

FOURTH BRIO MEETING & IDEA LEAGUE ETHICS

:: GIACOMO ZANOTTI

Abstract

Report on the Fourth BRIO Meeting Idea League Ethics Working Group Workshop held at the Lecco Campus of Politecnico di Milano on 4-5 July 2024.

Keywords

BRIO; Trustworthy AI; Ethical aspects of AI.

How to Cite

Zanotti, G. Fourth BRIO Meeting & Idea League Ethics Working Group Workshop. *The Reasoner*, 19(1). <https://doi.org/10.54103/1757-0522/27117>

The *Fourth BRIO Meeting & Idea League Ethics Working Group Workshop* was conceived as a collaboration between the project [BRIO – Bias, Risk and Opacity in AI](#) (PRIN MUR), and the *Ethics Working Group of IDEA League*.

The workshop took place at the Lecco Campus of Politecnico di Milano on the 4th and 5th of July 2024, and its main aim was to create a fertile ground for interdisciplinary discussions around the ethical, technical, and societal aspects of Trustworthy AI.

The first speaker was Saskia Nagel (RWTH Aachen University), who gave a talk on “Trusting relationships – the case of human-technology interaction in medicine”. Looking at the specific context of human-technology interaction in medicine, Nagel analysed

THE REASONER 19(1), January 2025.

50

<https://doi.org/10.54103/1757-0522/27117>

This work is licensed under a [Creative Commons](#) “Attribution-ShareAlike 4.0 International” license.



BY SA

the conceptual tenability and the moral costs of extending the notion of trust to individual technological artifacts. After that, Giacomo Zanotti gave a talk on risk and uncertainty in AI, presenting the developments of the research conducted at Politecnico di Milano in collaboration with Daniele Chiffi and Viola Schiaffonati, and Emanuele Bottazzi Grifoni investigated the issue of understanding in LLMs by drawing on later Wittgenstein’s philosophy of language. Karl de Fine Licht (Chalmers University of Technology) was the last speaker of the day, presenting a work on “Generative AI in higher education”. Starting from an analysis of the actual capabilities of current Generative AI tools, with a focus on their prospects of deployment in higher education, de Fine Licht considered some potential positive and negative aspects of such deployment, discussing some strategies to collectively evaluate its legitimacy.

The first talk of the second day was given by Ben Wagner (TU Delft), who presented his work “Interlocking governance mechanisms: towards robust accountability of AI systems”. Wagner argued that effective AI governance needs interlocking mechanisms – e.g., transparency databases, audit mechanisms, and financial and corporate reporting – to ensure accountability. After that, Ekaterina Kubyshkina presented her work (in collaboration with Mattia Petrolo and Giuseppe Primiero) on hyperintensionality in Trustworthy AI, distinguishing on a formal basis two sources of hyperintensionality in trust. The last talk of the morning presented the developments of the collaboration between Alkemy (Davide Posillipo) and BRIO’s unit at the University of Milano (Greta Coraglia, Giuseppe Primiero, Francesco Genco), discussing the theory behind the BRIOxAlkemy tool for the evaluation of fairness-related risk.

The last two talks of the workshop were given by Salvatore

Giugliano and Daniele Porello. Giugliano presented a work (co-authored with Andrea Apicella, Francesco Isgrò and Roberto Prevede) on the improvement of the performance of ML through XAI techniques, while Porello focused on the issue of classification biases and presented an ontological framework for protected predicates.

The event was successful and fostered fruitful discussions of different aspects of Trustworthy AI at the intersection of computer science, logic, and philosophy.

GIACOMO ZANOTTI 

Università degli Studi di Pavia