

Web Search: Techniques, algorithms and Applications

Eneko Agirre <e.agirre@ehu.edu>
German Rigau <german.rigau@ehu.es>



Content

- Introduction (German Rigau)
 - 2:30h (24/02: 15h - 17:30h)
- Tools and Applications (German Rigau & Eneko Agirre)
 - Tools: 2:30h (26/02: 15h - 17:30h)
 - Applications: 5h (03/03, Wednesday 21: 15h - 17:30h)
- NLP for Ontologies & large-scale Knowledge Bases (Eneko Agirre)
 - 5h (05/03, 10/03: 15h - 17:30h)
- Ontologies & large-scale Knowledge Bases for NLP (German Rigau)
 - 5h (24/03, 26/03: 15h - 17:30h)

Content

- Introduction (German Rigau)
- Tools and Applications
 - Basic Tools (German Rigau)
 - Tokenization
 - Sentence Splitting
 - Language Identifiers
 - Lemmatization, POS tagging
 - Named Entity Recognizers and Categorizers (NERC)
 - Parsing
 - Word Sense Disambiguation (WSD)
 - Semantic Role Labelling (SRL)

Content

- Applications (Eneko Agirre)
 - Information Retrieval (IR)
 - Cross-lingual Information Retrieval (CLIR)
 - Question Answering (QVA)
 - Information Extraction (IE)
 - Classification
 - Summarization
 - Machine Translation

Content

- NLP for ontologies (Eneko Agirre)
 - Ontology learning
 - Terminology extraction
 - Learning hierarchies
 - Learning semantic relations
 - Specific domains: biomedicine

Content

- Ontologies & Knowledge-based for NLP (German Rigau)
 - Words & Works
 - Ontologies:
 - Mikrokosmos
 - Large-scale Knowledge Bases:
 - WordNet & EuroWordNet
 - More large-scale resources
 - ConceptNet, Framenet, VerbNet, PropBank, ...
 - Building Wordnets
 - WordNet extensions:
 - SUMO ontology, eXtended WordNet, Meaning project
 - Reasoning