Semantic Information Retrieval

**Motivation**

- **Vocabulary Mismatch Problem**
  - Essay: I like baking cakes...
  - Semantic Relatedness: pastries, confectionery, food processing...

**Goals**

- Improve the performance of IR on domain-specific document collections:
  - recall (closing the vocabulary gap)
  - precision (especially for the first 10 ranks)

**Software**

- Darmstadt Knowledge Processing Repository: UIMA components for NLP, IR, and semantic relatedness
- Wikipedia API (JWPL) & Wiktionary API (JWKTL)
- Dextract: Test data creation for semantic relatedness

**Methods**

- Integrating semantic relatedness into IR
- Combining linguistic knowledge sources, e.g. German wordnet, and Web 2.0 knowledge sources, e.g. Wikipedia

**Publications**


**People**

- **Principal investigators**
  - Prof. Dr. Iryna Gurevych
  - Prof. Dr. Max Mühlhäuser

- **Project staff**
  - Christof Müller (coordinator)
  - Torsten Zesch

http://www.ukp.tu-darmstadt.de