

Three birds (in the LOD cloud) with one stone: BabelNet, Babelfy and the Wikipedia Bitaxonomy

Tiziano Flati, Andrea Moro and Roberto Navigli
surname@di.uniroma1.it



SAPIENZA
UNIVERSITÀ DI ROMA

We present the current status of **linguistic resources** developed in our research group and published as linked data and linguistic services **in the LOD cloud**, namely **BabelNet** [1], **Babelfy** [2] and the **Wikipedia Bitaxonomy** [3]. We introduce them in terms of their salient aspects and explain how they connect to the world of LOD by means of querying, exploring and exporting data into **RDF format**.



babelnet.org

BabelNet^{2.5}

A very large multilingual encyclopedic dictionary and semantic network

Seamless integration of multilingual encyclopedic and lexicographic knowledge

Java APIs for programmatic access and SPARQL endpoint for querying Linked Data



A semantic network consisting of 9.3 million meanings and 7.7 million images

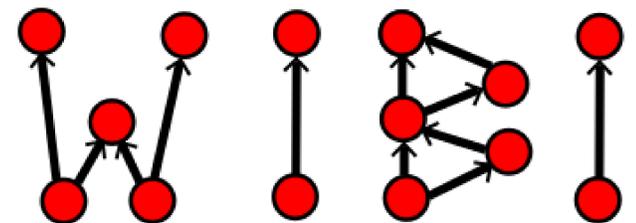
Unified, integrated access to Wikipedia, Wikidata, WordNet, Wiktionary, OmegaWiki and Open Multilingual WordNet

Multiple textual definitions in many languages

50 languages covered: English, Spanish, Chinese, Hindi, Arabic, Portuguese, Russian, Japanese and many others!



wibitaxonomy.org



An integrated taxonomy of about 3.5M Wikipedia pages and 600K categories, aligned to each other

nearly full coverage of pages and categories via a unified, 3-phase approach

State-of-the-art results

on both pages and categories when compared to all available knowledge resources like DBpedia, YAGO, MENTA, WikiNet and WikiTaxonomy

Self-contained resource

(no additional resources or supervision required) and virtual absence of supervision, making WiBi replicable on any new version of Wikipedia

RDF Export facility

Seamless conversion into standard RDF format (turtle, rdf-xml, n-triple)



Unified, multilingual, graph-based approach to Entity Linking and Word Sense Disambiguation

Fares well both on long texts and short sentences

Loose identification of candidate meanings coupled with a densest subgraph heuristic



babelfy.org

fy Babelfy
"Word sense disambiguation and entity linking together!"



RESTful Java API

Programmatically retrieve disambiguated text with a few Java lines!

References:

- [1] Navigli, R., Ponzetto, S.P.: **BabelNet: The automatic construction, evaluation and application of a wide-coverage multilingual semantic network**. Artificial Intelligence 193, 217–250 (2012)
- [2] Moro, A., Raganato, A., Navigli, R.: **Entity Linking meets Word Sense Disambiguation: a Unified Approach**. Transactions of the Association for Computational Linguistics (TACL) 2, 231–244 (2014)
- [3] Flati, T., Vannella, D., Pasini, T., Navigli, R.: **Two Is Bigger (and Better) Than One: the Wikipedia Bitaxonomy Project**. In: Proc. of ACL 2014. pp. 945–955. Baltimore, Maryland



The authors gratefully acknowledge the support of ERC Starting Grant MultiJEDI No. 259234.



The authors also acknowledge support from the LIDER project (No. 610782), a Coordination and Support Action funded by the European Commission under FP7.

mlider