



Natural Language Processing for the Semantic Web

Dr. Diana Maynard

d.maynard@dc.shef.ac.uk

Dr. Johanna Völker

voelker@informatik.uni-mannheim.de

LREC 2010

Who we are ...

Dr. Diana Maynard

Senior researcher, NLP group at the
University of Sheffield, UK

Member of the GATE development team

<http://www.dcs.shef.ac.uk/~diana/>



Dr. Johanna Völker

Postdoc, Artificial Intelligence group,
University of Mannheim, Germany

Funded by Margarete-von-Wrangell
habilitation grant

http://ki.informatik.uni-mannheim.de/people/johanna_voelker.html



Motivation

- **Semantic Web** requires machine-interpretable semantics in order to process textual information on the internet
- How do we get **machine-interpretable semantics**?
- **NLP techniques** such as information extraction, disambiguation, term recognition etc. give meaning to unstructured text
- When this meaning is linked to an ontology, it (a) becomes **reusable** across the Semantic Web and (b) enables processes such as **reasoning** to be carried out
- Ultimately, **Open Linked Data** enables this information to be exposed, shared and connected via dereferencable URIs
- All this leads eventually to the Holy Grail of **knowledge understanding**

Challenges

- The **web** is huge and constantly growing
- Only (semi-)automatic methods for adding meaning to the **unstructured information** are thus feasible
- Is NLP technology **robust** and **scalable** enough to cope with these demands?
- How can **NLP experts** understand ontologies and domain experts and provide relevant tools in the face of constantly changing needs?
- How can **ontologists** make use of NLP toolkits without requiring NLP expertise?

Aims of this tutorial

- Introduce the essential concepts of **ontologies and the semantic web** to NLP and language specialists
- Describe methodologies for manual and automatic **ontology construction** and **population**
- Show how traditional NLP techniques such as **information extraction** can be “made semantic”
- Use GATE as an example NLP toolkit to demonstrate real life **applications** for semantic web development

<http://www.dcs.shef.ac.uk/~diana/courses/lrec-nlp-semweb-tutorial.html>

Schedule

From	To	
09:00	09:15	Introduction
09:15	10:00	Ontologies and the Semantic Web
10:00	10:30	GATE and Controlled Language Ontology Editing
10:30	11:00	Coffee Break
11:00	12:00	Semantic Annotation and Evaluation
12:00	12:45	Ontology Learning and Information Extraction
12:45	13:00	Conclusion