

## **Abstract for the general public**

If big data is the new oil, then loss of privacy is the new climate crisis: in the age of personal data being collected and processed on an unprecedented scale, the individual is faced with an ever-growing power of big business and the opaque decision-making systems it employs. Founded on the premise of countervailing power voiced by Ken Galbraith, the project seeks to prove that artificial intelligence can be used to empower the modern-day citizen by providing automated compliance analysis of legal text, such as terms of service or privacy policies, performed in a transparent manner by XAI (eXplainable Artificial Intelligence). This analysis would also include a qualitative assessment of the clarity of expression, thus accounting for the required linguistic transparency dimension in the overall assessment of legal compliance. This research can become a significant step towards developing a technology freely available to all, in the form of mobile apps, Internet browser plugins or standalone desktop applications that will allow a quick, easy and transparent user-level compliance checking of any electronic terms of service contract or privacy policy of a service provider, providing simple and reliable explanations regarding all compliance issues which have been identified.

The present research is based upon the experience gathered by the Principal Investigator over the course of the projects developed over the years 2018-2019 under the Claudette framework (a portmanteau for [unfair] Clause Detector) and hosted at the European University Institute. Born out of a series of experiments on training a machine-learning classifier to recognize unfair contract terms, the Claudette project has developed into a vast framework encompassing the analysis of both terms of service and privacy policies for compliance with the relevant law (in particular, the Unfair Terms Directive and the GDPR). This research has also demonstrated that not only is it possible to train a machine-learning system to provide explanations to its unfairness assessments; using a neural network which draws from a knowledge base of possible legal explanations actually improves the efficiency of identifying unfair clauses in contracts. Significant progress was also made towards making the system useful to citizens across Europe by experimenting with different methods of rendering the system capable of supporting multiple languages, both by utilizing the existing training corpora and through semi-automated creation of new ones.

Building upon this basis, the planned research will make two considerable steps forward. First, in order to demonstrate that XAI can be used for transparent assessment of compliance of terms of service and privacy policies, knowledge bases will be built on the basis of the normative framework developed on the basis of the Unfair Terms Directive and the General Data Protection Regulation, including all official guidelines and relevant case law of the European Court of Justice. These knowledge bases will then be used to test the possibility of building a fully explainable compliance-testing AI system. Second, in order to analyse and quantify the aspects of linguistic transparency, to be included in the overall compliance assessment, the project will attempt to reach three main goals: a) Analysis of the concept of transparency against the background of the UCTD and the GDPR and elaboration (against this background) of the set of criteria for distinguishing between transparent and non-transparent pieces of contractual design; b) identify the most accurate framework for linguistic analysis of readability in regard to terms of service and privacy policies and attempt to translate the selected method into a machine-learning algorithm; c) empirical analysis of a representative sample of clauses from terms of service and from privacy policies.