

Natural Language Processing for Multilingual Legal Analytics

Research project and work plan

Project ADELE (DG JUSTICE, Grant Agreement 101007420) is premised on the ongoing paradigm shift towards cognitive computing and human-centered AI which is transforming many socio-economic activities, including justice. The project applies legal analytics (LA) – a blend of data science, machine learning and natural language processing techniques – to judicial decisions: it aims to (1) develop methods to extract knowledge and engage in outcome predictions, and therefrom (2) build a pilot tool to support legal research and decision-making processes in the judiciary. The project will begin with the analysis of four datasets of judgments respectively of Italian and Bulgarian case law in the fields of Intellectual Property, and Unfair Commercial Practices (WP2). Annotation guidelines will be developed and used to tag the datasets. Domain-specific ontologies and citation networks will be built. In WP3 tagged documents will be processed with machine learning algorithms to extract legal arguments and predict possible outcome of new cases. Two German datasets will be collected and used as a language transferability test. WP4 will develop a pilot tool embedding knowledge extraction and outcome prediction functionalities. The tool will be subject to (1) a preliminary test on a beta-version of the tool; (2) a validation test on the final pilot (WP5). Tests will take place both in Bulgaria and Italy with specialized end-users. The main outputs of ADELE will be (1) a LA methodological framework and (2) a pilot tool to support activities of justices in the Italian and Bulgarian courts. These novel solutions will have a strong impact on the efficacy and transparency of judicial decision-making by developing automated means that enhance the understanding of the law, facilitate judicial tasks and reduce the courts' workload. It will contribute to developing a trustworthy AI by producing ethical and robust LA methods for the judiciary and will foster scientific debate on the use of AI technologies in the legal field.

The objective of the present position is to carry out activities in the context of ADELE's WP3 and to coordinate with the related activities carried out in the other WPs. In particular, the work will include working with and extending AI methods to extract knowledge and engage in outcome predictions, working with the multi-lingual datasets provided by the other partners in the ADELE consortium, and contributing to the definition of such datasets and to the integration of the methods/algorithms into the pilot tools.

All activities are to be carried out at the University of Bologna's Department of Computer Science and Engineering (DISI) and in collaboration with the other ADELE consortium partners, in particular: Università degli Studi di Torino, Università di Lussemburgo, European University Institute, Sayuz Yuristite v Bulgaria, Bondatsiya Libre, and Apis Europe.