

XAILA 2021 - the Fourth Explainable & Responsible AI in Law (XAILA) Workshop at ICAIL2021 - the 18th International Conference on Artificial Intelligence and Law, São Paulo, Brazil

Preface

The 4th International Workshop on *eXplainable and Responsible AI and Law (XAILA2021@ICAIL)* was held at the 18th International Conference on Artificial Intelligence and Law (ICAIL 2021) organized by the Law School of the University of São Paulo, Brazil (entirely online), on 21 June 2021. The idea of the XAILA series of workshops (1st edition XAILA at JURIX 2018 in Groningen, 2nd edition XAILA at JURIX 2019 in Madrid, 3rd edition XAILA at JURIX 2020 in Brno (online)) is to provide an interdisciplinary platform for the discussion of ideas with respect to explainable AI, algorithmic transparency, comprehensiveness, interpretability and related topics.

This edition of XAILA was the first one to accompany the ICAIL conference and one among the 11 workshops that were attached to ICAIL 2021. This edition of the workshop attracted considerable attention from the community with 853 registered participants (in comparison, the main conference attracted 1240 registrations). We are convinced that these numbers show that the problems relevant for the workshop are perceived as not only as one of the most theoretically significant ones, but also as one of the most pressing ones on the practical side. In particular, the problem of adequate legal regulation of AI is acknowledged as one of the most prominent issues, in European Union and elsewhere. For obvious reasons, the development of such a regulatory framework is an extremely complex issue, and it requires a deep understanding of how AI systems work.

Consequently, the XAILA workshop, rooted in the AI and Law community, attains even more interdisciplinary character. It aims to join the normative perspective of legal theory and ethics on the one hand, and the formal and computational approach represented by AI and Law research and practice. What is more, the problems of explainability, transparency and understandability of computational legal systems create a natural platform to integrate the classical symbolic approach and the more recent computational approach, the latter being represented first and foremost by the machine learning models. One of the most important issues for the AI and Law community is to explore the balance of advantages and

disadvantages of the two approaches while developing systems supporting the work of lawyers and enriching our understanding of legal reasoning. However, the interests of the XAILA community go beyond the legal applications of AI, while the problems of transparency, explainability and responsibility, and the adequate regulation of these issues relevant for any context intelligent systems operation. The embracing of this broader involves, again, the creation of stronger links between different actors. The communication between the practically oriented field of engineering and the academic area of research in AI and Law is now hindered by the diverging terminology used by these communities and the identification of different purposes.

We may identify at least six fields which would potentially benefit from the increase of the flow of information between them in connection with the research and practice on explainable and responsible AI. One of them is the AI and Law research with its tradition of interdisciplinary approach towards the modeling of legal reasoning. Second, the perspective brought by the general research on AI should be taken into consideration, and properly accommodated to the purposes of the legal community. Third, we should mention the quickly developing sector of Legal Tech, focused on supplying legal practice with intelligent solutions aiming at the optimization of legal tasks performance. Naturally, the broadly understood field of legal practice, as the recipient of the said solutions, is the fourth relevant actor. It should be stressed that this field is very diverse as it encompasses not only the commercial sector of law entrepreneurship, but also the public administration, the judiciary and the legislation. Each of these subfields have their specific needs and purposes, which moreover vary from jurisdiction to jurisdiction. Fifth, the academic field of legal theory and philosophy may offer an important contribution to the ongoing discussion and explainable AI and Law, as this area of research is concerned with the elaboration of models of legal reasoning and the discussion of the most general legal concepts, such as fairness, liability and responsibility, important for the sake of development of the regulatory frameworks. Sixth, in our opinion the discussion on explainable AI and Law could benefit from the insights brought by cognitive science research. The latter has already established connections with legal theory which resulted in interesting research results concerning the mechanisms of legal decision making. This research may be of particular relevance for investigation of human-computer interaction in connection with the use of intelligent support tools.

The workshop program included two presentations by invited speakers and five by authors presenting their research.

Our first invited speaker was Wojciech Wiewiórowski, who gave a lecture entitled 'Data protection aspects of Explainability of AI: Between transparency and fairness'. He is acting as the EU's European Data Protection Supervisor, and Adjunct professor in the Faculty of Law and Administration of the University of Gdańsk. He provided the audience with insightful, inspiring and up-to-date perspectives on the developing regulatory perspective on AI in the European Union, and how it is influencing the AI industry and research.

The second invited speaker was Katie Atkinson, with a lecture entitled 'The Landscape and Challenges for Explainability in AI and Law'. She is Professor of Computer Science and Dean of the School of Electrical Engineering, Electronics and Computer Science at the University of Liverpool. She clearly and comprehensively showed how the field AI & Law has been developing explainable AI methods all along, and why ever more impact of research in AI & Law for legal practice and AI generally can be expected.

The first paper presentation was by Trevor Bench-Capon (University of Liverpool) on the paper entitled 'Using Issues to Explain Legal Decisions'. In the paper, he explains how traditional AI & Law approaches, and in particular those based in case-based reasoning approaches, are relevant for machine learning outcome prediction methods.

The second paper was presented by Salvatore Sapienza (University of Bologna), presenting 'Explanations in Risk Analysis: Responsibility, Trust and the Precautionary Principle'. He emphasises the need for reliability and trustworthiness in the high impact domain of health risk analysis, using the regulation of food safety in Europe as a case study.

The third paper 'Socially Responsible Virtual Assistant for Privacy Protection: Implementing Trustworthy' was written by Alžběta Krausová (Czech Academy of Sciences) and seven coauthors. In the paper, the management of privacy settings is used as an example setting where a virtual assistant can support a person's rights and self-determination.

The fourth paper by Davide Carneiro (Politécnico do Porto) and four coauthors is entitled 'A Conversational Interface for interacting with Machine Learning models'. In the paper, a conversational chatbot is used to approach the need for better scrutiny and accountability of machine learning AI methods. An analysis of legal and ethical considerations is used as a starting point.

The fifth presentation by Cor Steging (University of Groningen) and two coauthors was on the paper 'Discovering the Rationale of Decisions: Experiments on Aligning Learning and Reasoning'. The paper presents a knowledge-driven method for evaluating and adjusting the rationale used by black box machine learning systems.

The workshop organizers would like to thank the Program Committee members for their work in the review process. We are also grateful to the extensive efforts of the ICAIL 2021 organization, especially hard in these times of hybrid conference organization. We are happy to furthermore thank the invited speakers, the authors of papers and all participants to the workshop.

Thanks to all of you, the meeting was stimulating and thought-provoking, and we hope to meet many of you at a future edition of the workshop series.

Program committee

Michał Araszkiwicz, Jagiellonian University in Kraków, Poland
Martin Atzmueller, Osnabrück University, Germany
Floris Bex, Utrecht University, The Netherlands
Jörg Cassens, University of Hildesheim, Germany
Enrico Francesconi, IGSG-CNR, Italy
Grzegorz J. Nalepa, Jagiellonian University in Kraków, Poland
Jose Palma, University of Murcia, Spain
Monica Palmirani, CIRSFD, Italy
Juan Pavón, Universidad Complutense de Madrid, Spain
Radim Polcák, Masaryk University, Czechia
Marie Postma, Tilburg University, The Netherlands
V́ctor Rodríguez Doncel, Universidad Polit́cnica de Madrid, Spain
Ken Satoh, National Institute of Informatics and Sokendai, Japan
Jaromir Savelka, Carnegie Mellon University, USA
Piotr Skrzypczynski, Poznan University of Technology, Poland
Bart Verheij, University of Groningen, The Netherlands

XAILA 2021 Organizing Committee

Grzegorz J. Nalepa, Jagiellonian University in Kraków, Poland
Bart Verheij, University of Groningen, the Netherlands

Michał Araszkiewicz, Jagiellonian University in Kraków, Poland
Martin Atzmueller, Osnabrück University, Germany