



AI Frontiers Conference: 8 June 2019

The **Rhodes Artificial Intelligence Lab** (RAIL) is a Rhodes Scholar-led group of graduate students at Oxford working to apply machine learning for social good. We work with partner organisations around the world to create artificial intelligence solutions to complex social challenges, from improving microscopy techniques for diagnosing diseases to addressing toxic content in online conversations. Our members are Rhodes Scholars and graduate students working across mathematical, computational and social sciences.

Frontiers for AI in Society

Artificial intelligence and machine learning techniques are becoming more widespread at a rapid rate, and their applications are increasingly diverse. For governments, artificial intelligence represents a significant challenge – particularly to employment and labour participation rates, as automation threatens to displace some types of manual work. At the same time, AI offers unprecedented opportunities to improve governance and promote better policy outcomes. Used creatively and responsibly, disruptive technologies can increase freedom, enhance people's lives and help governments to perform effectively. Manipulated or misunderstood, however, they can present a real threat to civil liberties and distort economic and social relations.

RAIL's first annual conference will highlight the latest frontiers in artificial intelligence, addressing the intersections between artificial intelligence and society and bringing together some of the world's leading experts on how to adapt to, regulate, and harness AI.

This conference aims to highlight new frontiers in the impact of artificial intelligence on society, focusing on three cutting-edge issues:

1. The potential impact of AI on industrialisation and development in low-income countries, and how governments can adapt to automation
2. The increasing use of AI to promote healthy conversations online, and how governments can combat abuse, toxicity, political interference, and disinformation
3. The impact of AI on creativity and the future of artistic production



RHODES ARTIFICIAL INTELLIGENCE LAB

Date: Saturday 8 June 2019

Expected number of attendees: 80-100

Final programme:

9:30am Guests arrive (coffee in the Atlantic room)

9:45am Welcome – RAIL Directors and CEO of the Rhodes Trust

10:00am **Panel 1: AI x Development**

With most effort focused on the impact of automation and artificial intelligence on the future of work in advanced economies, comparatively less attention has been paid to the potential impact of disruptive technologies on industrialisation in low-income countries. Are concerns about the possible loss of jobs and obstacles to industrialisation legitimate, or overblown?

Speakers:

Arnab Thakur leads Frontier Technologies for NITI Aayog, the Government of India's premier strategy and policy institution, and is a co-author of India's National Strategy on AI.

Prof Benno Ndulu is the Academic Director of the Pathways for Prosperity Commission, and a former Governor of the Central Bank of Tanzania.

Prof Stephen Gelb is Senior Research Fellow at the Overseas Development Institute, where he leads work on private sector development, and an economist with thirty years of experience in research and policy in South Africa and elsewhere.

11:00am Break

11:15am **Panel 2: AI x Public Discourse**

The development of new machine learning techniques to develop and disseminate "fake news" and political disinformation is an object of rising global concern. At the same time, artificial intelligence models are being used to identify harmful content and promote healthier engagement online. How can governments



RHODES ARTIFICIAL INTELLIGENCE LAB

harness and regulate these technologies for good, without compromising political freedoms?

Speakers:

Dr Lucas Graves is Acting Director of Research at the Reuters Institute for Journalism, and has been at the forefront of research on the fact-checking movement in the United States and around the world – including in his 2016 book *Deciding What's True: The Rise of Political Fact-checking in American Journalism*.

Dr Stefan Theil is a Postdoctoral Research Fellow at the Bonavero Institute of Human Rights, where his research focuses on whether the differences in online versus offline behaviour require the law to revisit the principles and assumptions underlying human rights.

Rishabh Shukla is a Research Engineer at Factmata, where he works on NLP and deep learning models to tackle fake news and misinformation and to detect hateful, toxic and biased content on the internet.

12:15pm Lunch

13:00pm **Panel 3: AI x Creativity (in partnership with the TORCH Futures Thinking Network and the OII Research Project on Creative Algorithmic Intelligence)**

Algorithms can now generate original music, text and images - but are they really "creative"? What would this even mean? And what are the implications for the future of creative industries?

Speakers:

Anne Ploin is a Research Assistant at the Creative Algorithmic Intelligence project in the Oxford Internet Institute, where her work explores creative human/AI collaborations and the automation potential of creative work.

Dr Isis Hjorth is a cultural sociologist and Research Associate at the Oxford Internet Institute, where she specialises in analysing emerging practices associated with networked technologies. She



RHODES ARTIFICIAL INTELLIGENCE LAB

has previously researched participatory digital tools for artistic production.

Marie von Heyl studied Fine Art at Kunsthochschule Berlin Weißensee and the Royal Academy in London. Her work draws from the poetic friction and productive misunderstandings that emerge when different models of reality collide, overlap or do not quite fit together.

Dr Paul Duckworth is a Postdoctoral Research Assistant in Machine Learning in the Oxford Department of Engineering Science and a member of the Machine Learning Research Group. His interests include the impact of machine learning technologies on society and the future of human work, and our understanding of algorithmic (creative?) intelligence.

14:30pm Guests depart