Towards OntoLex-Lemon editing in VocBench 3

MANUEL FIORELLI, ARMANDO STELLATO, TIZIANO LORENZETTI, ANDREA TURBATI[+],
PETER SCHMITZ, ENRICO FRANCESCONI, NAJEH HAJLAOUI, BRAHIM BATOUCHE[*]

ABSTRACT: Released in early May 2016, the OntoLex-Lemon model is a suite of agreed-upon RDF vocabularies for the representation of ontology lexicons. The OntoLex-Lemon model, as well as its predecessor Monnet *lemon*, has also been used (actually "misused", from a certain point of view) to represent lexical-semantic resources (e.g. wordnets), dictionaries and, more broadly, to serve as a cornerstone of the Linguistic Linked Open Data cloud. The number of users potentially interested in editing or consuming OntoLex-Lemon data is thus very large. However, common ontology and RDF editors are inconvenient because of the complex design patterns embodied in the OntoLex-Lemon model, which relies heavily on reification and indirection. In this paper, we discuss our ongoing work to extend the collaborative thesaurus and ontology editor VocBench 3 with facilities tailored to the OntoLex-Lemon model, while retaining its large feature set and the wide modelling spectrum offered by RDF.

Keywords: Lemon, OntoLex, VocBench, Lexicon, RDF.

1. Introduction

The W3C Community Group Ontology-Lexicon[1] (OntoLex) published its final report[2] in early May 2016, defining the OntoLex-Lemon[3] model: a suite of RDF vocabularies (called modules) for the representation of lexicons for ontologies, in accordance with Semantic Web[4] best practices. The modules of OntoLex-Lemon cover aspects such as morphology, syntax-semantics mapping, variation, translation, and linguistic metadata. This rich linguistic characterization of ontologies is unattainable with widely deployed models on the Semantic Web (e.g. RDFS and SKOS-(XL) labels), and it enables a wide range of ontology-driven NLP applications (e.g. knowledge verbalization, semantic parsing, question answering…)[5]. Outside of its originally intended scope, OntoLex-Lemon (as well as its predecessors) has also been used to represent and interlink lexicons, lexical-semantic resources and, in general, language resources in the Linguistic Linked Open Data (LLOD) cloud[6].

There is thus a large and varied group of people potentially interested in consuming, editing or otherwise interacting with datasets that use the OntoLex-Lemon model. Unfortunately, common

---

[+] Dept. of Enterprise Engineering, University of Rome Tor Vergata, Via del Politecnico 1, 00133 Rome, Italy. {fiorelli, turbati}@info.uniroma2.it, stellato@uniroma2.it, tiziano.lorenzetti@gmail.com

[*] Publications Office of the European Union, Luxembourg. {Peter.SCHMITZ, Enrico.FRANCESCONI}@publications.europa.eu, {Najeh.HAJLAOUI, Brahim.BATOUCHE}@ext.publications.europa.eu

Institute of Legal Information Theory and Techniques (ITTIG), Consiglio Nazionale delle Ricerche (CNR), Via dei Barucci 20, 50127 Florence, Italy

[1] Ontology-Lexica Community Group: <https://www.w3.org/community/ontolex/> (last consulted: 23/01/2018)

[2] Lexicon Model for Ontologies: <https://www.w3.org/2016/05/ontolex/> (last consulted: 23/01/2018)

[3] J. MCCRAE., J. BOSQUE-GIL, J. GRACIA, *et al.*, *The OntoLex-Lemon Model: Development and Applications*, in "Proceedings of eLex 2017 conference", Leiden, The Netherlands, 19-21 September 2017, pp. 19-21

[4] T. BERNERS-LEE, J.A. HENDLER, O. LASSILA, *The Semantic Web: A new form of Web content that is meaningful to computers will unleash a revolution of new possibilities*, in «*Scientific American*», vol. CCLXXXIV, n. 5, 2001, pp. 34-43, doi: https://doi.org/10.1038/scientificamerican0501-34

[5] P. CIMIANO, C. UNGER, J. MCCRAE, *Ontology-Based Interpretation of Natural Language*, in *Synthesis Lectures on Human Language Technologies*, vol. VII, n. 2, 2014, pp. 1-178, doi: http://dx.doi.org/10.2200/S00561ED1V01Y201401HLT024

[6] Aa.Vv., *Linked Data in Linguistics*, edited by C. Chiarcos, S. Nordhoff, S. Hellmann, Springer, Berlin, Heidelberg, 2012, doi: https://doi.org/10.1007/978-3-642-28249-2

ontology and RDF editors are highly inconvenient for those purposes, since they lack the facilities and the appropriate abstraction over the OntoLex-Lemon model that are required to easily create and understand the often complex modelling patterns demanded by OntoLex-Lemon.

Conversely, systems tailored to (a specific application of) OntoLex-Lemon often lose the flexibility of the RDF model, which allows loading and mixing arbitrary vocabularies, as well as additional features of ontology/RDF editors (unless they are recreated in the purpose-built system).

Searching for a solution to the problem above, we explored a third way in which an ontology/RDF editor is extended to better support the OntoLex-Lemon model, while remaining compatible with the underlying dynamics of the editor. Indeed, the exploitation of the existing ecosystem was advocated[7] as an important benefit of Linked Data adoption in Linguistics.

Specifically, we are working on extending the collaborative ontology and thesaurus editor VocBench 3[8]. In this paper, we describe the work that has already been done, and lay down the design of the features that are still required to achieve our final goal. We first present an early customization effort[9] that was actually aimed at validating the extensibility of VocBench (via the so-called *custom forms*) more than at creating a full-fledged OntoLex-Lemon editor. Subsequently, we report on the integration of a module for effective management of the metadata module (LIME). The lessons we learned with the *custom form* experiment and the identified gaps motivated the implementation of the support for OntoLex-Lemon as a first-class citizen inside VocBench 3, on par with other modelling vocabularies such as RDFS, OWL, SKOS and SKOS-XL. This work is still in the design phase, which will be presented in this paper.

This effort is being done in the context of the Public Multilingual Knowledge Management Infrastructure (PMKI) action, launched by the European Commission (EC) to promote the Digital Single Market in the European Union (EU). PMKI aims to share maintainable and sustainable Language Resources making them interoperable in order to support language technology industry, and public administrations, with multilingual tools able to improve cross border accessibility of digital services

The objective of PMKI is to implement a proof-of-concept infrastructure to expose and to harmonize internal (European Union institutional) and external multilingual lexicons aligning them in order to facilitate interoperability. Additionally, the project aims to create a governance structure to extend systematically the infrastructure by the integration of supplementary public multilingual taxonomies/terminologies.

2. State of the Art on Linguistic Resources and Language Representation

This is a very broad field, as the following definition suggests: «The term linguistic resources refers to (usually large) sets of language data and descriptions in machine readable form, to be used in building, improving, or evaluating natural language (NL) and speech algorithms or systems»[10].

---

[7] C. CHIARCOS, J. MCCRAE, P. CIMIANO, C. FELLBAUM, *Towards Open Data for Linguistics: Linguistic Linked Data*, in *New Trends of Research in Ontologies and Lexical Resources. Theory and Applications of Natural Language Processing*, edited by A. Oltramari, P. Vossen, L. Qin, *et al.*, Springer, Berlin, Heidelberg 2013, pp. 7-25, doi: https://doi.org/10.1007/978-3-642-31782-8_2

[8] A. STELLATO, A. TURBATI, M. FIORELLI, *et al.*, *Towards VocBench 3: Pushing Collaborative Development of Thesauri and Ontologies Further Beyond*, in "Proceedings of the 17th European Networked Knowledge Organization Systems (NKOS) Workshop", Thessaloniki, Greece, September 21st 2017, pp. 39-52

VocBench web site: <http://vocbench.uniroma2.it/> (last consulted: 25/01/2017)

[9] M. FIORELLI, T. LORENZETTI, M.T. PAZIENZA, *et al.*, *Assessing VocBench Custom Forms in Supporting Editing of Lemon Datasets*, in *Language, Data, and Knowledge*, edited by J. Gracia, F. Bond, J.P. McCrae, *et al.*, Lecture Notes in Artificial Intelligence, vol (X)CCCXVIII, Springer, Cham, pp.237-252, doi: https://doi.org/10.1007%2F978-3-319-59888-8_21

[10] Aa.Vv., *Survey of the State of the Art in Human Language Technology*, edited by R.A. Cole, J. Mariani, H. Uszkoreit, *et al.*, Cambridge University Press, Cambridge, UK 1997

There have been multiple efforts in the past aimed at achieving consensus among different theoretical perspectives and design approaches. The Text Encoding Initiative (TEI)[11] and the LRE-EAGLES (Expert Advisory Group on Linguistic Engineering Standards) project[12] are just a few. They intended to promote the reuse of existing (partial) linguistic resources and the development of new ones for those languages and domains where linguistic resources were unavailable, and creating a cooperative infrastructure to collect, maintain, and disseminate linguistic resources.

WordNet[13] is lexico-semantic resource for American English, which defined a model for subsequent wordnets in other languages[14]. These wordnets clearly separate words, senses and glosses, and are characterized by diverse semantic relations like hyponymy and meronymy.

Lexical Markup Framework[15] (LMF) is a more recent effort, become an ISO standard (LMF; ISO 24613:2008), which supports the representation of monolingual, bilingual or multilingual lexical resources. It covers different aspects, such as morphology, syntax, semantics, and translation.

With the advent of the Semantic Web and Linked Open Data, a number of models have been proposed to enrich ontologies with information about how vocabulary elements have to be expressed in natural language. These include the Linguistic Watermark framework[16], LexOnto[17], LingInfo[18], LIR[19], LexInfo[20] and Monnet *lemon*[21]. The *lemon* model envisions an open ecosystem in which ontologies and lexicons for them co-exist, both of which are published as data on the Web.

In 2012, the OntoLex W3C Community Group was chartered to define an agreed specification informed by the aforementioned models, whose designers are all involved in the community group.

The OntoLex-Lemon model is primarily based on the ideas found in Monnet *lemon*, which was already adopted by a number of lexicons[22]. More specifically, OntoLex-Lemon consists of a number of vocabularies corresponding to different modules: core, synsem, decomp, vartrans, lime. The core module (Figure 1) retains from Monnet *lemon* the separation between the lexical and the ontological layer (following Buitelaar[23] and Cimiano *et al.*[24]), where the ontology describes the semantics of the domain and the lexicon describes the morphology, syntax and pragmatics of the words used to express the domain in a language. A lexicon consists of lexical entries with a single syntactic class (part-of-speech) to which a number of forms are attached (e.g. the singular/plural

---

11 Text Encoding Initiative: <http://www.tei-c.org> (last consulted: 23/01/2018)

12 EAGLES: <http://www.ilc.cnr.it/EAGLES/home.html> (last consulted: 23/01/2018)

13 G.A. MILLER, R. BECKWITH, C. FELLBAUM, *et al.*, *Introduction to WordNet: An On-line Lexical Database*, in «*International Journal of Lexicography*», vol. III, n. 4, 1990, pp. 235-244, doi: https://doi.org/10.1093/ijl/3.4.235 ; *WordNet. An Electronic Lexical Database*, edited by C. Fellbaum, The MIT Press, Cambridge, MA, USA 1998

14 Wordnets in the World: <http://globalwordnet.org/wordnets-in-the-world/> (last consulted: 02/02/2018)

15 G. FRANCOPOULO, M. GEORGE, N. CALZOLARI, *et al.*, *Lexical Markup Framework (LMF)*, in "Proceeding of the 5th International Conference on Language Resources and Evaluation (LREC 2006)", Genoa, Italy, 24-26 May 2006, pp. 233-236

16 M.T. PAZIENZA, A. STELLATO, A. TURBATI, *Linguistic Watermark 3.0: an RDF framework and a software library for bridging language and ontologies in the Semantic Web*, in "Proceedings of the 5th Workshop on Semantic Web Applications and Perspectives (SWAP2008)", Rome, Italy, 15-17 December 2008; A. OLTRAMARI, A. STELLATO, *Enriching Ontologies with Linguistic Content: an Evaluation Framework*, in "Proceedings of OntoLex 2008", Marrakech, Morocco, 31 May 2008

17 P. CIMIANO, P. HAASE, M. HEROLD, *et al.*, *LexOnto: A Model for Ontology Lexicons for Ontology-based NLP*, in "Proceedings of the OntoLex07 Workshop (held in conjunction with ISWC'07)", Busan, Korea, 11 November 2007

18 P. BUITELAAR, T. DECLERCK, A. FRANK, *et al.*, *LingInfo: Design and Applications of a Model for the Integration of Linguistic Information in Ontologies*. In "Proceedings of OntoLex06", Genoa, Italy, 27 May 2006

19 E. MONTIEL-PONSODA, G. AGUADO DE CEA, A. GÓMEZ-PÉREZ, *et al.*, *Enriching ontologies with multilingual information*, in «*Natural Language Engineering*», vol. XVII, n. 3, 2011, pp. 283-309, doi: https://doi.org/10.1017/S1351324910000082

20 P. CIMIANO, P. BUITELAAR, J. MCCRAE, *et al.*, *LexInfo: A declarative model for the lexicon-ontology interface*, in «*Web Semantics: Science, Services and Agents on the World Wide Web*», vol. IX, n. 1, 2011, pp. 29-51, doi: https://doi.org/10.1016/j.websem.2010.11.001

21 J. MCCRAE, G. AGUADO-DE-CEA, P. BUITELAAR, *et al.*, *Interchanging lexical resources on the Semantic Web,* in «*Language Resources and Evaluation*», vol. XLVI, n. 4, 2012, pp. 701–719, doi: https://doi.org/10.1007/s10579-012-9182-3

22 L. BORIN, D. DANNELLS, M. FORSBERG, *et al.*, *Representing Swedish Lexical Resources in RDF with lemon*, in "Proceedings of the ISWC 2014 Posters & Demonstrations Track", Riva del Garda, Italy, October 21, 2014, pp. 329-332; R. NAVIGLI, S. PONZETTO, *BabelNet: The automatic construction, evaluation and application of a wide-coverage multilingual semantic network*, in «*Artificial Intelligence*», vol. CXCIII, 2012, pp. 217-250, doi: https://doi.org/10.1016/j.artint.2012.07.001; J. ECKLE-KOHLER, J. MCCRAE, C. CHIARCOS, *lemonUby – A large, interlinked, syntactically-rich lexical resource for ontologies*, in «*Semantic Web*», vol. VI, n. 4, 2015, pp. 371-378, doi: http://doi.org/10.3233/SW-140159

23 P. BUITELAAR, *Ontology-based semantic lexicons: Mapping between terms and object descriptions*, in *Ontology and the Lexicon*, edited by C.R. Huang, N. Calzolari, A. Gangemi, *et al.*, Cambridge University Press, Cambridge, United Kingdom, 2010

24 P. CIMIANO, J. MCCRAE, P. BUITELAAR, *et al.*, *On the Role of Senses in the Ontology-Lexicon*, in *New Trends of Research in Ontologies and Lexical Resources. Theory and Applications of Natural Language Processing* edited by A. Oltramari, P. Vossen, L. Qin, *et al.*, Springer, Berlin, Heidelberg, 2013, pp. 43-62, doi: https://doi.org/10.1007/978-3-642-31782-8_4

forms of a noun), and each form has a number of representations (string forms), e.g. written or phonetic representation. While an entry can be linked directly to an entity in an ontology, usually the binding between them is realized by a lexical sense resource where pragmatic information such as domain or register of the connection may be recorded. Lexical concepts were introduced in the model to represent the "semantic pole of linguistic units, mentally instantiated abstractions which language users derive from conceptions"[25]. They are intended to represent abstractions in existing lexical resources such as synsets in wordnets.

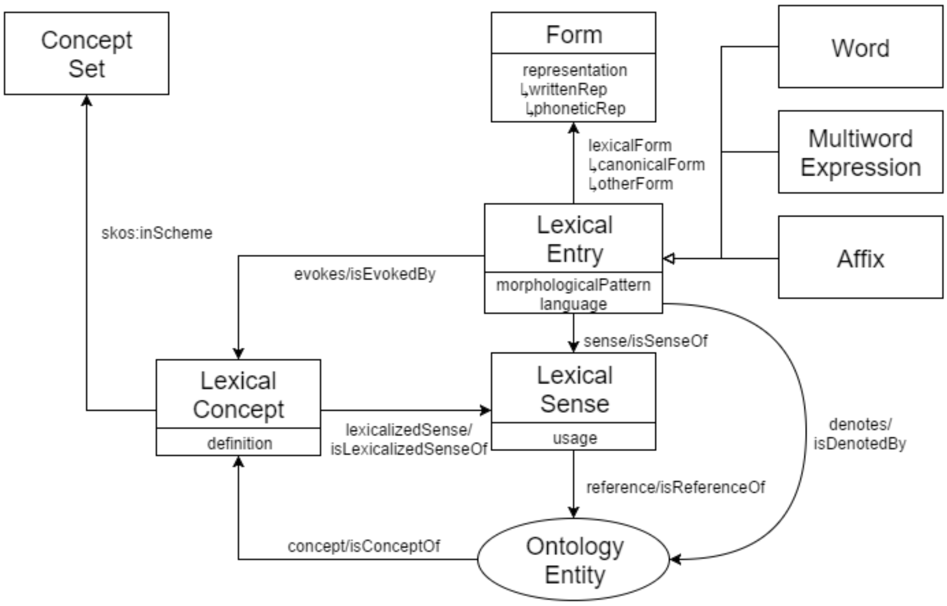

Figure 1 The OntoLex-Lemon core module

The synsem module (left side of Figure 2) allows to associate a lexical entry with a syntactic frame (representing a stereotypical syntactic context for the entry), while an ontology mapping can be used to bind syntactic and semantic arguments together.

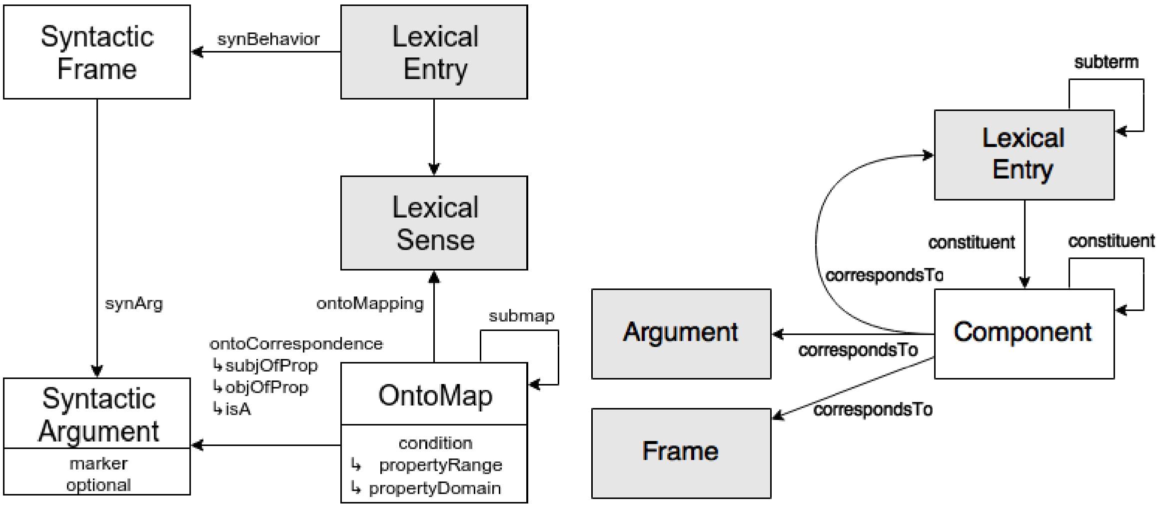

Figure 2 The syntax-semantics module (synsem) on the left and the decomposition module (decomp) on the right

The decomp module (right side of Figure 2) is concerned with the decomposition of a lexical entry into its constituents (i.e. tokens). The class decomp:Component models these constituents, which in turn correspond to lexical entries. This indirection allows recording inside a component

---

[25] V. EVANS, *Lexical concepts, cognitive models and meaning-construction*, in «*Cognitive Linguistics*», vol. XVII, n. 4, 2006, pp. 491-534, doi: https://doi.org/10.1515/COG.2006.016

information such as the fact that the entry "autonomo"@es occurs with feminine gender inside "comunidad autonoma"@es. We can also represent parse trees, by subdividing a component into its constituents.

Because of the lack of space, we will not introduce vatrans and lime, but necessary information about them will be provided briefly later. Additionally, Fiorelli *et al.*[26] describes the design of (a release candidate version of) LIME under the perspective of metadata-based discovery and exploitation of linguistic information in different tasks, including ontology mediation[27].

While Semantic Web practitioners recognized the benefits of linguistic information, linguists in turn acknowledged[28] that the adoption of Semantic Web technologies could benefit the publication and integration of language resources. This led to the formation of the Linguistic Linked Open Data (LLOD) cloud. There is thus a convergence of interests and results between these two communities. Unsurprisingly, recent discussions[29] on OntoLex-Lemon were focused on improving its suitability to encode (legacy) language resources, departing from its original focus on ontology lexicons.

3. VocBench

VocBench[30] is a web-based collaborative thesaurus (and ontology)[31] editor supporting access control, history, and structured validation workflows. The latter, in particular, is relevant to some large organizations that need it to enforce quality control over proposed changes, such as in the case of the Food and Agriculture Organization (FAO) of the United Nations and its Agrovoc[32] thesaurus. VocBench is based on Semantic Web standards such as RDF, OWL and SKOS(-XL). VocBench is powered by the Knowledge Management and Acquisition platform Semantic Turkey[33], which handles data management, persistence and most of its functionality.

While previously focused on SKOS-XL thesauri, VocBench 3, being developed in the context of an action funded by the ISA^2 work programme, has now a wider scope, encompassing generic ontology/RDF editing, while at the same time retaining and improving much appreciated characteristics such as collaboration, validation and publication workflow. In previous versions, these capabilities were implemented either by the VocBench web application or by a dedicated extension of Semantic Turkey. Conversely, in VocBench 3 these distinguishing features were completely integrated into Semantic Turkey, making them available to any application developed on top of this platform. Therefore, VocBench became a mere user interface for Semantic Turkey.

---

[26] M. FIORELLI, A. STELLATO, J.P. MCCRAE, *et al. LIME: The Metadata Module for OntoLex*, in *The Semantic Web. Latest Advances and New Domains*, edited by F. Gandon, M. Sabou, H. Sack, *et al.*, Lecture Notes in Computer Science, vol. M(X)LXXXVIII, Springer, Cham 2015, pp. 321-336, doi: https://doi.org/10.1007/978-3-319-18818-8_20

[27] M. FIORELLI, M.T. PAZIENZA, A. STELLATO, *A meta-data driven platform for semi-automatic configuration of ontology mediators*, in "Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC'14)", Reykjavik, Iceland, 26-31 May 2014

[28] C. CHIARCOS, J. MCCRAE, P. CIMIANO, *et al.*, *Towards Open Data for Linguistics: Linguistic Linked Data*, in *New Trends of Research in Ontologies and Lexical Resources. Theory and Applications of Natural Language Processing*, edited by Oltramari A., Vossen P., Qin L., *et al.*, Springer, Berlin, Heidelberg 2013, pp. 7-25, doi: https://doi.org/10.1007/978-3-642-31782-8_2

[29] K. FAHAD, A. BELLANDI, F. BOSCHETTI, *The Challenges of Converting Legacy Lexical Resources to Linked Open Data using Ontolex-Lemon:The Case of the Intermediate Liddell-Scott Lexicon*, in "Proceedings of OntoLex-2017 1st Workshop on the OntoLex Model (co-located with LDK-2017)", Galway, Ireland, 18 June, 2017; S. STOLK, *OntoLex and Onomasiological Ordering: Supporting Topical Thesauri*, in "Proceedings of OntoLex-2017 1st Workshop on the OntoLex Model (co-located with LDK-2017)", Galway, Ireland, 18 June, 2017 ; J. BOSQUE-GIL, J. GRACIA, E. MONTIEL-PONSODA, *Towards a Module for Lexicography in OntoLex*, in «*DICTIONARY News*», vol. VII, 2017, pp.7-12 < http://kdictionaries.com/kdn/kdn25.pdf#page=7>

[30] A. STELLATO, S. RAJBHANDARI, A. TURBATI, *et al.*, *VocBench: A Web Application for Collaborative Development of Multilingual Thesauri*, in The *Semantic Web. Latest Advances and New Domains*, edited by F. Gandon, M. Sabou, H. Sack, *et al.*, Lecture Notes in Computer Science, vol. M(X)LXXXVIII, Springer, Cham 2015, pp. 38-53, doi: https://doi.org/10.1007/978-3-319-18818-8_3

[31] Since version 3

[32] C. CARACCIOLO, A. STELLATO, A. MORSHED, *et al.*, *The AGROVOC Linked Dataset*, in «*Semantic Web Journal*», vol. IV, n. 3, 2013, pp. 341-348, doi: https://doi.org/10.3233/SW-130106

[33] M.T. PAZIENZA, N. SCARPATO, A. STELLATO, *et al.*, *Semantic Turkey: A Browser-Integrated Environment for Knowledge Acquisition and Management*, in «*Semantic Web Journal*», vol. III, n. 3, 2012, pp. 279-292, doi: https://doi.org/10.3233/SW-2011-0033

4. Custom forms for OntoLex-Lemon

Custom forms are a new feature introduced in VocBench 3 to support personalized forms for the creation of new resources. They can be further distinguished as either *custom constructors*, used to create an instance per se, or as *custom ranges*, used to create a new resource that is assigned to a property of another resource. *Custom constructors* are suitable when the customization applies to every instance of a class: for example, they allow to require that every Student is associated with a student identifier. On the other hand, *custom ranges* express a customization that is bound to a specific property. The property skos:note is a typical example: from an axiomatic perspective, it can hold any type of value (i.e. IRI, bnode, or literal), however a *custom range* can be associated with that property to support the specific pattern: *Documentation as a Related Resource Description* expressed in the SKOS-PRIMER, section 4.2[34]

We validated the design of *custom forms* through its application to the OntoLex-Lemon use case. In particular, we evaluated to which extent we could map existing *lemon* patterns for the ontology-lexicon interface[35] to *custom forms*. This use case is particularly interesting, because:

1. each entry of the ontology lexicon is associated with a complex graph
2. these graphs are instances of a few templates associated with known design patterns

The result of this experimentation is a collection of custom forms[36] that should be used to customize the range of the property lime:entry, which relates a lime:Lexicon to its ontolex:LexicalEntry(s).

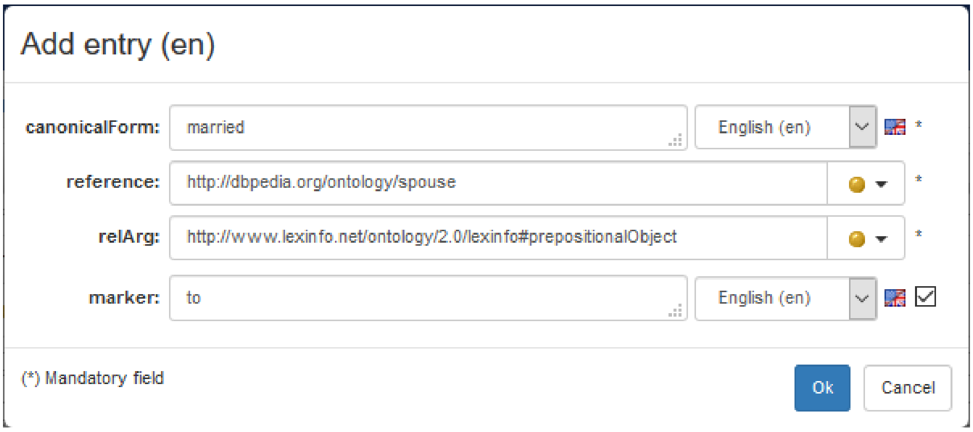

Figure 3 An already filled custom form describing the relational adjective "married"

Figure 3 illustrates a custom form filled with information binding the property dbo:spouse to the (relational) adjective “married (to)”. The first two fields in the form hold the canonical (uninflected) form (i.e. married) and the ontology reference (i.e. dbo:spouse), respectively. To understand the other two fields, we should consider that the stereotypical syntactic behavior of a relational adjective is a predicative frame with a prepositional argument marked by a given preposition. In our example, the stereotypical sentence is “x is married to y” and the marker (fourth field) is the preposition “to”, while the type of syntactic argument is lexinfo:prepositionalObject. In this pattern, x and y are mapped, respectively, to the semantic subject and object of the ontology property. In general, the correspondence between syntactic and semantic arguments can be altered.

[34] World Wide Web Consortium (W3C): SKOS Simple Knowledge Organization System Primer. In: World Wide Web Consortium (W3C). Available at: http://www.w3.org/TR/skos-primer

[35] J.P. McCrae, C. Unger, *Design Patterns for Engineering the Ontology-Lexicon Interface*, in *Towards the Multilingual Semantic Web*, edited by P. Buitelaar, P. Cimiano, Springer, Berlin, Heidelberg, 2014, pp. 15-30, doi: https://doi.org/10.1007/978-3-662-43585-4_2

[36] Lemon VocBench Custom Forms: <https://bitbucket.org/art-uniroma2/lemon-vb-customforms> (last consulted: 26/01/2018)

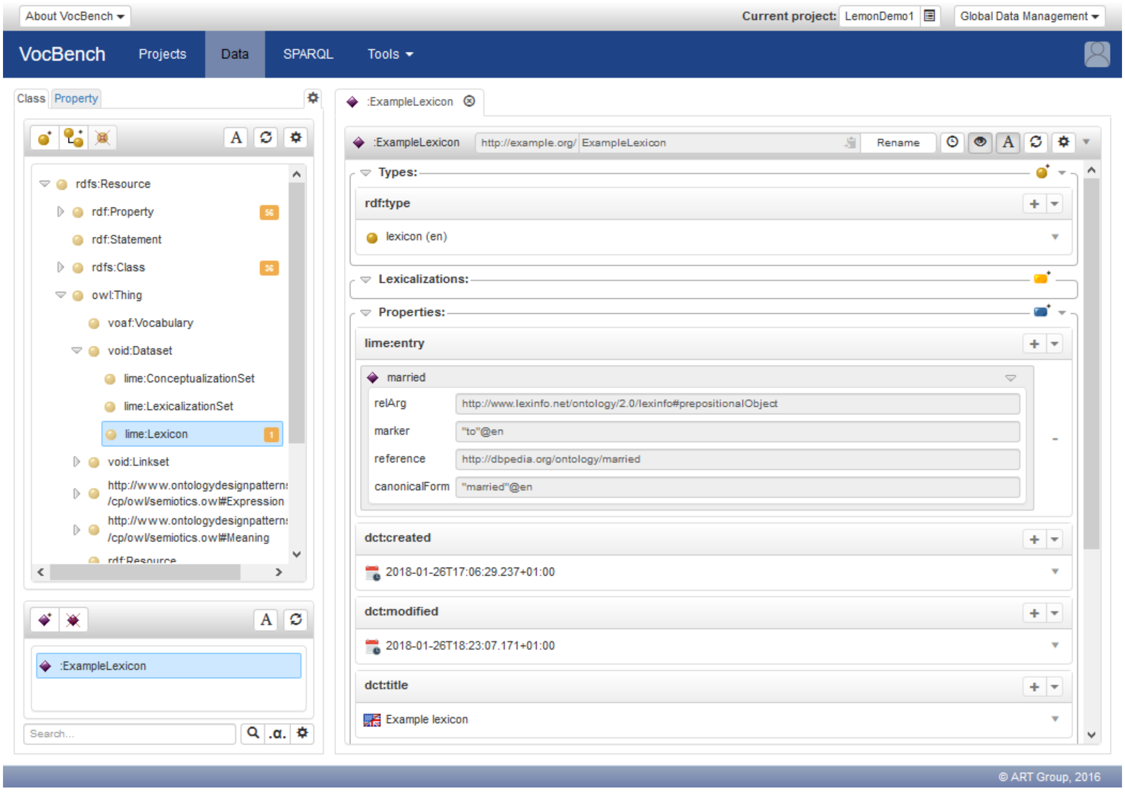

Figure 4 VocBench 3 Data View focused on a lexicon

After describing *custom forms* in isolation, we show how they are integrated into the rest of VocBench 3. In order to start editing OntoLex-Lemon data, it necessary to import the corresponding OWL vocabularies. Being a generic ontology editor, VocBench 3 shows the newly imported classes and properties in the corresponding (hierarchical views), and support their use in the description of resources. Figure 4 shows the *Data View* of VocBench 3 focused on an instance of the class lime:Lexicon, i.e. an OntoLex-Lemon lexicon. The *Resource View* on the right side of the window layouts the RDF description of the lexicon, roughly consisting in a number of property-value pairs. By means of built-in facilities, the user can set properties such as the title of the lexicon, its description, and so on. Ideally, any RDF vocabulary can be mixed and matched inside VocBench 3. It is worth to notice that the picture also shows the (optional) bookkeeping of resource-level version information, having VocBench 3 automatically set the properties dct:created and dct:modified.

The property lime:entry relates a lexicon to its entries. Its instantiation is cumbersome, because the user should define a number of intermediate resources and then arrange them in a complex pattern. The *custom forms* for OntoLex-Lemon solve this problem by providing additional forms tailored to the construction of lexical entries describing an ontology lexicon. These forms are also used to ease the interpretation of lexical entries. In the example, lexical entries are rendered via their lemma (instead of their URI), while the multi-value preview of an entry (resembling a filled-form) condenses information that may be located several steps away from the entry (e.g. the literal "married"@en associated indirectly via an ontolex:Form) or encoded implicitly (e.g. the mapping between syntactic and semantic arguments is realized by unifying their objects).

5. LIME metadata exporter

The OntoLex-Lemon model complements existing metadata vocabularies such as Dublin Core Metadata Terms and the Vocabulary of Interlinked Datasets, with its LIME module, which supports

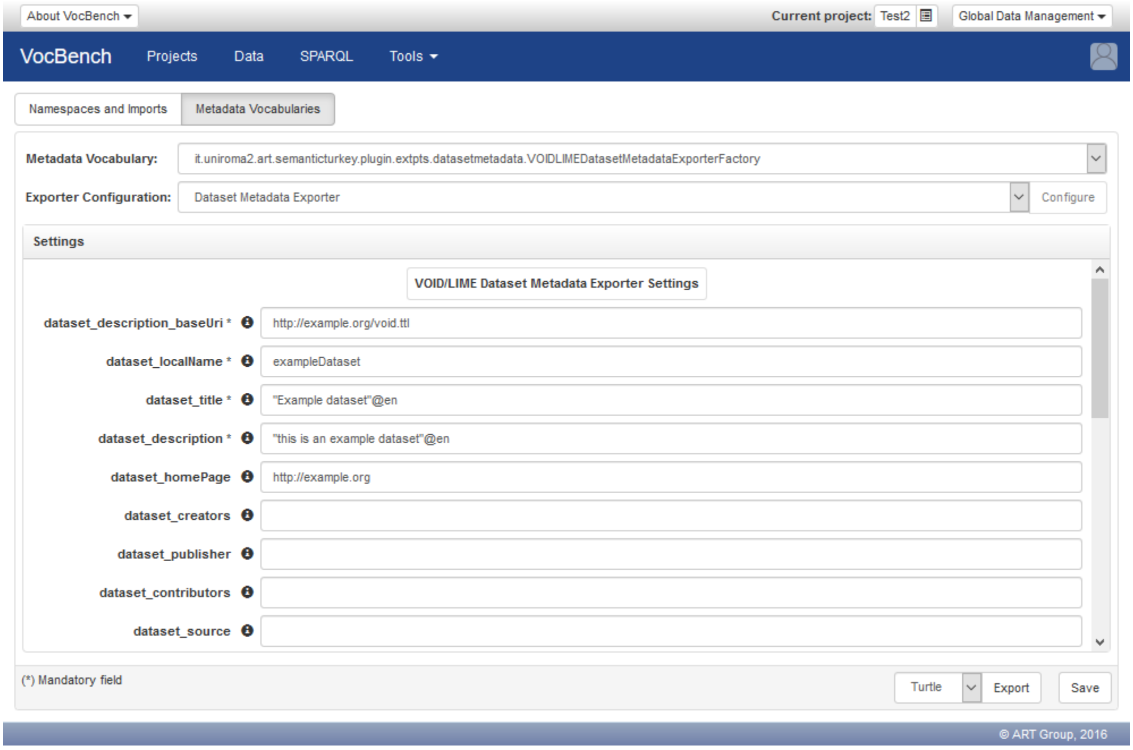

Figure 5 Metadata export framework to produce LIME metadata

the representation of specific metadata about the ontology-lexicon interface. Descriptive metadata about a lexicon (mostly represented using existing vocabularies) are expected to be entered manually. Additionally, LIME defines a number of statistics that would be hard – or at least annoying – to compute and then represent appropriately. We overcome this problem by having a hybrid approach: the user enters some metadata manually (see Figure 5), while the generated description will include a number of statistics that are computed through the LIME API[37].

6. Design of missing features

The work on LIME is somehow separated from the one required by other modules, because LIME is about linguistic metadata rather than linguistic information. The latter was partially tackled by the development of the *custom forms* for OntoLex-Lemon, which however do not completely satisfy the need for a comprehensive OntoLex-Lemon editor. Firstly, they are unable to enforce some constraints (e.g. the language of a lexical entry should match the one of its containing lexicon), they do not support (explicitly) variable-number components (e.g. event verbs, which do not have a fixed number of participants), and they do not support (in the multi-value preview) some ontology lexicon design patterns (whose decoding is ambiguous). We will address most of these issues, when revising and extending the *custom form* mechanism. Nonetheless, *custom forms* by design do not cover some aspects of the system that should be customized for a refined end-user experience.

Indeed, those refinements are the core of our ongoing work on the support for the OntoLex-Lemon model as a first-class citizen of VocBench 3 like other lexicalization models such as RDFS and SKOS(-XL). This first-class support consists in a number of features, which we will implement in the next months. In the following sections, we describe the (early) design of these missing features.

---

[37] M. FIORELLI, M.T. PAZIENZA, A. STELLATO, *An API for OntoLex LIME datasets*. In “Proceedings of OntoLex-2017 1st Workshop on the OntoLex Model (co-located with LDK-2017)”, Galway, Ireland, 18 June, 2017

6.1. Project metadata and configuration

A newly created project must be associated with a *semantic model* telling the nature of the semantic backbone (e.g. OWL ontology vs SKOS thesaurus) and a *lexicalization model* telling which mechanism (e.g. RDFS, SKOS or SKOS-XL labels) is used to associate linguistic information with the semantic resources. OntoLex-Lemon is mainly concerned with the representation of rich lexicons for ontologies, thesauri and RDF datasets in general, therefore we might characterize it as an additional *lexicalization model*. This characterization might work in practice, nonetheless someone may argue that it is at least misleading when a lexicon is edited as an autonomous language resource without any connection to an ontology or other lexicalized dataset. In fact, we might replace this semantic/lexicalization model dichotomy in favor of a more flexible schema.

If we solve the remaining issues with the *custom forms* we have already developed, we may consider to include them in our environment for OntoLex-Lemon editing, instead of providing a native implementation of their capabilities. However, this inclusion presupposes that they will be distributed together with VocBench 3, and moreover configured automatically when a specific type of project is created. Independently from our resolution on the *custom forms*, we should provide a similar auto-configuration facility for importing relevant ontologies inside a project.

6.2. Rendering based on lexicalizations and conceptualizations

Among the uses of the *lexicalization model* we should notice the selection of a suitable *rendering engine*, a component producing human-friendly representations of resources based on their lexicalizations. This feature is of paramount importance when resources have just an alphanumeric identifier, which is often the case in multilingual datasets: a *rendering engine* aware of users' language preferences will actually produce a localized view over a multilingual dataset specifically crafted for each user. When an OntoLex-Lemon lexicon is used to lexicalize a reference dataset or to bind words to lexical concepts, a dedicated *rendering engine* should be used. The provision of a new *rendering engine* is in fact a quite simple activity, already foreseen by VocBench through a dedicated extension point. Nonetheless, white-box modification of the system is requested to implement the automatic suggestion of this rendering engine based on project metadata. Similarly, an intrusive modification of the system will be required to render lexical entries themselves (i.e. showing them with their canonical form in place of their potentially unintelligible identifiers).

6.3. Dedicated visualizations

Under the perspective of an ontology editor, lexicons are just instances of the class lime:Lexicon, while lexical entries in a specific lexicon are just its values for the property lime:entry. An extension for OntoLex-Lemon might introduce more immediate visualizations, like it has been done for the SKOS model. In the case of SKOS, for example, instead of presenting concepts as a flat list of instances, a dedicated panel supported i) the selection of concepts found in given concept schemes, ii) the presentation of concepts in a tree reflecting hierarchical information encoded in SKOS-specific object properties. In the case of OntoLex-Lemon, a dedicated panel might present entries belonging to specified lexicons. A lexicon is a usually long list of entries, which could be browsed more easily if subdivided by means of an alphabetic index.

A similar argument could be raised for the class ontolex:ConceptSet, which is a collector of ontolex:LexicalConcepts. In fact, since these classes extend skos:ConceptScheme and skos:Concept, respectively, existing panels related to SKOS can be used. However, dedicated panels (possibly extending the ones for SKOS) could be introduced for reasons including i) separate

Figure 6 A fragment of the Resource View showing current support for decomposition

OntoLex-Lemon concept sets from generic SKOS concept schemes, ii) additional filtering (e.g. based on linguistic annotations), iii) represent different semantic relations between lexical concepts.

A vartrans:TranslationSet groups together translations of lexical entries. As a minimum, we need a list of available translation sets, while their content could be browsed using the existing resource view. Each translation is then modeled as a resource, so a slight extension is required to have them rendered suitably in the resource view. A dedicated *custom form* could be used to enable the multi-value preview of a translation resource visiting the resource view itself.

6.4 Search aware of OntoLex-Lemon

VocBench provides a sophisticated search function, which is aware of predefined *lexicalization models*. We will extend this function, in order to find resources associated with an OntoLex-Lemon entry that matches the search criteria. Differently from other *lexicalization models*, we might also be interested in searching lexical entries themselves. The rationale is that lexical entries are modeled as resources (in contrast to literals), and they can be used in multiple lexicalizations/conceptualizations as well as be put in complex relationships with other entries.

6.5 Generic visualization/editing of the syntax and semantics interface

The syntax-semantic interface is probably the piece of information whose encoding in OntoLex-Lemon requires the most complex patterns, while resulting in a rather implicit representation. In OntoLex-Lemon, the correspondence between syntactic and semantic arguments is encoded through the unification of corresponding arguments. The multi-value preview of the *custom forms* for the lemon ontology lexicon patterns decodes this complex information as an easy-to-read form. However, in general we do not have an efficient mechanism to show this correspondence. Let us suppose that a lexical entry is shown in the *Resource View*: neither semantic arguments nor syntactic arguments would be visible, because they are introduced indirectly by the syntactic frames and the ontology mappings associated with the lexical entry. Currently, the user can only open the description of these associated resources in new *Resource Views*, therefore there is no chance of having syntactic and semantic arguments displayed together in the same panel.

A possible solution to this problem is to enable nesting the *resource views* of the dependent resources (e.g. syntactic behaviors and ontology mappings) inside the view associated with a resource (e.g. the lexical entry). Indeed, this feature is generally useful, as its presence in some ontology/RDF editors suggest. While nested *Resource Views* solve the problem of not being able to see semantic and syntactic arguments together, they do nothing to facilitate the identification of the correspondence between them. Unfortunately, the flexibility of OntoLex-Lemon makes an intuitive visualization in the general case very difficult, if not impossible.

6.6. Support for lexical entry decomposition

Lexical entries can be multi-word expressions. The VARTRANS module allows to represent the tokenization of these lexical entries, as well as their syntactic parse tree. Dedicated facilities should be provided to improve visualization and editing with respect to the basic offering of VocBench 3.

Figure 6 shows how the *Resource View* renders the tokenization of the compound lexical entry "chief executive officer". Individual tokens are values of the property decomp:constituent, which are unordered as per RDF specification (as can be seen in the picture). The relative position of the tokens may be encoded with the properties rdf:_N, each holding the n-th token. Furthermore, these different pieces of information belong to different sections of the *Resource View*. A section dedicated to decomposition may directly list the components in the proper order (when defined), and provide operations for inserting a token in the right position. The representation of syntactic parse trees is even more complex, because components have components themselves (unless they are leaves). The possibility to nest *Resource Views* that we have envisaged before may benefit users, nonetheless we think it is better to employ the bracket notation or a tree representation (see Figure 7). The bracket notation can also allow users to create parse trees and tokenizations more easily, as an alternative to the explicit creation and connection of many intermediate resources.

6.7. Smart suggestions and use of NLP tools

An OntoLex-Lemon editor should disburden users from tasks that can be automated (at least partially). Let us exemplify this capability with the task of decomposing the lexical entry "chief executive officer". VocBench could use a tokenizer to split the multi-word expression into tokens, then use a postagger and a lemmatizer to identify relevant lexical entries. The system might create non existing entries, while the choice between already existing homograph entries (with the same part of speech) could be delegated to the user. In absence of these components for a given natural language and domain, intelligent suggestions can be generated anyway by looking for lexical entries whose canonical and alternative forms match substrings of the text we want to decompose.

6.8. Support for redundant property chains and reified relations

The OntoLex-Lemon model often provides alternative patterns to represent a certain bit of information with progressively higher level of accuracy. For example, the relation between an ontology concept and a lexical entry can be expressed as a single triple, or instead reified via the introduction of a sense. A similar discourse applies to translations. VocBench should visualize the information irrespectively from the chosen pattern, while redundant data should be kept in sync: for example, if a reified relationship is removed, should we delete the non-reified version as well?

[NP [JJ chief] [JJ executive] [NN officer]]

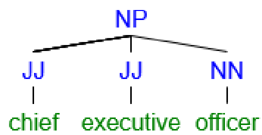

Figure 7 Bracket notation and corresponding graphical representation

6.9. OntoLex-Lemon integrity constraint validation

VocBench 3 has a subsystem for recognizing violations of constraints including: i) some that are beyond the expressivity of the underlying ontology language, ii) informal constraints, iii) resource-specific policies. One common use case is identifying and repairing problems in data loaded from external sources (not benefiting from interactive checks by VocBench). Having OntoLex-Lemon as a first-class citizen of VocBench 3 requires the introduction of checks for new constraints.

6.9. General-purpose improvements to the Resource View

Evaluating OntoLex-Lemon editing with the base VocBench 3 system, we discovered a few limitations of the *Resource View* to be addressed. Firstly, it is not possible (in general) to create a new resource when setting the value of a property, unless a dedicated custom range is fired. This limitation affects the usability of the system when the object is reified. Secondly, the *Resource View* does not support the properties rdf:_N, since their definition cannot be found in any ontology.

7. Related Work

C*ustom forms* use the knowledge acquisition framework CODA[38], and leverage its transformation language PEARL as a form definition language. Form customization can be found in other ontology editors such as TopBraid Composer[39] (TBC) and Protégé[40] (in version 3.x, which influenced WebProtégé[41]). Additionally, the former inspired our notion of nested *Resource Views*.

The *lemon* design patterns for ontology lexicons were implemented as a domain specific language (DSL). Although it can be compiled to RDF, this pattern language is used as an alternative, more concise mechanism to represent ontology lexicons. These representations are often saved as text files, which can be version-controlled like source code. Furthermore, they are often associated with non-RDF editing environments, from text-editors to dedicated systems, such as Lemonade[42]. The latter supports the generation of snippets for lexical entries and cross-language checks.

Lemon source[43] is a wiki-style collaborative ontology lexicon editor (based on Monnet *lemon*): it uses NLP components to enrich the linguistic content encoded in plain labels, and promotes the reuse of lexical entries defined in existing resources (e.g. Princeton WordNet). LexO[44] is another collaborative *lemon* editor, which is being developed with a special focus on the requirements of Humanities (e.g. references to texts).

---

38 M. FIORELLI, M.T. PAZIENZA, A. STELLATO, *et al.*, *CODA: Computer-aided ontology development architecture*, in «*IBM Journal of Research and Development*», vol. LVIII, n. 2/3, 2014, pp. 14-1, doi: https://doi.org/10.1147/JRD.2014.2307518

39 TopBraid Composer Maestro Edition: <https://www.topquadrant.com/tools/ide-topbraid-composer-maestro-edition/> (last consulted: 01/02/2018)

40 J. GENNARI, M. MUSEN, R. FERGERSON, *et al.*, *The evolution of Protégé-2000: An environment for knowledge-based systems development*, in «*International Journal of Human-Computer Studies*», vol. LVIII, n. 1, 2003, pp. 89-123, doi: https://doi.org/10.1016/S1071-5819(02)00127-1

41 T. TUDORACHE, C. NYULAS, N.F. NOY, *et al.*, *WebProtégé: A Collaborative Ontology Editor and Knowledge Acquisition Tool for the Web*, in «*Semantic Web*», vol. IV, n.1, 2013, pp. 89-99

42 M. RICO, C. UNGER, *Lemonade: A Web Assistant for Creating and Debugging Ontology Lexica*, in *Natural Language Processing and Information Systems*, edited by C. Biemann, S. Handschuh, A. Freitas, *et al.*, Lecture Notes in Computer Science, vol. M(X)CIII, Springer, 2015, pp. 448-452, doi: https://doi.org/10.1007/978-3-319-19581-0_45

43 E. MONTIEL-PONSODA, J. MCCRAE, P. CIMIANO, *Collaborative semantic editing of linked data lexica*, in "Proceedings of Eighth International Conference on Language Resources and Evaluation (LREC 2012)", 21/05/2012 - 27/05/2012, Estambul, Turquía, pp. 2619-2625

44 A. BELLANDI, A. WEINGART, *Developing LexO: a Collaborative Editor of Multilingual Lexica and Termino-Ontological Resources in the Humanities*, in "Proceedings of Language, Ontology, Terminology and Knowledge Structures Workshop (LOTKS 2017)", Montpellier, France, 19 September 2017

8. Conclusions

Existing ontology/RDF editors are not very usable in conjunction with the OntoLex-Lemon model, which requires complex models that are difficult to build resource-by-resource. In particular, we set ourselves to solve this problem in relation to the collaborative thesaurus and ontology editor VocBench 3. We described the work done so far, highlighting its limitations and the lessons we learned. Then, we laid down the preliminary design of the missing features that need to be implemented in order to offer a refined OntoLex-Lemon editor.

Acknowledgements

This work has been drafted under the 2016.16 action[45], which is supported by the ISA^2 Programme[46]. ISA^2 is a EUR 131 million programme of the European Commission, supporting the modernization of public administrations in Europe through the development of eGovernment solutions. More than 20 solutions are already available[47], with more to come soon. All solutions are open source and available free of charge to any interested public administration in Europe.

References

Aa.Vv., *Survey of the State of the Art in Human Language Technology*, edited by Cole R.A., Mariani J., Uszkoreit H., *et al.,* Cambridge University Press, Cambridge, UK 1997

Aa.Vv., *WordNet. An Electronic Lexical Database*, edited by C. Fellbaum, The MIT Press, Cambridge, MA, USA 1998

Aa.Vv., *Linked Data in Linguistics*, edited by Chiarcos C., Nordhoff S., Hellmann S., Springer, Berlin, Heidelberg, 2012, doi: https://doi.org/10.1007/978-3-642-28249-2

Bellandi A., Weingart A., *Developing LexO: a Collaborative Editor of Multilingual Lexica and Termino-Ontological Resources in the Humanities*, in “Proceedings of Language, Ontology, Terminology and Knowledge Structures Workshop (LOTKS 2017)”, Montpellier, France, 19 September 2017

Berners-Lee T., Hendler J.A, Lassila O., *The Semantic Web: A new form of Web content that is meaningful to computers will unleash a revolution of new possibilities*, in «*Scientific American*», vol. CCLXXXIV, n. 5, 2001, pp. 34-43, doi: https://doi.org/10.1038/scientificamerican0501-34

Borin L., Dannells D., Forsberg M., *et al.*, *Representing Swedish Lexical Resources in RDF with lemon*, in "Proceedings of the ISWC 2014 Posters & Demonstrations Track", Riva del Garda, Italy, October 21, 2014, pp. 329-332

Bosque-Gil J., Gracia J., Montiel-Ponsoda E., *Towards a Module for Lexicography in OntoLex*, in «*DICTIONARY News*», vol. VII, 2017, pp.7-12 <http://kdictionaries.com/kdn/kdn25.pdf#page=7>

Buitelaar P., Declerck T., Frank A., *et al.*, *LingInfo: Design and Applications of a Model for the Integration of Linguistic Information in Ontologies*. In “Proceedings of OntoLex06”, Genoa, Italy, 27 May 2006

Buitelaar P., *Ontology-based semantic lexicons: Mapping between terms and object descriptions*, in *Ontology and the Lexicon*, edited by Huang C.R., Calzolari N., Gangemi A., *et al.*, Cambridge University Press, Cambridge, United Kingdom, 2010

Caracciolo C., Stellato A., Morshed A., *et al.*, *The AGROVOC Linked Dataset*, in «*Semantic Web Journal*», vol. IV, n. 3, 2013, pp. 341-348, doi: https://doi.org/10.3233/SW-130106

---

[45] https://ec.europa.eu/isa2/actions/overcoming-language-barriers_en

[46] https://ec.europa.eu/isa2/

[47] https://ec.europa.eu/isa2/solutions_en

CHIARCOS C., MCCRAE J., CIMIANO P., FELLBAUM C., *Towards Open Data for Linguistics: Linguistic Linked Data*, in *New Trends of Research in Ontologies and Lexical Resources. Theory and Applications of Natural Language Processing*, edited by Oltramari A., Vossen P., Qin L., *et al.*, Springer, Berlin, Heidelberg 2013, pp. 7-25, doi: https://doi.org/10.1007/978-3-642-31782-8_2

CIMIANO P., HAASE P., HEROLD M., *et al.*, *LexOnto: A Model for Ontology Lexicons for Ontology-based NLP*, in "Proceedings of the OntoLex07 Workshop (held in conjunction with ISWC'07)", Busan, Korea, 11 November 2007

CIMIANO P., BUITELAAR P., MCCRAE J., *et al.*, *LexInfo: A declarative model for the lexicon-ontology interface*, in «*Web Semantics: Science, Services and Agents on the World Wide Web*», vol. IX, n. 1, 2011, pp. 29-51, doi: https://doi.org/10.1016/j.websem.2010.11.001

CIMIANO P., MCCRAE J., BUITELAAR P., *et al.*, *On the Role of Senses in the Ontology-Lexicon*, in *New Trends of Research in Ontologies and Lexical Resources. Theory and Applications of Natural Language Processing* edited by Oltramari A., Vossen P., Qin L., *et al.*, Springer, Berlin, Heidelberg, 2013, pp. 43-62, doi: https://doi.org/10.1007/978-3-642-31782-8_4

CIMIANO P., UNGER C., MCCRAE J., *Ontology-Based Interpretation of Natural Language*, in *Synthesis Lectures on Human Language Technologies*, vol. VII, n. 2, 2014, pp. 1-178, doi: http://dx.doi.org/10.2200/S00561ED1V01Y201401HLT024

ECKLE-KOHLER J., MCCRAE J., CHIARCOS C., *lemonUby – A large, interlinked, syntactically-rich lexical resource for ontologies*, in «*Semantic Web*», vol. VI, n. 4, 2015, pp. 371-378, doi: http://doi.org/10.3233/SW-140159

EVANS V., *Lexical concepts, cognitive models and meaning-construction*, in «*Cognitive Linguistics*», vol. XVII, n. 4, 2006, pp. 491-534, doi: https://doi.org/10.1515/COG.2006.016

FAHAD K., BELLANDI A., BOSCHETTI F., *The Challenges of Converting Legacy Lexical Resources to Linked Open Data using Ontolex-Lemon:The Case of the Intermediate Liddell-Scott Lexicon*, in "Proceedings of OntoLex-2017 1st Workshop on the OntoLex Model (co-located with LDK-2017)", Galway, Ireland, 18 June, 2017

FIORELLI M., PAZIENZA M.T., STELLATO A., *A meta-data driven platform for semi-automatic configuration of ontology mediators*, in "Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC'14)", Reykjavik, Iceland, 26-31 May 2014

FIORELLI M., PAZIENZA M.T., STELLATO A., *et al.*, *CODA: Computer-aided ontology development architecture*, in «*IBM Journal of Research and Development*», vol. LVIII, n. 2/3, 2014, pp. 14-1, doi: https://doi.org/10.1147/JRD.2014.2307518

FIORELLI M., STELLATO A., MCCRAE J.P., *et al. LIME: The Metadata Module for OntoLex*, in *The Semantic Web. Latest Advances and New Domains*, edited by Gandon F., Sabou M., Sack H., *et al*., Lecture Notes in Computer Science, vol. IXLXXXVIII, Springer, Cham 2015, pp. 321-336, doi: https://doi.org/10.1007/978-3-319-18818-8_20

FIORELLI M., LORENZETTI T., PAZIENZA M.T., *et al.*, *Assessing VocBench Custom Forms in Supporting Editing of Lemon Datasets*, in *Language, Data, and Knowledge*, edited by Gracia J., Bond F., McCrae J.P., *et al.*, Lecture Notes in Artificial Intelligence, vol (X)CCCXVIII, Springer, Cham, 2017, pp.237-252, doi: https://doi.org/10.1007%2F978-3-319-59888-8_21

FIORELLI M., PAZIENZA M.T., STELLATO A., *An API for OntoLex LIME datasets*. In "Proceedings of OntoLex-2017 1st Workshop on the OntoLex Model (co-located with LDK-2017)", Galway, Ireland, 18 June, 2017

FRANCOPOULO G., GEORGE M., CALZOLARI N., *et al.*, *Lexical Markup Framework (LMF)*, in "Proceeding of the 5th International Conference on Language Resources and Evaluation (LREC 2006)", Genoa, Italy, 24-26 May 2006, pp. 233-236

GENNARI J., MUSEN M., FERGERSON R., *et al.*, *The evolution of Protégé-2000: An environment for knowledge-based systems development*, in «*International Journal of Human-Computer Studies*», vol. LVIII, n. 1, 2003, pp. 89-123, doi: https://doi.org/10.1016/S1071-5819(02)00127-1

MCCRAE J., AGUADO-DE-CEA G., BUITELAAR P., *et al.*, *Interchanging lexical resources on the Semantic Web*, in «*Language Resources and Evaluation*», vol. XLVI, n. 4, 2012, pp. 701–719, doi: https://doi.org/10.1007/s10579-012-9182-3

MCCRAE J.P., UNGER C., *Design Patterns for Engineering the Ontology-Lexicon Interface*, in *Towards the Multilingual Semantic Web*, edited by Buitelaar P., Cimiano P., Springer, Berlin, Heidelberg, 2014, pp. 15-30, doi: https://doi.org/10.1007/978-3-662-43585-4_2

MCCRAE J., BOSQUE-GIL J., GRACIA J., *et al.*, *The OntoLex-Lemon Model: Development and Applications*, in "Proceedings of eLex 2017 conference", Leiden, The Netherlands, 19-21 September 2017, pp. 19-21

MILLER G.A., BECKWITH R., FELLBAUM C., *et al.*, *Introduction to WordNet: An On-line Lexical Database*, in «*International Journal of Lexicography*», vol. III, n. 4, 1990, pp. 235-244, doi: https://doi.org/10.1093/ijl/3.4.235

MONTIEL-PONSODA E., AGUADO DE CEA G., GÓMEZ-PÉREZ A., *et al.*, *Enriching ontologies with multilingual information*, in «*Natural Language Engineering*», vol. XVII, n. 3, 2011, pp. 283-309, doi: https://doi.org/10.1017/S1351324910000082

MONTIEL-PONSODA E., MCCRAE J., CIMIANO P., *Collaborative semantic editing of linked data lexica*, in "Proceedings of Eighth International Conference on Language Resources and Evaluation (LREC 2012)", 21/05/2012 - 27/05/2012, Estambul, Turquía, pp. 2619-2625

NAVIGLI R., PONZETTO S., *BabelNet: The automatic construction, evaluation and application of a wide-coverage multilingual semantic network*, in «*Artificial Intelligence*», vol. CXCIII, 2012, pp. 217-250, doi: https://doi.org/10.1016/j.artint.2012.07.001

OLTRAMARI A., STELLATO A., *Enriching Ontologies with Linguistic Content: an Evaluation Framework*, in "Proceedings of OntoLex 2008", Marrakech, Morocco, 31 May 2008

PAZIENZA M.T., STELLATO A., TURBATI A., *Linguistic Watermark 3.0: an RDF framework and a software library for bridging language and ontologies in the Semantic Web*, in "Proceedings of the 5th Workshop on Semantic Web Applications and Perspectives (SWAP2008)", Rome, Italy, 15-17 December 2008

PAZIENZA M.T., SCARPATO N., STELLATO A., *et al.*, *Semantic Turkey: A Browser-Integrated Environment for Knowledge Acquisition and Management*, in «*Semantic Web Journal*», vol. III, 2012, n. 3, pp. 279-292, doi: https://doi.org/10.3233/SW-2011-0033

RICO M., UNGER C., *Lemonade: A Web Assistant for Creating and Debugging Ontology Lexica*, in *Natural Language Processing and Information Systems*, edited by Biemann C., Handschuh S., Freitas A., *et al.*, Lecture Notes in Computer Science, vol. M(X)CIII, Springer, 2015, pp. 448-452, doi: https://doi.org/10.1007/978-3-319-19581-0_45

STELLATO A., RAJBHANDARI S., TURBATI A., *et al.*, *VocBench: A Web Application for Collaborative Development of Multilingual Thesauri*, in The *Semantic Web. Latest Advances and New Domains*, edited by Gandon F., Sabou M., Sack H., *et al.*, Lecture Notes in Computer Science, vol. M(X)LXXXVIII, Springer, Cham 2015, pp. 38-53, doi: https://doi.org/10.1007/978-3-319-18818-8_3

STELLATO A., TURBATI A., FIORELLI M., *et al.*, *Towards VocBench 3: Pushing Collaborative Development of Thesauri and Ontologies Further Beyond*, in "Proceedings of the 17th European Networked Knowledge Organization Systems (NKOS) Workshop", Thessaloniki, Greece, September 21st 2017, pp. 39-52

STOLK S., *OntoLex and Onomasiological Ordering: Supporting Topical Thesauri*, in "Proceedings of OntoLex-2017 1st Workshop on the OntoLex Model (co-located with LDK-2017)", Galway, Ireland, 18 June, 2017

TUDORACHE T., NYULAS C., NOY N.F., *et al.*, *WebProtégé: A Collaborative Ontology Editor and Knowledge Acquisition Tool for the Web*, in «*Semantic Web*», vol. IV, n. 1, 2013, pp. 89-99

WORLD WIDE WEB CONSORTIUM (W3C): SKOS Simple Knowledge Organization System Primer. In: World Wide Web Consortium (W3C). Available at: http://www.w3.org/TR/skos-primer

Webography

Ontology-Lexica Community Group: <https://www.w3.org/community/ontolex/> (last consulted: 23/01/2018)

Lexicon Model for Ontologies: <https://www.w3.org/2016/05/ontolex/> (last consulted: 23/01/2018)

VocBench web site: <http://vocbench.uniroma2.it/> (last consulted: 25/01/2017)

Text Encoding Initiative: <http://www.tei-c.org> (last consulted: 23/01/2018)

EAGLES: <http://www.ilc.cnr.it/EAGLES/home.html> (last consulted: 23/01/2018)

Wordnets in the World: <http://globalwordnet.org/wordnets-in-the-world/> (last consulted: 02/02/2018)

Lemon VocBench Custom Forms: <https://bitbucket.org/art-uniroma2/lemon-vb-customforms> (last consulted: 26/01/2018)

TopBraid Composer Maestro Edition: <https://www.topquadrant.com/tools/ide-topbraid-composer-maestro-edition/> (last consulted: 01/02/2018)