Robotics Seminar Series

AI For Deception, Manipulation & Coercion: Towards a Theory and Regulation

Peter Asaro
Associate Professor
The New School

This talk will provide an overview of my recent work on deception, manipulation and coercion by AI and artificial agents. I start from a review of the literature in analytic philosophy on coercion, and its related concepts of manipulation and deception. I also consider mass media theories of propaganda, advertising and persuasion, and well as HRI and HCI theories of “nudging” and their relation to the ethics of artificial agents influencing the behavior of human agents. I then consider how these concepts are challenged by emerging AI and robotics technologies, including targeted marketing, Large Language Models, companion chatbots and social robots. From this I aim to develop a theoretical framework for how these concepts might apply to AI and artificial agents as they become increasingly sophisticated and capable. The goal of this framework is to inform policy and the regulation of these systems so as to assess and avoid the most the most harmful risks posed by these systems as they become further integrated into society.

I will also leave time to discuss my work in advocacy around robotics and AI, including the Campaign to Stop Killer Robots, and working with scientists and engineers to advance socially just and ethical technology.

Professor Peter Asaro is a philosopher of science, technology and media. His work examines artificial intelligence and robotics as a form of digital media, the ethical dimensions of algorithms and data, and the ways in which technology mediates social relations and shapes our experience of the world.

His current research focuses on the social, cultural, political, legal and ethical dimensions of automation and autonomous technologies, from a perspective that combines media theory with science and technology studies. He has written widely-cited papers on autonomous weapons from the perspective of just war theory and human rights, and the legal and moral issues raised by law enforcement robots and predictive policing. Prof. Asaro's research also examines agency and autonomy, liability and punishment, and privacy and surveillance as it applies to consumer robots, industrial automation, smart buildings, UAVs and drones, and autonomous vehicles. His research has been published in international peer reviewed journals and edited volumes, translated into French, German, Korean and Braille, and he is currently researching a book that interrogates the intersections between advanced robotics, and social and ethical issues.

In 2009, Prof. Asaro co-founded the International Committee for Robot Arms Control (ICRAC) which has been advocating for an international ban on autonomous weapon systems, and which in 2012 joined a coalition of NGOs to form the Campaign to Stop Killer Robots. The Campaign has been successful in initiating discussions of autonomous weapons at the United Nations Convention on Conventional Weapons (CCW), and seeks to advance those talks to treaty negotiations.

FRIDAY
Jan 12
10-11AM

Rogers Hall 230
FREE
Refreshments Served

OSU Robotics
robotics.oregonstate.edu

Oregon State University