

QUALITY AND ADVANCES OF XAI

Warsaw University
of Technology



**WEDNESDAY & THURSDAY,
SEPTEMBER 21-22**

Warsaw University of Technology (WUT)

PHD SCHOOL



This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 860621.

QUALITY AND ADVANCES OF XAI



DAY 1

Wednesday, September 21, 9:15–22:00



9:15–10:45 **ON THE NOTION OF EXPLANATION IN THE
CONTEXT OF XAI**

*Marcin Będkowski / Assistant Professor, Institute of Polish
Language, University of Warsaw*

Any sufficiently advanced technology is indistinguishable from magic"... unless it has been explained. The aim of the talk is to provide an answer to the question: What does it mean to explain something (especially an AI model)? In the first part of the talk, the general notion of explanation from the perspective of the methodology and philosophy of science will be discussed. Explanation will be described in relation to other types of reasoning, various demarcation problems (concerning, for example, the distinction between explanation and argumentation or interpretation), and typical cognitive biases associated with this type of reasoning. In the second part, examples and selected problems related to the explanation of AI models as well as building transparent and trustworthy systems will be presented.

10:45–11:15 **COFFEE BREAK**



11:15–12:45 **ARGUMENTATION, ETHOS AND DIALOGUE**

*Marcin Koszowy / Professor, Warsaw University of
Technology*

The aim of the talk is to explore the relation between argumentative, ethotic (i.e. related to the character of a speaker) and dialogical structures in natural communication for the purpose of studying explainable AI. The main claim is that in order to fully capture and adequately model ethotic arguments (such as appeals to authority or expert opinion), the study of their dialogical setting is required. The need for a systematic study of the link between arguments, ethos and dialogue will be justified by discussing some examples of appealing to expertise and authority in argumentation. For that purpose, the argument schemes approach, the profiles of dialogue method, and the OVA (Online Visualisation of Arguments) diagrammatic representation of dialogical and inferential aspects of argumentation will be employed.

QUALITY AND ADVANCES OF XAI



12:45–14:30 LUNCH



14:30–16:00 **EXPLAINABLE REASONING AS A
COMBINATION OF ARTIFICIAL AND
COLLECTIVE INTELLIGENCE**

*Marcin Wozniak / Chief Executive Officer, Optimum Pareto
Foundation*

Explainable artificial intelligence contains not only the challenge of interpreting an algorithm's inferences but also of the practical use of these interpretations by society. The second aspect will be crucial in determining the nature of the emerging technological civilization. Will access to such systems and the ability to interact with them be inclusive and democratic, or will it remain in the hands of the few? Rather than treating this problem as an additional module to be built over and above the basic functionality of inference processes. The Swarmcheck system relies on the power of public discourse to produce and test inferences through argumentation. The presentation will consist of two parts: the theoretical framework of the Swarmcheck system and a demonstration of how it works through collaborative argumentation mapping.

16:00–16:30 COFFEE BREAK



16:30–18:00 **TEAM WORK**

TUTORS / Ewelina Gajewska, Katarzyna Budzyska

18:00–20:30 DINNER

20:30–22:00 SOCIAL ACTIVITY

QUALITY AND ADVANCES OF XAI



DAY 2

Thursday, September 22, 9:15–22:00



9:15–10:45 **FROM EXPLAINABLE AI TO RESPONSIBLE AI: THE INDUSTRY PERSPECTIVE**

*Roberto Diñeiro/ AI Director, Accenture and NL4XAI
Secondment Supervisor*

An AI system is biased when it produces decisions that unexpectedly impact certain groups more than others in a disproportionate way, such as when an AI model outcome fails to properly represent personhood or identity of an individual or when it unfairly impacts a person's ability to access resources. A global landscape on AI standards and regulations plus consumer pressure require that executives consider Responsible AI from design. Implementing Responsible AI brings to the different industries trust that AI is being executed correctly, helps realize the operational advantage of creating scalable and efficient systems and enables mandatory regulatory compliance.

10:45–11:15 COFFEE BREAK



11:15–12:45 **LEGAL ASPECTS OF EXPLAINABLE AI: INDIVIDUAL RIGHTS IN A WORLD OF ALGORITHMIC SYSTEMS**

*Marcin Rojszczak/ Assistant Professor, Faculty of
Administration and Social Sciences, Warsaw University of
Technology*

Algorithmic systems influence the life of an increasing part of society. In the case of automatic decisions that can lead to significant legal effects for an individual, a question arises about the transparency of such actions. Hence, the issue of the explanation of automatic decisions has also been the subject of intensive research in legal sciences for several years now. The aim of the lecture will be to present the key requirements for XAI from a legal perspective. This will also be an opportunity to discuss whether the currently presented models of explanation of automatic decisions can in fact be considered sufficient, taking into account the purpose for which they are introduced – that is the protection of individuals' rights.

QUALITY AND ADVANCES OF XAI



12:45-14:30 LUNCH



14:30-16:00 **DATA, CAPITALISM AND AI: WHY AN ELSEC APPROACH MATTERS**

Ulises Cortes/ High Performance Artificial Intelligence Group Manager, Barcelona Supercomputing Center and NL4XAI External Ethical Advisor

Artificial Intelligence has experienced an exponential boost in its powers and capacities in the last few decades. AI-based tools are no longer a part of the dystopian future those science fiction movies depict. Today, data extraction, social and psychological profiling, and behaviour control have provided astonishing business expansion for companies giving an additional impulse to commercialization and market capitalism. Experts fear that AI-based technologies not only have the potential to reinforce the capitalistic roots of the economy but also recast the world of politics. Will examine why an ELSEC approach is essential to deliver trustworthy and fair Ai-based applications that fit with the EU values.

16:00-16:30 COFFEE BREAK



16:30-18:00 **MENTORING SESSION**

MENTORS/ Marcin Będkowski, Katarzyna Budzynska, Roberto Diñeiro, Marcin Koszowy, Marcin Rojszczak, Marcin Wozniak

18:00-18:30 SWEET BREAK

18:30-21:00 SOCIAL ACTIVITY

21:00-22:00 DINNER