PhD seminar:

"Introduction to the semantic annotation platform INCEpTION"

In the context of the PhD program in *Applied Linguistics: Argumentation in Professional Practice*

ECTS: 1
Number of participants: max. 20 (full!)

Program:

Two half days via ZOOM:

Wednesday, June 9, 2021, 8:30 to 12:30
Thursday, June 10, 2021, 8:30 to 12:30

Topic, focus:

The investigation of specific phenomena in written texts often requires compiling task-specific corpora and creating or extending task-specific knowledge bases. Researchers profit from automatic, semi-automatic, and manual annotation of linguistic structures and semantic or argumentative information. Presently, researchers require a broad range of skills and tools to address such annotation tasks, especially when annotation is an essential part of the research process and used for data-driven investigation. Explorative manual annotation serves as a starting point for the development of systematic classifications and annotation schemes. Multi-layer and concurrent annotation of elements and linking between elements and layers help to gain deeper insights and support modeling of complex phenomena.

Manual annotations for a small set of texts can also serve as input for machine learning algorithms to allow for systematic annotation of larger collections to test the validity of the classification system.

Main value added from a theoretical and methodological perspective:

Multi-layer annotation is a prerequisite for comprehensive and hermeneutic modeling of large text collections. The combination of automatic basic linguistic annotation (e.g., Part-of-Speech) and manual data-driven annotation (i.e., marking, labeling, and linking words and phrases), allows researchers to develop scientifically sound perspectives. Explicit machine-readable annotation is also a prerequisite for reproducible research.

Main value added from a practical perspective:

INCEpTION is a text-annotation environment for various kinds of annotation tasks on written text. INCEpTION is a state-of-the-art open-source web application. It allows several users to work on the same annotation project and it can handle several annotation projects at a time. It provides a recommender system to help you create annotations faster and easier.

Learning objective:

Participants are familiar with the concept of entities and relations in INCEpTION and understand the general workflow of annotation.

Participants are familiar with the concept of multi-layer annotation.

Participants can use the platform to upload raw and pre-annotated text collections and apply standard automatic linguistic annotation (e.g., POS, NER).

Participants can use the framework to model the phenomena in their own research using multiple annotation layers, i.e., existing ones and newly created ones.

Evaluation: Attendance to the whole series is required in order to obtain credits. Each seminar will involve an in-course activity which will be assessed with a pass/ fail evaluation.

Bio notes/ Profiles

Ute Winchenbach is a collaborator/developer of the INCEpTION project https://inception-project.github.io at UKP Lab at TU Darmstadt. She holds an M.Sc. in Computer Science from TU Darmstadt.