Oliver Ferschke und Iryna Gurevych: **Extracting Arguments from Web Discourse**

Public discourse on the web has many faces. From online editions of newspapers to the discussions of anonymous web users, there is a broad spectrum of knowledge sources covering current events with a varying degree of objectivity and reliability. In this talk, we will focus on two different types of web discourse arising from two different domains.

First, we will address the analysis of user discussions on Wikipedia.
In the absence of an editorial board, Wikipedia users have to decide on their own how an article should be further developed and what particular content should be included. The discussion pages give valuable insights regarding the arguments that users provide in this context. We created corpora of manually annotated discussion pages, analyzed the discourse and trained classifiers for automatically identifying dialog acts aimed at article improvement.

Second, we will discuss an ongoing project on argument extraction from online articles from the educational domain. We created a corpus of articles from online newspapers covering current educational topics. In a two-step approach, we first identify arguments in these texts and then annotate them according to a newly designed annotation schema. With the resulting data we plan to train an argument extraction system for augmenting a search engine for the educational domain.