

Advanced Topics:

Information Retrieval

Natural Language Processing

WinterTerm 2015/2016

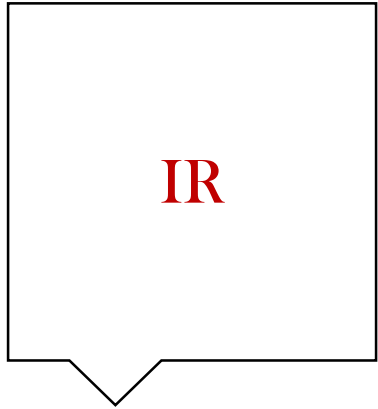
Web Technology and Information Systems

Bauhaus-Universität Weimar
Faculty of Media



NLP

- The branch of **computer science** focuses on developing **systems** that allow **computers** to **communicate** with **human natural language**.
- The process of computer analysis of **input** provided in a **natural language**, and **conversion** of this input into a useful form of **representation**.
- Application: e.g., **Machine Translation**.



- Information Retrieval (IR) is **finding material** (usually documents) of an **unstructured** nature (usually text) that satisfies an **information need** from within **large collections** (usually stored on computers).
- Application: e.g., **Search Engine**.

What can
we do with
NLP and
IR?

**YOU
CAN CHANGE
THE WORLD**



Selected Topics

- Discourse Analysis.
- Sentiment Analysis.
- Argumentation Mining.
- Word Embedding.
- Automated Essay Scoring.
- Clustering of News Articles.
- Relation Extraction.
- Textual Entailment.
- Wikipedia in IR and NLP.
- Multi-document Summarization.
- Computational Ethics.
- Known-item Refinding.
- Justification for Recommendations.



Seminar Deliverables

Choosing a Topic

Topic Presentation
30 min

Paper Presentation
45 min

Lab Presentation
30 min

Topic

Topic Presentation

- Definition
- Examples
- Models
- Applications

Paper Presentation

- The Problem
- Motivation
- Main contribution
- Method
- Exp. and Results
- Future work

Lab Presentation

- The Problem
- Implementation
- Exp. and Results



Policy

- Attendance is obligatory.
- Each student should give three talks.
- Simple implementation of IR/NLP method.
- Grad is based on your performance in the lab and presentation.



Organization

- Next meeting:
 - Our group will give you short introductions about IR and NLP.
 - You will have three topics from the list you are interested in.
- Next ...
 - Topic presentations.
 - Paper presentations.
 - Lab presentations.
- Dates for the seminar talks are to be determined.



Questions