These two URIs represent the singular masculine and singular feminine forms of the Spanish word *entendedor*.

- http://kdictionaries.com/id/lexiconES/entendedor-adj-form-1
- http://kdictionaries.com/id/lexiconES/entendedor-adj-form-2

If the dictionary contains two different adjectival endings, as with *entendedor* which has different endings for the feminine and masculine forms (*entendedora* and *entendedor*), and they are not explicitly mentioned, then we use numbers in the URI to describe them. If the gender is explicitly mentioned, then the URIs would be:

- http://kdictionaries.com/id/lexiconES/entendedor-adj-form
- http://kdictionaries.com/id/lexiconES/entendedora-adj-form

In addition, it should be considered that the aim of triplifying the XML was for all these headwords with senses, forms and translations, to connect and be identified and linked following the SW standards.

One of the last steps of complexity was to develop a generic XSLT which can triplify all the different languages of this dictionary series and store the complete data in a triple store. The question remains whether the design of such a universal XSLT is possible while taking into account the differences in languages or the differences in dictionaries.

4. Application and exploration

We tried to investigate also whether the automated resource linking could help with the translation of one dictionary into another the language. Two bilingual dictionaries were considered - English(en)-German(de) and German(de)-English(en).

For the word *bank* the following translations are found:

Bank (de) – *bank* (en) – German to English
bank (en) – *Bank* (de) – English to German

The URI of the translation from German to English was designed to look like:

- http://kdictionaries.com/id/tranSetDE-EN/Bank-n-SE00006116-sense-bank-n-Bank-n-SE00006116-sense-TC00014378-trans

And the one for the translation from English to German would be:

- http://kdictionaries.com/id/tranSetEN-DE/bank-n-SE00006110-sense-Bank-n-bank-n-SE00006110-sense-TC00014370-trans

In this case, both represent the same translation but have different URIs because they were generated from different dictionaries (in accordance with the translation order) that need to be mapped to each other so as to represent the same concept.

The word *Bank* in German can mean either a bench or a bank in English. When either of these English senses is translated back into German the result is the German word *Bank*. It is, however, not possible to determine which sense out of the two was translated unless the URI that contains the sense ID is included. It is also important to maintain the order of translation (source-target) but later map both translations to the same sense and same concept. This is difficult to establish automatically.

5. Future work

The actual overlap and automatic linking of the dictionary resources remains to be tested. There are also some lexicographic elements which were not covered by the new OntolexKD model and need to be added.

There is also the necessity to verify and check for differences between KD's XML dataset and the derived KD's triplified dataset. For this, SPARQL queries need to be created that validate and verify the resulting RDF.

References

Bosque-Gil, J., J. Gracia, and A. Gómez-Pérez. 2016. Linked data in lexicography. *Kernerman Dictionary News* 24, 19-24.

Bosque-Gil, J., J. Gracia, E. Montiel-Ponsoda, and G. Aguado-de Cea. 2016. Modelling multilingual lexicographic resources for the Web of Data: The K Dictionaries case. In *Proceedings of GLOBALEX 2016 Workshop at the 10th Language Resources and Evaluation Conference (LREC 2016), Portorož, (Slovenia)*.

Gracia, J. 2015. Multilingual dictionaries and the Web of Data. *Kernerman Dictionary News* 23, 1-4.

Klimek, B., and M. Brümmer. 2015. Enhancing lexicography with semantic language databases. *Kernerman Dictionary News* 23, 5-10.

TIAD shared task 2017 – Translation Inference Across Dictionaries

The first shared task on Translation Inference Across Dictionaries was aimed to explore best methods and techniques for automatic generation of new bilingual dictionaries based on existing resources. It relied on extracts from 15 bilingual dictionaries of K Dictionaries (KD) for developing three new language pairs that were validated against existing KD data and by human translators.

TIAD 2017 was organized by Noam Ordan, Morris Alper and Ilan Kererman (KD) and Jorge Gracia (OEG, Madrid Politechnic University). The results were presented in a workshop co-located with the Language, Data and Knowledge conference at NUI Galway on June 18, 2017 by four teams:

- Kathrin Donandt, Christian Chiarcos and Maxim Ionov; Goethe University, Frankfurt
- Tom Knorr; Neurocollective, San Francisco CA
- Thomas Proisl, Philipp Heinrich, Stefan Evert and Besim Kabashi; Erlangen University
- Uliana Sentsova; National Research University Higher School of Economics, Moscow

The papers are published as part of the LDK 2017 Workshop Proceedings http://ceur-ws.org.

Noam Ordan

https://tiad2017.wordpress.com/

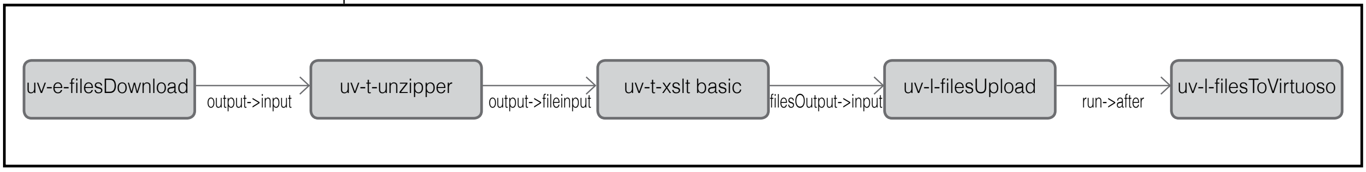

Figure 2: UnifiedViews pipeline used to triplify XML