

Seminar Topics: Information Extraction

English topics!

Viktor Hangya

hangyav@cis.lmu.de

▶ Topic: **Deep Learning for Emerging Named Entity Recognition**

▶ Overview:

- ▶ NER systems perform well on data similar to what they were trained on and can detect frequent well behaving NEs.
- ▶ New NEs emerge day-by-day which are often hard to detect for humans as well.
- ▶ Give a brief overview of the WNUT2017 shared task. How was the data created?
- ▶ Describe a participating system and give an analysis of the results. What were the main challenges?

▶ Suggested papers:

- ▶ **Results of the WNUT2017 Shared Task on Novel and Emerging Entity Recognition** Derczynski et al., 2017, *Proceedings of the 3rd Workshop on Noisy User-generated Text*
- ▶ **Distributed Representation, LDA Topic Modelling and Deep Learning for Emerging Named Entity Recognition from Social Media** Jansson and Liu, 2017, *Proceedings of the 3rd Workshop on Noisy User-generated Text*

- ▶ Topic: **Information Extraction from CVs**
- ▶ Overview:
 - ▶ Many companies receive a lot of CVs which have to be grouped by certain aspects and forwarded to the right departments.
 - ▶ Although they have some structure, it varies CV-by-CV making it harder to extract certain information.
 - ▶ Describe the task: motivation, data, systems, results?
- ▶ Suggested paper:
 - ▶ **Resume Information Extraction with Cascaded Hybrid Model** Yu et al., 2005, *Proceedings of the 43rd Annual Meeting of the ACL*

- ▶ Topic: **User Opinion Extraction**
- ▶ Overview:
 - ▶ People express their opinions about various entities using social media which are important information to companies and service providers.
 - ▶ The task is to extract opinion expressions and entities or their aspects, and decide if they are related.
 - ▶ Introduce a system performing opinion extraction. How can different information be extracted and their relationship decided? What level of performance can be achieved?
- ▶ Suggested paper:
 - ▶ **Opinion Mining on the Web by Extracting Subject-Aspect-Evaluation Relations** Kobayashi et al., 2006,
American Association for Artificial Intelligence

- ▶ Topic: **Clinical Temporal Relation Extraction with Neural Networks**
- ▶ Overview:
 - ▶ Temporal relation extraction is the task of extracting dates of certain events.
 - ▶ Introduce the developed neural system for the task and present the achieved results.
 - ▶ Based on the results what conclusions can be drawn?
- ▶ Suggested paper:
 - ▶ **Investigating the Challenges of Temporal Relation Extraction from Clinical Text** Galvan et al., 2016,
Proceedings of the 9th International Workshop on Health Text Mining and Information Analysis

▶ Topic: **Knowledge Base Completion**

▶ Overview:

- ▶ IE techniques can be used to build huge knowledge graphs containing various attributes of entities and their relations. This knowledge can be exploited for various downstream tasks.
- ▶ Most of the existing KGs are incomplete: entities are missing attributes and relations.
- ▶ How can we fill missing information in KGs? What is the relationship between the task and question answering systems?

▶ Suggested paper:

- ▶ **Knowledge Base Completion via Search-Based Question Answering** West et al., 2014,
Proceedings of the 23rd international conference on World wide web

- ▶ Topic: **Question Answering using KGs**
- ▶ Overview:
 - ▶ Answer natural language questions using the information extracted with various IE methods.
 - ▶ How can a system understand a question and use KGs to find the answers?
 - ▶ Introduce a neural model for the task!
- ▶ Suggested paper:
 - ▶ **Question Answering over Freebase with Multi-Column Convolutional Neural Networks** Dong et al., 2014,
Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing