

## The EXplainable AI in Law Workshop XAILA 2018

<http://xaila.geist.re>

at Jurix 2018 <http://jurix2018.ai.rug.nl>

**Organizers:** Grzegorz J. Nalepa, Martin Atzmueller, Michał Araszkiewicz, Paulo Novais

**Abstract:** Humanized AI emphasizes transparency and explainability in AI systems. These perspectives have an important ethical dimension, that is most often analyzed by philosophers. However, in order for it to be fruitful for AI engineers, it has to be properly focused. The intersection of Law and AI that makes it possible, as it provides a conceptual framework for ethical concepts and values in AI systems. A significant part of AI and Law research during the last two decades was devoted to operationalization of legal thinking with values. These results may now be reconsidered in a broader context, concerning the development of HAI systems and their social impact. It is a timely issue for the AI and Law community.

### **Motivation for the workshop and description**

Humanized AI (HAI) includes important perspectives in AI systems, including transparency and explainability (XAI). Another one is the affective computing paradigm. These perspectives have an important ethical dimension. While ethical discussion is conducted by many philosophers, in order for it to be fruitful for engineers in AI, it has to be properly focused with specific concepts and operationalized.

We believe, that it is the intersection of Law and AI that makes such an endeavor possible. Together, this lays foundations and provides a conceptual framework for ethical concepts and values in AI systems. Therefore, when discussing ethical consequences and considerations of transparent and explainable AI systems, including affective systems, we should focus on the legal conceptual framework. A significant part of AI and Law research during the last two decades was devoted to operationalization of legal thinking with values. These results may now be reconsidered in a broader context, concerning the development of XAI systems and their social impact. As such it is a very timely issue for the AI and Law community.

Our objective is to bring people from AI interested in XAI/HAI topics (possibly with broader background than just engineering) and create an ample space for discussion with people from the field of legal scholarship and/or legal practice. As many members of the AI and Law community join both perspectives, the JURIX conference should be assessed as perfect venue for the workshop. Together we would like to address some questions like:

- non-functional design choices for explainable and transparent AI systems (including legal requirements)
- legal requirements for AI systems in specific domains
- legal consequences of black-box AI systems
- legal criteria for explainable and transparent AI systems
- possible applications of XAI systems in the area of legal policy deliberation, legal practice, teaching and research
- ethical and legal implications of the use of AI systems in different spheres of societal life
- the relation of XAI and argumentation technologies
- XAI models and architectures
- the understanding of the notions of explanation and transparency in XAI
- risk-based approach to analysis of AI systems and the influence of XAI on risk assessment
- incorporating ethical values into AI systems and the legal interpretation and consequences of this process
- XAI, privacy and data protection
- possible legal aspects and consequences of affective systems
- legal requirements and risks in AI applications
- XAI, certification and compliance

**List of members of the program committee (tentative):**

Martin Atzmueller, Tilburg University, The Netherlands  
Michał Araszkievicz, Jagiellonian University, Poland  
Kevin Ashley, University of Pittsburgh, USA  
Szymon Bobek, AGH University, Poland  
Jörg Cassens, University of Hildesheim, Germany  
David Camacho, Universidad Autonoma de Madrid, Spain  
Pompeu Casanovas, Universitat Autonoma de Barcelona, Spain  
Colette Cuijpers, Tilburg University, The Netherlands  
Rafał Michalczak, Jagiellonian University, Poland  
Teresa Moreira, University of Minho Braga, Portugal  
Paulo Novais, University of Minho Braga, Portugal  
Grzegorz J. Nalepa, AGH University, Jagiellonian University, Poland  
Tiago Oliveira, National Institute of Informatics, Japan  
Martijn von Otterlo, Tilburg University, The Netherlands  
Adrian Paschke, Freie Universität Berlin, Germany  
Jose Palma, Univesidad de Murcia, Spain  
Monica Palmirani, Università di Bologna, Italy  
Radim Polčák, Masaryk University, Czech Republic  
Marie Postma, Tilburg University, The Netherlands  
Juan Pavón, Universidad Complutense de Madrid, Spain  
Ken Satoh, National Institute of Informatics, Japan  
Erich Schweighofer, University of Vienna, Austria  
Piotr Skrzypczyński, Poznań University of Technology, Poland  
Dominik Ślęzak, Warsaw University, Poland  
Michal Valco, University of Presov, Slovakia  
Tomasz Żurek, Maria Curie-Skłodowska University of Lublin, Poland

**Important dates:**

Submission: 23.11.2018  
Notification: 30.11.2018  
Camera-ready: 07.12.2018  
Workshop: 12.12.2018

**Submission and proceedings:**

Please submit using the dedicated EasyChair installation

<https://easychair.org/conferences/?conf=xaila2018>

We accept long (8 pages) and short (4 pages) papers in PDF. Please use the IOS Press format.

<http://www.iospress.nl/service/authors/latex-and-word-tools-for-book-authors/>

Workshop proceedings will be made available by CEUR-WS.

A post workshop journal publication is considered.