

Natural Language Processing for the analysis of judicial principles of law

Research project and work plan

This research project builds upon the data-and-AI paradigm shift that has taken place in the last two decades: Artificial Intelligence (AI) technologies, as applied to vast datasets (big data), have already transformed many aspects of our economic and social life, delivering many useful applications (smart assistants, self-driving cars, healthcare robots etc). The area of law and justice is also being affected, as new powerful systems become available to practitioners for enhancing legal cognition and practice. In this context, the emerging field of legal analytics (LA) develops AI, natural language processing (NLP) and machine learning applications to extract legal knowledge, infer undiscovered relations, and engage in data-driven predictions. The following lines of research have been pursued:

- Identification and representation of legal knowledge, defining types of annotations, concepts, rules, links and their integration with conceptual ontologies, providing visual maps linking cases and concepts.
- Machine learning with legal texts, managing complex unstructured datasets, classifying paragraphs of texts, extracting legal rules, comparing rules across jurisdictions.
- Extraction of arguments from cases, including argument schemes, claims, substantive legal factors So far, applications have been developed generally as commercial initiatives for the private sector.

The main objective of this research activity is to build a system able to process a decision or a tax audit act and produce a customised legal assessment on the relevant judicial principles of law in the document at hand. This would enable research, carried out in a team of AI experts and legal experts in the context of the POLINE European project funded under the JUST-2022-EJUSTICE Call (“Principles Of Law In National and European VAT”), on the correspondence of such principles with those existing at national and EU level or the selection of more relevant or more recent principles. The ultimate objective of POLINE is to increase European taxpayers’ awareness of the correct application of VAT law, support them in deciding whether it is making an appeal or claiming a violation of European law, and empower citizens, by facilitating their active role in the enforcement of VAT law and pursuit of a fair taxation.

The methodology foresees the design of a supervised classification system, which operates through a training set and a test set whose development is also part of the project activities. In the first months of the project, the research will focus on the development of the dataset and on the study of the state of the art in technologies enabling the detection of principles of law in legal decisions, and the measure of similarity between principles of law, considering lexical and semantic features. Selected and possibly extended semantic text similarity techniques will then be used to automatically identify similarities and divergences between

principles of law. The last phases of the project will focus on evaluation and integration of the developed solution with the other methods developed by other partners of the project.

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 POLINE consortium partners, in particular: Università degli Studi di Torino, Orebro University, European University Institute, and Apis Europe.