELRA), Paris, France, pp 233–236

[14] Gurevych I, Eckle-Kohler J, Hartmann S, Matuschek M, Meyer CM, Wirth C (2012) UBY - A Large-Scale Unified Lexical-Semantic Resource. In: Proceedings of the 13th Conference of the European Chapter of the Association for Computational Linguistics (EACL 2012), Avignon, France, Association for Computational Linguistics, Stroudsburg, PA, USA, pp 580–590

[15] Kipper K, Korhonen A, Ryant N, Palmer M (2008) A Large-scale Classification of English Verbs. Language Resources and Evaluation 42:21–40

[16] Laparra E, Rigau G (2009) Integrating WordNet and FrameNet using a Knowledge-based Word Sense Disambiguation Algorithm. In: Proceedings of the 7th International Conference on Recent Advances in Natural Language Processing (RANLP-2009), Borovets, Bulgaria, Association for Computational Linguistics, Stroudsburg, PA, USA, pp 208–213

[17] Matuschek M, Meyer CM, Gurevych I (2013) Multilingual Knowledge in Aligned Wiktionary and OmegaWiki for Computer-Aided Translation. Translation: Corpora, Computation, Cognition (TC3), vol 3, no 1 pp 87–118

[18] McCrae J, , Montiel-Ponsoda E, Cimiano P (2012) Integrating WordNet and Wiktionary with lemon. In: [7], pp 25–34

[19] McCrae J, Aguado-de Cea G, Buitelaar P, Cimiano P, Declerck T, Gómez-Pérez A, Gracia J, Hollink L, Montiel-Ponsoda E, Spohr D, Wunner T (2012) Interchanging lexical resources on the Semantic Web. Language Resources and Evaluation 46:701–719,

[20] McCrae JP, Cimiano P (2013) Mining translations from the web of open linked data. In: Proceedings of the Joint Workshop on NLP&LOD and SWAIE: Semantic Web, Linked Open Data and Information Extraction. Associated with the 9th International Conference on Recent Advances in Natural Language Processing (RANLP 2013), Hissar, Bulgaria, Association for Computational Linguistics, Stroudsburg, PA, USA, pp 8–11

[21] de Melo G, Weikum G (2008) Language as a Foundation of the Semantic Web. In: Bizer C, Joshi A (eds) Proceedings of the Poster and Demonstration Session at the 7th International Semantic Web Conference (ISWC 2008), CEUR, Karlsruhe,

Germany, CEUR WS, vol 401

[22] de Melo G, Weikum G (2010) Towards Universal Multilingual Knowledge Bases. In: Bhattacharyya P, Fellbaum C, Vossen P (eds) Principles, Construction, and Applications of Multilingual Wordnets. Proceedings of the 5th Global WordNet Conference (GWC 2010), Mumbai, India, Narosa Publishing, New Delhi, India, pp 149–156

[23] Meyer CM, Gurevych I (2011) What Psycholinguists Know About Chemistry: Aligning Wiktionary and WordNet for Increased Domain Coverage. In: Wang H, Yarowsky D (eds) Proceedings of the 5th International Joint Conference on Natural Language Processing (IJCNLP), Asian Federation of Natural Language Processing, Chiang Mai, Thailand, pp 883–892

[24] Narayanan S, Baker C, Fillmore C, Petruck M (2003) FrameNet Meets the Semantic Web: Lexical Semantics for the Web. In: Fensel D, Sycara K, Mylopoulos J (eds) The Semantic Web - ISWC 2003, Lecture Notes in Computer Science, vol 2870, Springer Berlin Heidelberg, pp 771–787,

[25] Tonelli S, Pighin D (2009) New Features for FrameNet - WordNet Mapping. In: Proceedings of the Thirteenth Conference on Computational Natural Language Learning (CoNLL-2009), Boulder, Colorado, USA, Association for Computational Linguistics, Stroudsburg, PA, USA, pp 219–227

[26] Unger C, Hieber F, Cimiano P (2010) Generating LTAG grammars from a lexicon-ontology interface. In: Proceedings of the 10th International Workshop on Tree Adjoining Grammars and Related Formalisms (TAG+10), Yale University, New Haven, Connecticut, USA, pp 61–68

[27] Van Assem M, Gangemi A, Schreiber G (2006) Conversion of WordNet to a standard RDF/OWL representation. In: Proceedings of the Fifth International Conference on Language Resources and Evaluation (LREC'06), Genoa, Italy, European Language Resources Association (ELRA), Paris, France, pp 237–242

[28] Windhouwer M, Wright SE (2012) Linking to linguistic data categories in ISOcat. In: [7], pp 99–107