

Seminar Hora Informaticae

Institute of Computer Science, Prague

Tuesday, December 3, 2024, 13.00 - 14.30 (1 - 2:30 PM) CET

Meeting Room 318, Address: Pod Vodárenskou věží 2, Prague 8

Meeting ID: 914 0834 4018, Passcode: 668534

https://cesnet.zoom.us/j/91408344018?pwd=x2QIz4F42BxlMSmWc1HOwHHA7Uw7PN.1

Robin Kopecký, Institute of Philosophy, Czech Academy of Sciences:

Robots as Moral Agents: Human Intuition, Programmed Ethics, and Real-World Implications.

This talk explores the concept of artificial intelligence and robots as moral agents, focusing on how they are evaluated and expected to behave in moral dilemmas. While robots are often perceived as more likely than humans to make utilitarian decisions, the moral evaluation of their actions depends heavily on context. I will examine foundational human intuitions about robotic minds, their moral status, and the implications of programmed morality interacting with human agents. Drawing on research from the past decade in the philosophy and psychology of technology, I will also present findings from our own studies, including children's perceptions of robotic minds. The discussion will cover specific applications such as autonomous cars, military decision-makers, robotic rescuers, healthcare robots, and AI in end-of-life decisions.

Robin Kopecký (https://philpeople.org/profiles/robin-kopecky) holds a PhD in Evolutionary Biology from the Faculty of Science at Charles University in Prague. He also holds an MA in Philosophy from the Faculty of Arts at Charles University. His research focuses on applied ethics and experimental philosophy, including topics such as the ethics of autonomous vehicles and moral intuitions about AI and robots in both adults and children.

HORA INFORMATICAE (meaning: TIME FOR INFORMATICS) is a broad-spectrum scientific seminar devoted to all core areas of computer science and its interdisciplinary interfaces with other sciences and applied domains. Original contributions addressing classical and emerging topics are welcome. Founded by Jiří Wiedermann, the seminar is running since 1994 at the Institute of Computer Science of the Czech Academy of Sciences in Prague.

https://www.cs.cas.cz/horainf