The INCEpTION TEXT Annotation Platform

Richard Eckart de Castilho, Jan-Christoph Kile, Ute Winchenbach, Iryna Gyurevych

Ubiquitous Knowledge Processing Lab (UKP-TUDA)
Dept. of Computer Science
Technische Universität Darmstadt

INCEpTION Platform

INCEpTION aims to support three functionalities which are commonly required for text annotation projects but typically not available in a single tool: corpus creation, text annotation, knowledge management.

The platform additionally provides assistive features such as machine learning recommenders to help users working on these tasks more efficiently.

Integrating these functionalities into a single comprehensive platform permits addressing tasks typically not found in generic annotation platforms, such as entity linking, knowledge base population, cross-document coreference annotation, etc.

Open Development

While most annotation tools are built in annotation projects, INCEpTION is an infrastructure software project and is not associated with any single annotation project. Acquiring early adopters and aligning with their use-cases is a key part of our mission.

This motivates our open development philosophy:

- All code is open and publicly available on GitHub under the liberal Apache License 2.0
- All development-related tasks and issues are publicly managed and discussed via GitHub
- Internal and private communication is kept at a minimum

Interoperability and Integration

Within the NLP and text mining landscapes, an annotation platform like INCEpTION only covers a part of the overall text analysis needs. Therefore, it is important that the platform is open and interoperable with external services and resources.

Integration goes beyond interoperability. E.g., when an external text mining platform wants to delegate annotation to the INCEpTION platform, it needs to be able to automatically set up annotation projects, data import, monitor the ongoing annotations and finally retrieve the annotated data for further use.