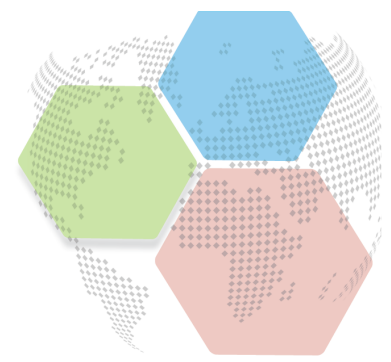


FRONTIERS OF SCIENCE: ARTIFICIAL INTELLIGENCE

THE ROYAL SOCIETY OF CANADA
AND THE ROYAL SOCIETY
OTTAWA, ONTARIO, CANADA



SCHEDULE

MONDAY, FEBRUARY 12	
6:00 pm	Evening Reception Fairmont Chateau Laurier, Zoé's
TUESDAY, FEBRUARY 13	
8:00 am	Breakfast Begins Fairmont Chateau Laurier, Québec Room
9:00 am	Welcoming Remarks
	Territorial Acknowledgment with Grandma Karen MacInnis , Knowledge Keeper and Elder
9:30 am	KEYNOTE PRESENTATION
	Yoshua Bengio , Department of Computer Science and Operations Research, Université de Montreal
	Why and How Could we Design Aligned and Provably Safe AI? This presentation will start by presenting the major risks associated with advanced AI systems, as we progress towards human-level competence - or more - on more and more cognitive tasks. This suggests two significant challenges to make sure that we avoid catastrophic outcomes: the technical alignment challenge (finding a way to design AI that behaves as intended, a yet unsolved problem) and the political challenge of coordination and democratic governance, at national and international levels (making sure that everyone follows the safety protocols and does not abuse the power of AI). I will then outline a machine learning research program to obtain quantitative and conservative risk evaluations to address the alignment challenge. This is based on deep learning methods for amortized inference of the Bayesian posterior over causal theories of the data available to the AI, and taking a risk management posture in which the AI takes conservative decisions accounting for the uncertainty around consequences of its actions and what humans really consider harmful.
10:15 am	Group Discussion
<i>Coffee served</i>	
11:00 am	The Great Debate Hosted by Yoshua Bengio 1. Do we have the necessary mechanisms for participatory action research on AI? 2. What is the scope of education needed to support AI literacy? And do we have a sense that there is a critical mass of the scientific community that is rising up to this challenge? 3. How can we promote an academic research environment that can tackle today's AI scale? Can we really justify continued development of large models?
12:00 pm	Lunch
1:00 pm	Breakout Groups The Impact of AI: AI and Ethics, AI and Health, AI and Privacy, AI and Deep Learning Problems include privacy concerns and the effect of energy consumption, among others. On the other hand, AI could lead to advancements in areas such as drug discovery and osmological analysis. We need to consider how both society and academia can develop AI while ensuring overall societal benefits. A sub-topic within this context is that even within academia and science, sometimes techniques are not discussed in a transparent and honest manner (e.g. glorying performance metrics) which makes them less reliable when used in practice, but also makes it difficult to apply them in science. What can we do as AI and data scientists to ensure transparency and accessibility? How do we foster easier comparison to enable access to AI technology within science? Perhaps a starting point is identifying cases where there is more benefit in using simpler models than directly jumping to deep learning.
2:30 pm	Breakout Group Presentations
4:00 pm	Cultural Activity (Skating, Ice Sculptures, Winterlude, Beavertails)
6:30 pm	Dinner Métropolitain Brasserie



WEDNESDAY, FEBRUARY 14		
8:00 am	Breakfast Fairmont Chateau Laurier, Québec Room	
9:00 am	PANEL PRESENTATIONS AI and its impacts, including regulatory aspects ?	
	Moderator Elissa Strome , Executive Director, Pan-Canadian AI Strategy, CIFAR	
	Panelists	Stephanie Dick , Simon Fraser University, Canada
		Atoosa Kasirzadeh , University of Edinburgh and Alan Turing Institute, UK
		Luke Stark , Western University, Canada
Erin Young , The Alan Turing Institute, UK		
10:00 am	Group Discussion	
11:00 am	KEYNOTE PRESENTATION	
	Doina Precup , Machine decision-making, McGill University Learning how to take decisions from preferences Learning how to take sequences of decisions under uncertainty is a hallmark of intelligence. Reinforcement learning, which draws inspiration from neuroscience, animal learning theory, control theory and economics, has been the method of choice for learning how to predict and act from interaction with an unknown environment. In reinforcement learning, the goal of the agent is expressed as the maximization of an expected numerical reward signal. However, when AI agents are trained by interacting with people, eliciting a numerical reward can be difficult and noisy. As a result, the method of choice is to allow a person to provide a preference between two different trajectories (for example, two answers to a query). But is this approach really justified? In this talk, I will discuss recent work in which we show that preferences can indeed be more expressive than numerical rewards, while still allowing for the existence of optimal behaviors. I will discuss recent algorithmic advances for learning such behaviors from data, in reinforcement-learning-style. Finally, I will discuss why alignment should be viewed as a sequential decision-making problem, rather than a one-step process, and why I am optimistic about the future AI-human coexistence.	
12:00 pm	Lunch	
1:00 pm	Breakout Groups Collaboration Between Countries and Disciplines: AI and Ethics; AI and Health; AI and Privacy; AI and Deep Learning	
	Interdisciplinary work is often difficult to publish (and fund) because it requires reviewing expertise from multiple distinct communities. How can we encourage collaborations across disciplines while supporting the careers of researchers involved in these collaborations (e.g., students, early-career faculty)? Different countries have distinct funding and recruiting priorities. How can we decrease the overhead of collaboration across countries to facilitate research projects that leverage each country's strengths? Are the benefits of interdisciplinary collaboration immediately clear to your areas? For example, in biology, researchers may benefit a lot by having AI collaborators, and they might also have possible venues for publication, whereas mainstream AI researchers will typically go for mainstream AI conferences, where an application in other areas <u>has to</u> be written in a way that's readable and accessible to a computer science audience. What other ways have people worked this out?	
2:30 pm	Breakout Group Presentations	
4:00 pm	Policy Discussion David Prodger, Denisse Albornoz and Nicole Mwananshiku	
5:00 pm	Light Dinner	
6:30pm	High Commission Reception Ottawa Art Gallery (OAG) – Canopés and Drinks reception	
THURSDAY, FEBRUARY 15		
8:00 am	Breakfast Fairmont Chateau Laurier, Québec Room	
9:00 am	Morning Wrap-up Co-Chairs and Panelists	
10:00 am	Reflection Discussion	
11:00 am	Closing Ceremony	
12:00 pm	Travel Out	