

You are cordially invited to attend the second McLaughlin College Union Debate that will be held on Wednesday, November 25th, 12:30 pm to 2:30 pm, that will consider a most consequential subject,

AI as an Existential Threat?

Artificial Intelligence (AI) research is accelerating, and its rapid development, innovations, and discoveries are already having an impact on society in quite dramatic ways such as autonomous vehicles, AI-generated music, poetry and storytelling, customer service bots and portals, and so on. The term “transformative AI” is used to describe a range of advances in AI that could impact on society in dramatic and difficult-to-reverse ways. Government policies and regulations will, undoubtedly, find it extremely difficult to keep up with the pace of technological progress with AI. Researchers are already working on advanced warning systems for any possible extreme events. However, *AI forecasting* based on *measuring AI progress* is at its early stages of development and its utility has been challenged by those who point out that it could never be able to account for the revolutionary breakthroughs and discoveries that have the potential to achieve AGI (Artificial General Intelligence), that will allow machines to adapt to a variety of situations to maximize their potential, or to achieve high-level machine intelligence (HLMI), to perform at the level of an average human adult on key cognitive measures necessary for economically relevant tasks, or to achieve “*superintelligence*,” that Nick Bostrom, states “greatly exceeds the cognitive performance of humans in virtually all domains of interest.” (*Superintelligence: Paths, Dangers, Strategies*. Oxford University Press, 2014) Leading, some would argue, to the “*AI control problem*” that may be unresolvable.

We have assembled an exceptional panel for our second McLaughlin College Union Debate that will consider the following proposition:

The rapidly accelerating advances in artificial intelligence (AI) could pose, quite likely, a serious existential threat to humankind in the not too distant future.

Please join us for what will undoubtedly prove to be a fascinating debate.

MODERATOR:



JAMES C. SIMEON, is Head of McLaughlin College and an Associate Professor in the School of Public Policy and Administration (SPPA), Faculty of Liberal Arts and Professional Studies, York University. He is a Member-at-Large of the Executive of the Canadian Association for Refugee and Forced Migration Studies (CARFMS) and a past President of CARFMS.

PANELISTS:



RUTH URNER, Assistant Professor, Department of Electrical Engineering and Computer Science, York University. Her areas of research interest are Machine Learning Theory, Societal Aspects of Machine Learning, Fairness in Machine Learning, and Interpretability in Machine Learning.

Ruth will speak against the proposition.



REGINA RINI, holds the Canada Research Chair in Philosophy of Moral and Social Cognition and is Associate Professor in the Department of Philosophy at York University. She studies how social norms should be negotiated in democratic societies, particularly in response to technological change. Her work has appeared in the *New York Times* and *Los Angeles Times* and she writes the regular ‘Morals of the Story’ column for the *Times Literary Supplement*. **Gina will speak in favour of the proposition.**



MATTHIJS MAAS, Postdoctoral Research Associate, Centre for the Study of Existential Risk, University of Cambridge, with research interests in adaptive global governance approaches for extreme technological risks, with a focus on high-stakes or destabilizing uses of AI technology in military contexts and the effects of AI technology on international law, among other areas.

Matthijs has agreed to argue against the proposition.



ALLAN WEISS, Associate Professor, Department of English and Department of Humanities, York University, is a creative writer and scholar who has published mainstream/literary, science fiction, and fantasy short stories in numerous periodicals and anthologies.

Allan will argue in favour of the proposition.



STUART ARMSTRONG, Research Fellow, Future of Humanity Institute, University of Oxford. Stuart has worked with researchers at DeepMind, to formalize AI desiderata in general models so that AI designers can include these safety methods in their designs and his collaboration with DeepMind on “[Interruptibility](#)” has been mentioned in over 100 media articles.

Stuart has agreed to argue both sides of the proposition.

Wednesday, November 25, 12:30 – 2:30 pm

CLICK here to register: <https://mycentre.apps01.yorku.ca/machform/view.php?id=613581>

For more information email Vicky Carnevale at vcarneva@yorku.ca