Preface

Annotation in Digital Humanities (annDH) is held on August 6-10, 2018 in Sofia, Bulgaria, as part of the 30th European Summer School in Logic, Language and Information (ESSLLI 2018).

The main topic of the workshop can be rephrased as: How can linguistics/computational linguistics help with annotation in DH?

Linguistic annotation is one of the core interfaces between linguistics and computational linguistics. It has also become a central interface between computational linguistics (CL) and digital humanities (DH). Texts are preprocessed and annotated, e.g. with parts of speech, for distant reading and other visualization applications, topic and network analyses, text mining and question answering for humanist research questions. In these applications the annotation is a means to an end and mostly invisible to the humanist researchers.

The annDH workshop makes an attempt to push the boundary of this interface and focus on annotation beyond the standard linguistic categories, looking at categories and relations relevant for humanist research questions themselves, such as metaphors, stereotypes, entities, causation of historical events, narratives, or philosophical reasoning. In this area, CL cannot necessarily provide tools, but instead it can provide methodology and best practices. Thus, lessons learned in linguistic annotation can be repurposed for annotation in DH. This includes CL support of the epistemological process of developing the annotation categories themselves, which are often inductively—or abductively—derived in a hermeneutically cyclic way. Also included in the scope of the workshop is research on the data types in the digital humanities, which often concern non-canonical language in the sense of historical language, literary style and other registers with idiosyncratic properties not commony dealt with in CL—and thus pose challenges for automated annotation.

The workshop had eleven submissions, out of which ten are presented at the workshop. Additionally the workshop comprises two discussion sections and an invited talk by Thierry Declerck on "Ontology-driven Annotation of Literary Texts". The topics of the presentations range from considerations concerning a generalized annotation tool via linguistic annotations for digital humanities projects to more specialized annotation tasks such as argument components, cultural artifacts, and quality of interventions. While most contributions focus on one specific annotation task, we hope that by bringing this audience together, we can foster a dialogue that will allow us to come up with guidelines for best practices for the robust annotation of a wide range of phenomena that go beyond simple categories in one way or another.

We would like to thank all members of the program committee of annDH for providing profound feedback to the authors.

August 6-10, 2018   Sandra Kübler
Sofia   Heike Zinsmeister