

Oliver Ferschke. "The Quality of Massive Open Online Collaboration", Herrenhausen Conference (Digital) Humanities Revisited – Challenges and Opportunities in the Digital Age, Hanover, Germany, 5.-7-12-2013

## **Abstract**

User generated content is the main driving force of the increasingly social web. Participatory and collaborative content production has largely replaced the traditional ways of information sharing and make up a large share of the daily information consumed by web users.

The main properties of user generated content are a low publication threshold and little or no editorial control. While this has positively affected the variety and timeliness of the available information, it causes an even higher variance in quality than the already heterogeneous quality of traditional web content.

In this project, we focus on the quality of collaboratively created texts. Using the example of Wikipedia, we investigate how the quality of articles can be assessed automatically and how we can apply language technology to facilitate quality assurance on a large scale.

In a first scenario, we analyze two corpora of Wikipedia article discussion pages extracted from the English Wikipedia and the Simple English Wikipedia, which we manually annotated with quality-directed speech act labels. We show how these corpora can be used to automatically identify the problems and solutions discussed by the community. We finally discuss possibilities how this approach can help to improve work coordination in Wikipedia and ultimately improve the quality assurance process.

In a second scenario, we focus on analyzing the article quality directly. Instead of applying abstract quality scores to each article, we approach the problem from a different direction and aim to identify concrete quality problems. Wikipedia already provides a rich set of quality flaw markers, which we extract on a large scale and use as training data for automatic quality flaw prediction, which can assist authors in improving the quality of their articles. In this context, we furthermore analyze the topic prevalence of individual flaw types, i.e. the phenomenon that particular flaws appear more often in articles from certain topics. This biased distribution negatively influences the data samples used for machine learning experiments and thus warrants further investigation.