
Claudio Delli Bovi
delibovi@di.uniroma1.it
bn.17381128n

Alessandro Raganato
raganato@di.uniroma1.it
bn.17381127n

http://lcl.uniroma1.it/sew

**SEW: What is it?**
- **SEW** (Semantically Enriched Wikipedia) [6] is a sense-annotated corpus automatically built from Wikipedia by exploiting its hyperlink structure along with the wide-coverage sense inventory of BabelNet [5].
- **SEW** constitutes both a large-scale Wikipedia-based semantic network and a sense-tagged dataset with more than 200 million annotations of over 4 million different concepts and named entities.

**SEW-EMBED: From Explicit to Embedded Representations**

Using SEW to build vector representations for BabelNet senses and Wikipedia pages:

\[
\text{Word Similarity: } \max_{\sigma \in \mathbb{S}} \sigma(s_1, s_2) \quad \text{Given two words } w_1 \text{ and } w_2 \text{, select the closest pair of senses}
\]

\[
\text{Lexical Specificity (LS) [3]: } \text{spec}(T, t, F, f) = -\log_2 P(X > f)
\]

**References**


SemEval-2017 Experiments

Subtask 1: Multilingual Word Similarity

Global score: 0.552 (4th position)

Subtask 2: Cross-Lingual Word Similarity

Global score: 0.558 (3rd position)