

OF PEOPLE AND CODE
Epistemological, Ontological, and Formal Approaches to AI

28 November 2024
Amphitheatre FCIências.ID, Ciências ULisboa

ABSTRACTS

Christian J. Feldbacher-Escamilla (University of Cologne)
AI for a Social World – A Social World for AI

Artificial intelligence (AI) is increasingly utilised not only in commercial settings but also in policy-making and various social sectors like the legal domain, the health sector, environmental research, and the like. This growing integration has raised concerns about transparency, leading to demands for explanation. A common response is the development of explainable AI, which aims to create systems that are transparent and interpretable. The spectrum ranges from black-box models that lack these qualities to newer AI systems that can readily provide explanations. The case of explainable AI illustrates the importance of adapting such systems to our specific needs. In this talk, we want to argue, however, that it is also important to look in the opposite direction, namely to ask how our knowledge of the social world, its structural features, and possibly also ways to adapt can help us to approach problems of AI. As we will argue, in particular, taking in a social perspective on problems of epistemic justification helps us to address long-lasting epistemic problems with the help of machine learning.

Roberta Ferrario (Laboratory for Applied Ontology – CNR)
Formal Ontologies: Enhancing computer-mediated communication between humans

In the last decades, human-to-human communication has dramatically changed, one of the main factors of change being the role of mediation played by information systems of different sorts and connected technologies. In the first part of the talk, I shall present the interdisciplinary research area of formal ontology, which conjugates philosophically-based conceptual analysis and axiomatic representation and I will particularly focus on foundational ontologies and their importance for systems' interoperability. I shall then try to show how formal ontologies may be used for apparently very distant applications, like digital twin and digital humanities. For the latter point I will leverage the ongoing work in a couple of projects we are currently carrying out in my research lab.

Ekaterina Kubyshkina (University of Milan)
Hyperintensional trust in AI

The notions of trust and trustworthiness in the field of AI are currently the focus of a collective, interdisciplinary effort for clarification. In this work, we contribute to this ongoing debate in two distinct yet interconnected ways. First, we identify two senses in which an agent might place trust in an AI system. The first considers the results of tests conducted on the system alongside the agent's expectations. The second extends this by factoring in the agent's "pragmatic" background when considering these tests. We argue that these two forms of trust can be understood in relation to well-known approaches in statistical inference: the first aligns with a frequentist interpretation, while the second reflects a Bayesian view of

trust. Secondly, we argue that both forms of trust are hyperintensional – meaning that, even when two objects (e.g., data, programs, systems) are necessarily equivalent, an agent may still distinguish between them based on the trust they elicit. Building on this analysis, we present a formal framework that enables reasoning about hyperintensional trust in the context of AI systems.

(Joint work with Mattia Petrolo and Giuseppe Primiero)

Catia Pesquita (University of Lisbon)

Neuro-symbolic AI and Scientific Discovery

In an age where an AI-based system can predict the 3d structure of a protein – the so-called protein fold, from a simple representation of the sequence of aminoacids that compose that protein, and where AI models detect breast cancer on scans earlier than clinicians, what should we be looking ahead to? Will AI become a scientific tool that supports scientists in making novel, perhaps groundbreaking discoveries? Or one that accomplishes that on its own? In this talk, I will discuss how scientific discovery, in particular in the life sciences, offers unique challenges which can easily become opportunities to drive the development of novel AI approaches and why I believe neuro-symbolic AI is needed to transform AI into a bona fide scientific tool.

Giuseppe Primiero (University of Milan)

Verification of trustworthiness for non-deterministic computations and their copies

In this talk I will overview our work in the area of verification of trustworthiness for non-deterministic computational systems. Principally motivated by the insurgence of machine learning systems, we have developed proof systems and relational semantics for the evaluation of their trustworthiness with respect to properties of interest. I will stress motivation and methodology and illustrate how these are translated in the definition of a trust operator for our logics. The most recent development generalizes these logics to a family of trust operators for copies of systems, i.e. when a certain behaviour is observable in a system and one wants to infer trustworthiness for other systems, e.g. trained on the same datasets or using the same learning model.

Camilla Quaresmini (Politecnico di Milano)

Mitigating bias in AI: The case of gender and epistemic diversity

In this talk, I explore the role of diversity in AI through two case studies that highlight how intelligent systems can reinforce biases through misrepresentation and credibility attribution. The first case focuses on gender diversity, illustrating how classification algorithms misgender individuals by enforcing outdated and static identity categories. The second case addresses epistemic diversity, where individuals or groups are unjustly denied credibility and excluded from contributing to or benefiting from knowledge based on their social characteristics. Through these examples, I will show how AI systems can perpetuate discrimination and propose conceptual tools, along with practical implementations, to mitigate these issues.