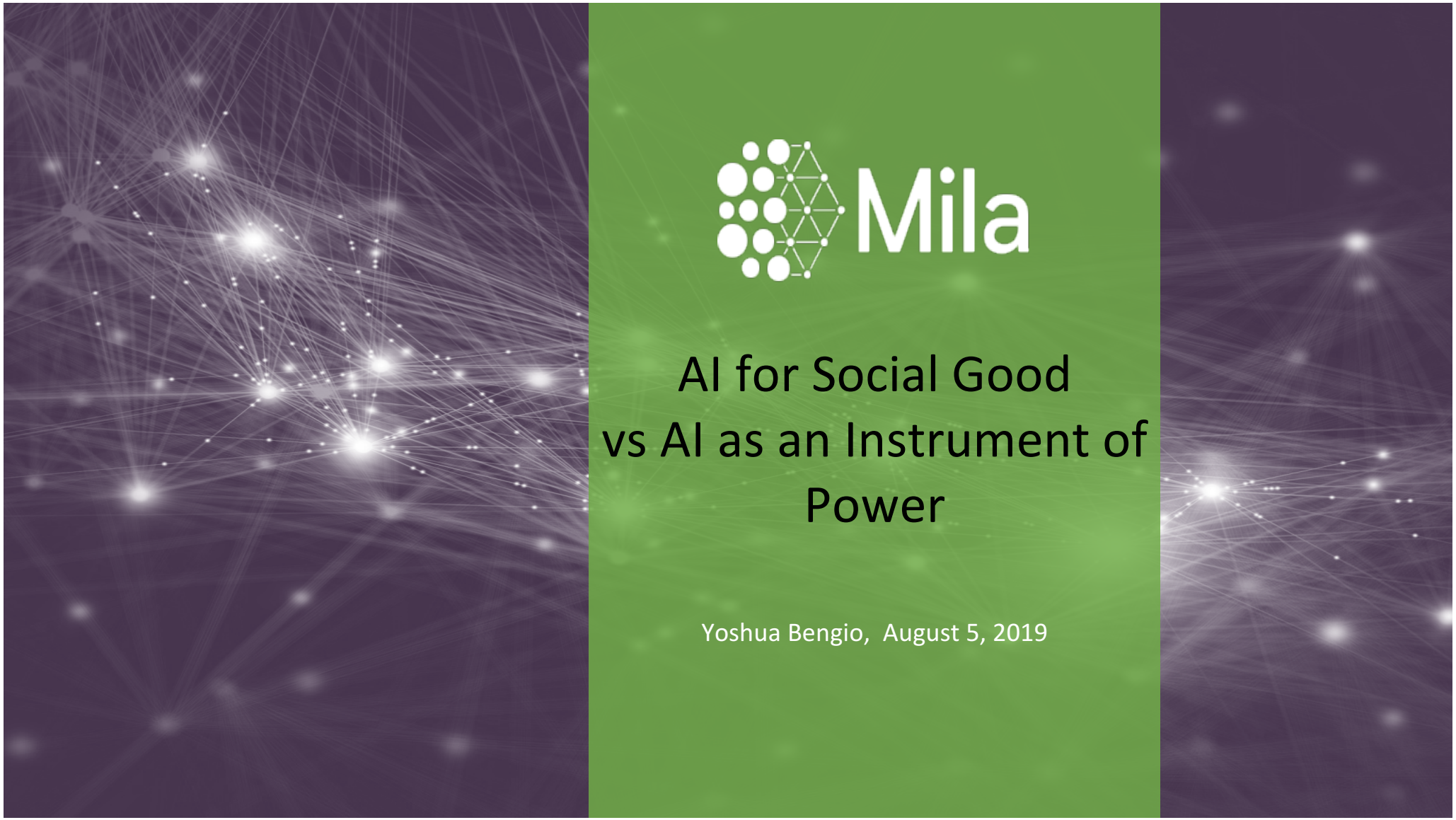



# AI for Social Good vs AI as an Instrument of Power

Yoshua Bengio, August 5, 2019

# Mila - Missions and Values

## Missions

**Academic research**

Contribute distinctively to basic research in deep learning and reinforcement learning

**Education & Talent**

Attract, train and retain a talent pool in the field of machine learning and artificial intelligence

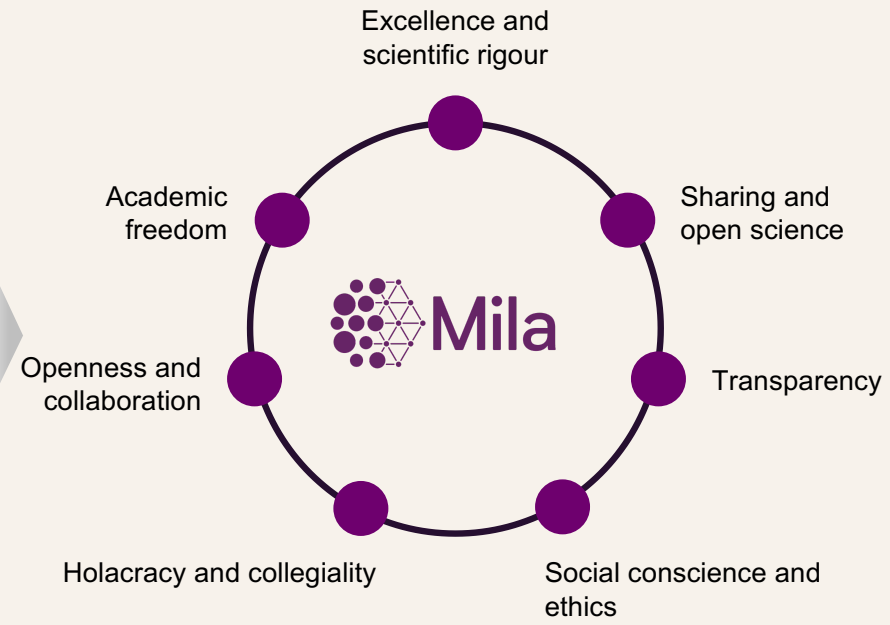
**Technological transfer**

Contribute to Québec's economic development through technology transfer and business innovation

**AI for Humanity**

Foster a dialogue on the socially responsible use of AI and the development of social and environmental applications

## Fundamental Values

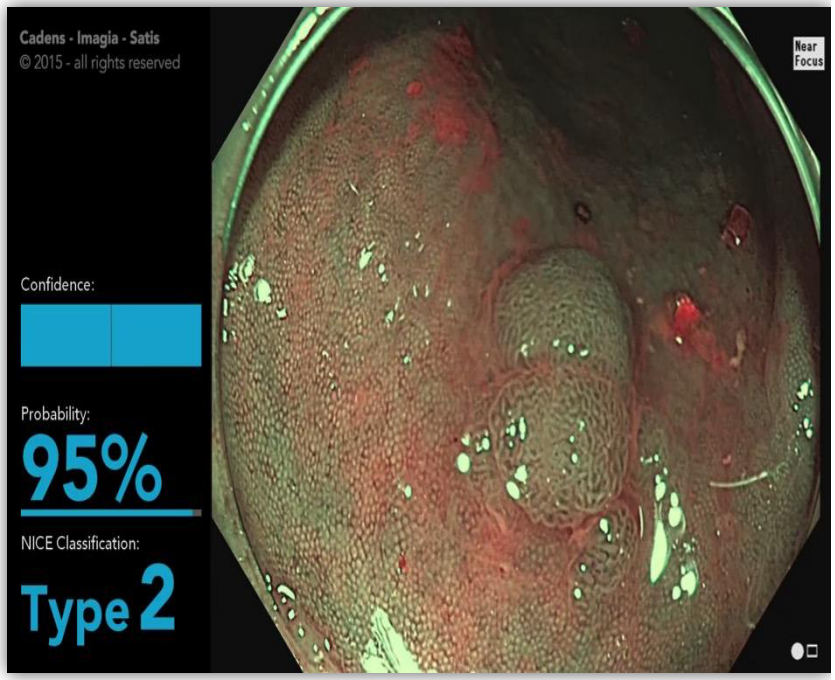



Mila

# Beneficial AI Potential

- Medical applications
- Environmental applications
- Educational applications
- Democratizing justice
- Humanitarian applications

Etc.




*Imagia detects cancer cells, helps doctors*

Mila

# Beneficial AI Activities at Mila


- Montreal Declaration for the Responsible Development of AI
- Recruitment incentives for diversity
- AI4G workshops @ NeurIPS 2018, ICLR 2019, ICML 2019, NeurIPS 2019
- AI & Climate Change workshops @ ICML 2019, NeurIPS 2019, ICLR 2020
- Democratize AI in developing countries, research interns program
- AI projects for healthcare research, climate change,  education, humanitarian applications
- AI Commons project, ***http://www.aicommons.com***
- Int'l School on Bias and Discrimination in AI, June 2019


# Montreal Declaration (2017)

1- **WELL-BEING** PRINCIPLE

2- **RESPECT FOR AUTONOMY** PRINCIPLE

3- **PROTECTION OF PRIVACY AND INTIMACY** PRINCIPLE

4- **SOLIDARITY** PRINCIPLE

5- **DEMOCRATIC PARTICIPATION** PRINCIPLE

6- **EQUITY** PRINCIPLE

7- **DIVERSITY INCLUSION** PRINCIPLE

8- **PRUDENCE** PRINCIPLE

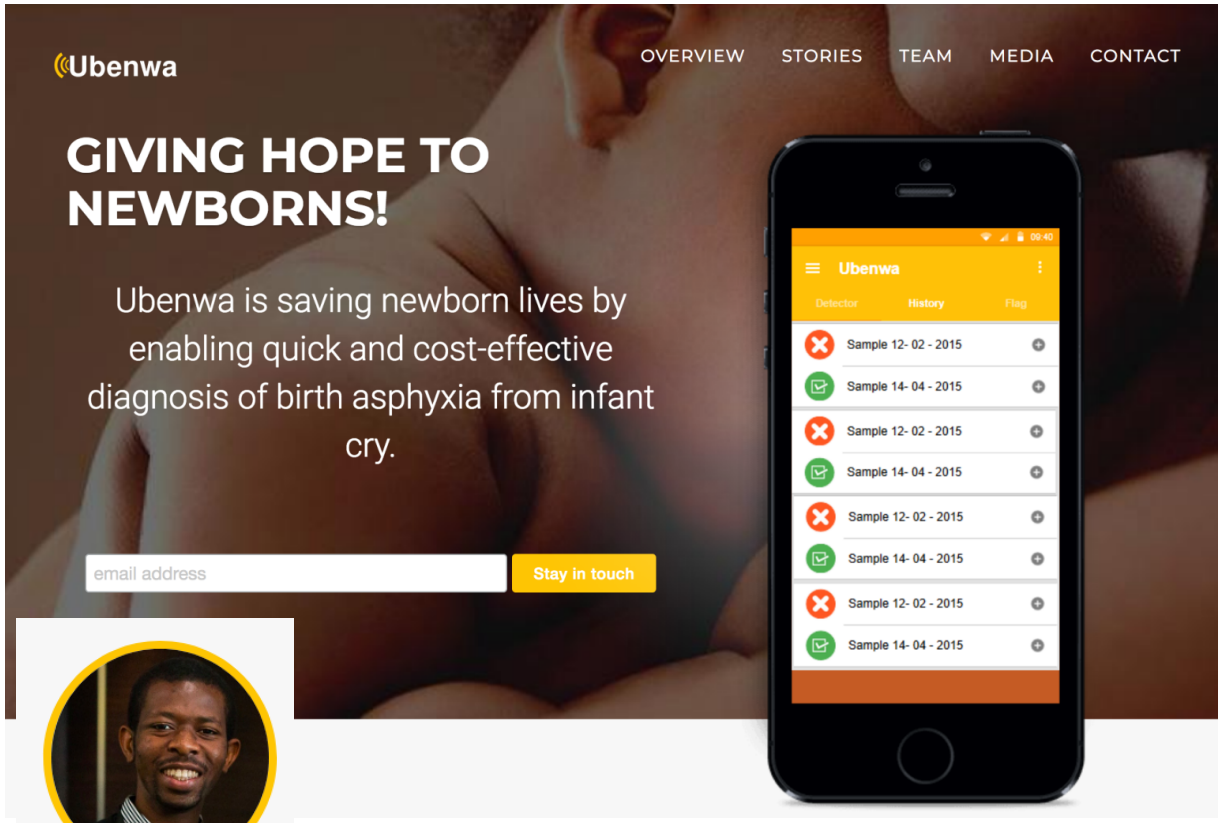
9- **RESPONSABILITY** PRINCIPLE

10- **SUSTAINABLE DEVELOPMENT** PRINCIPLE

https://www.montrealdeclaration-responsibleai.com/the-declaration


- International organization to facilitate AI for Social Good Applications
- Hub for connecting problem owners, machine learners, developers, start-ups & nonprofits, philanthropy
- Clearinghouse helping to prioritize projects, collect and validate data, bring people together

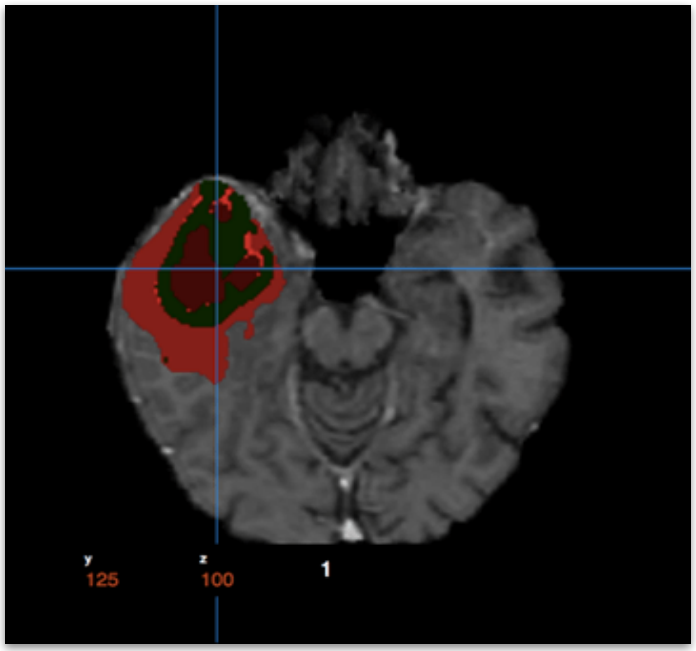
Mila




Mila

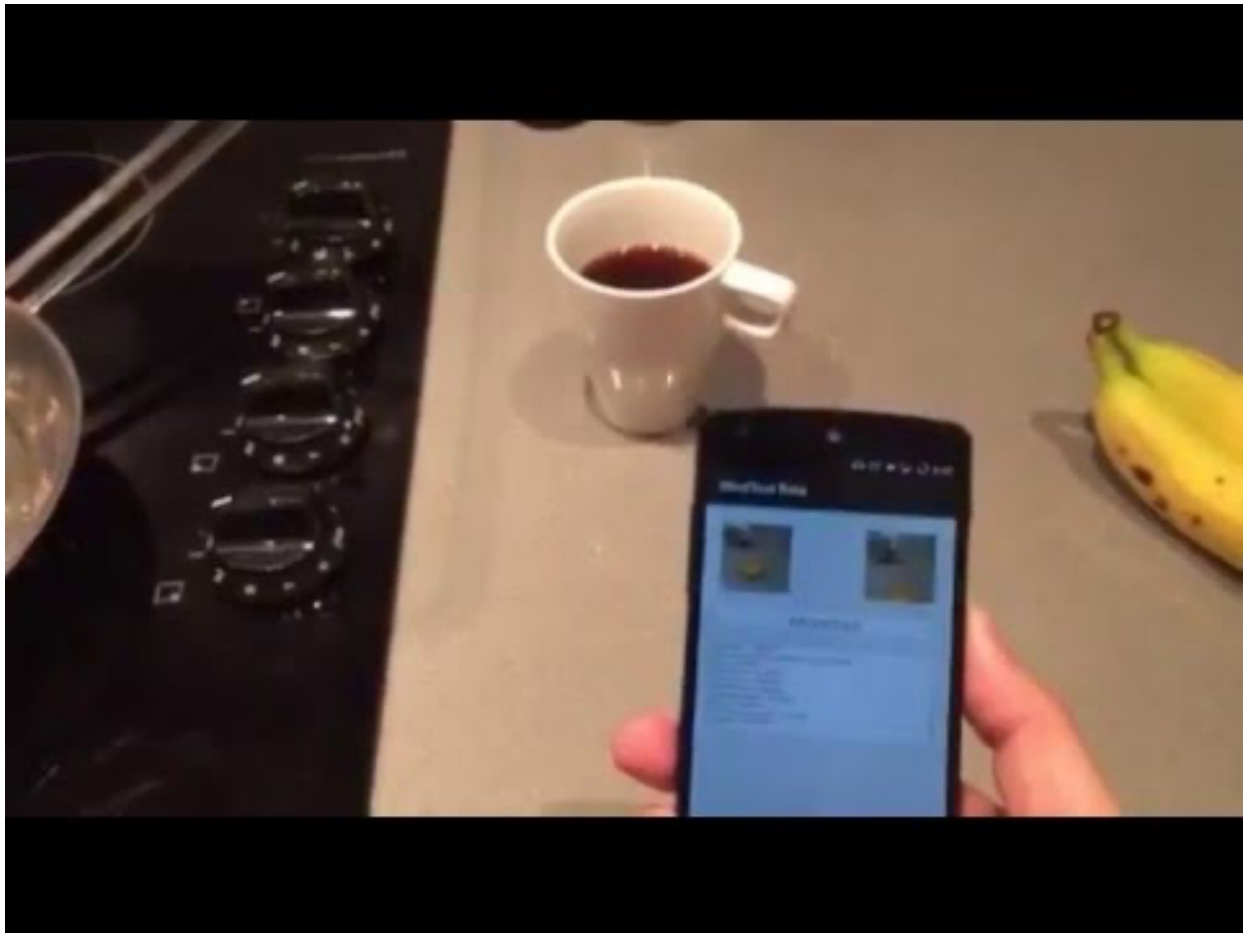
https://www.ubenwa.ai/index.html#page-top

# Assistance for Cancer Detection



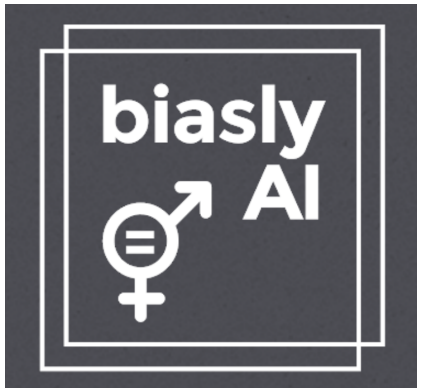

*Images:* http://qure.ai/index.html#whatwedo

Mila



**Blindtool app,** Joseph Cohen http://josephpcohen.com/w/blindtool-helping-the-blind-see/

Mila




Boys are agressive.

Girls like to cook.

Mila https://sites.google.com/view/biaslyai/home

# Environmental Applications of AI

- Optimizing energy resources
  - Smart grids, demand forecasting, minimize transportation costs
- Climate modelling
  - Predicting effect of climate change
- Accelerated R&D of new materials
- Conservation efforts
- Visceralization: visualize future impact of climate change


Mila

# Sufficient to Fight Climate Change: Political Will

- It may still technically be possible to limit warming to 1.5C if drastic action is taken now: Smith et al, Nature Communications, 15 jan 2019

- Politicians will act if there is sufficient priority given to climate change in the general population (dixit some politicians who want to do more but whose hands are tied by current opinions)

- How do we convince those who are not yet and increase the priority they give to this issue?


Mila

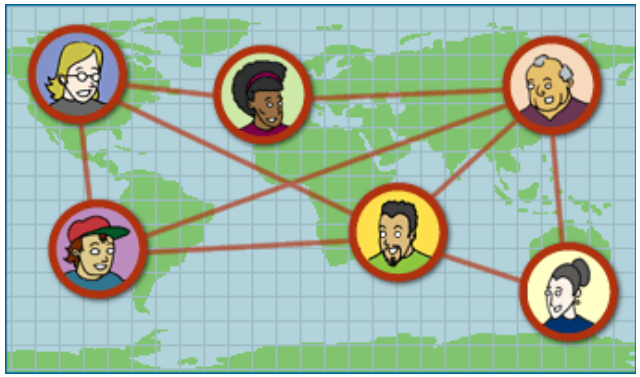
# The Climate Change Visualization Project

- We are working on a project that aims to generate **images** that depict **accurate**, **vivid**, and **personalized** outcomes of climate change using CycleGAN models.

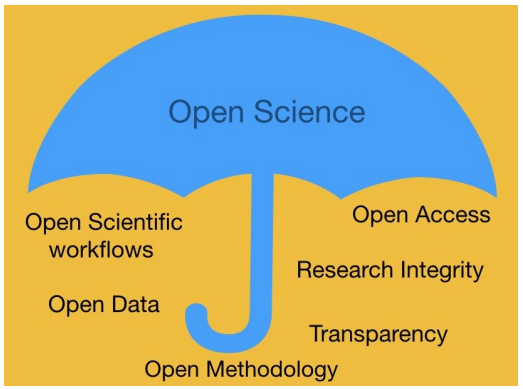
Team members: Karthik Mukkavilli, Sasha Luccioni, Victor Schmidt, Vahe Vardanyan in collaboration with Jennifer Chayes (Microsoft Research)

Mila

# From Democracy to Honest Rational and Inclusive Debates



- Replacing advertising by honest, rational and inclusive debates
- Social model for collective choice and resource allocation: the scientific community


Mila

# *Why Digital Contact Tracing?*

**Limitations** of manual tracing compared to mobile phones digital tracing:

- Covid-19 is most contagious pre-symptomatic (not detected by manual contact tracing)
- Poor coverage of population (& not all contacts are with known persons)
- Delay of days

**Advantage** of manual tracing:

- Human tracers use their judgement, can integrate many clues, including ones not observed by a phone

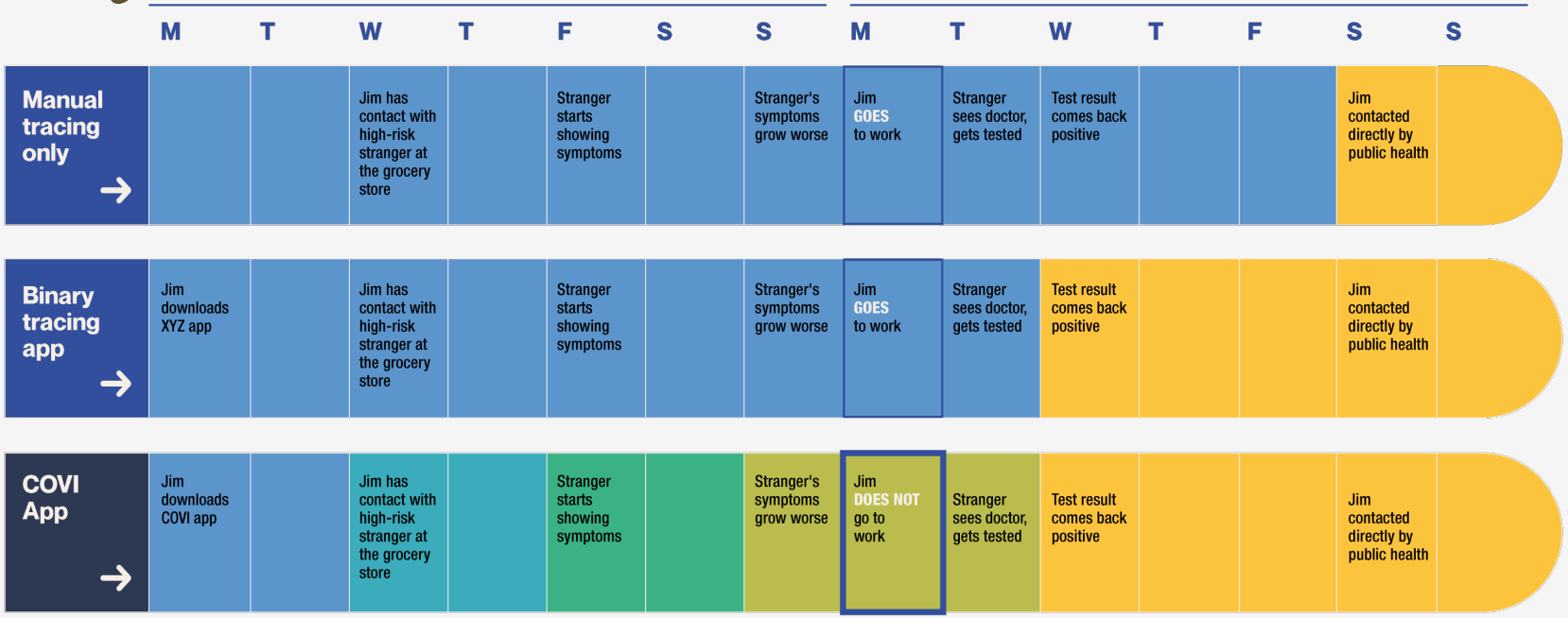
***Could ML-based digital tracing also integrate many clues to identify highest risks?***
***(Best of both worlds)***

# *Why ML-Based Risk Estimation?*

- Standard tracing (manual or digital) is binary and brittle: contact or no contact?
- In reality, many clues should be integrated before taking a decision about the risk level of a person:
  - **Symptoms: available many days before tests**
  - Prior medical conditions
  - Age, biological sex
  - **Risk levels of all the contacts**, and when these happened

- ML can be used to predict the infection status probability distribution, given these clues
- Provides ***early warning signals***, well before standard tracing (digital or manual) would raise a flag
  - Saves lives (lower R) and enables selective deconfinement
  - These signals and other information from the app can ***empower manual contact tracers*** to make informed decisions and extend their reach
  - Empowers public health management (e.g. testing more in some areas, based on early warnings of outbreaks)

# *Early Awareness Saves Lives*




- ML estimators can aggregate many weak clues, then warn a user that their risk has increased well before they would otherwise get contacted by public health manual tracing.
- Studies have shown that people change their behavior drastically and self-isolate as soon as they suspect they might be contagious.
- Early awareness would reduce the number of contagious contacts and the rate of spread of the virus.

# *Why a Privacy-First Design?*


- Wide adoption is necessary for significant impact
- Main concern of citizens is privacy
- It is possible to obtain **strong privacy** while enabling contact tracing and risk awareness propagation: two main concerns:
  - Privacy wrt authorities: governments should not know my trajectory and contact graph **(BIG BROTHER)**
  - Third party anonymity: others should not know my risk and personal data **(LITTLE BROTHER)**

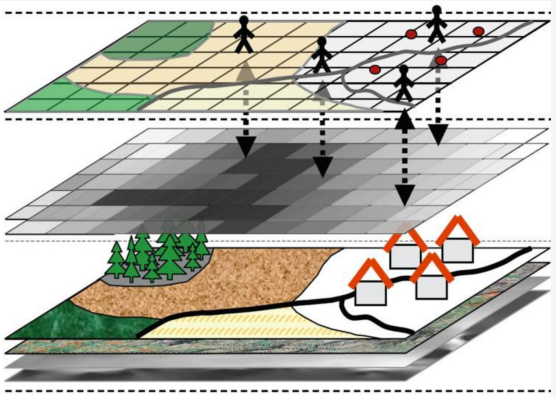
# *Social, Ethical & Governance Considerations*


- Human rights include:
  - Privacy
  - Safety and freedom
  - Dignity, avoid stigmatization
- This would not be compatible with an app which
  - Tells us we cross path with somebody probably infected
  - Tells us that infected people are frequently at some location

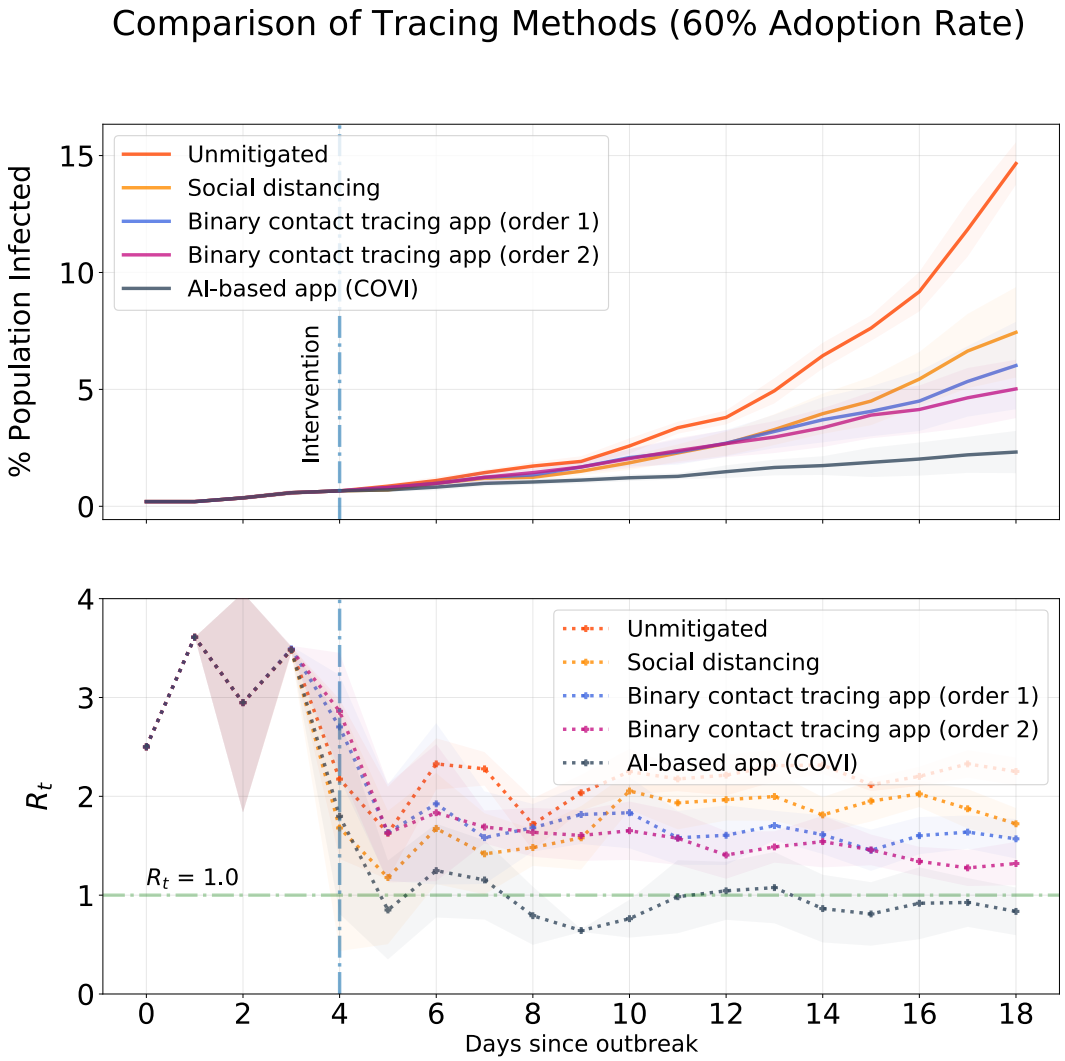
  (even if it would be tempting purely from the health point of view)
- Data should be shared only with well-informed consent
- Users should be able to ask that their data be removed from the database
  - Old data removed after three months or the end of the epidemic
- Managing these issues and the data should be by an independent & non-profit ***data-trust***

# *Epidemiological Generative Model*

- Stochastic agent-based model, implemented in Simpy
- Population of humans moves around according to a mobility distribution and stochastically meets at locations, creating contacts (with duration and distance characteristics)
- The mobility model captures the time spent in different types of locations (home, office, transportation, stores, hospitals, long-term care facilities, etc).
- Humans can be in several discrete states (susceptible, exposed, contagious, recovered) and hold continuous state variables (e.g. viral load, contagiousness).
- Symptoms stochastically arise due to flu/cold or Covid
- Temporal evolution of contagiousness
- Spread of the disease tracked via $R_0$ , attack rate etc
- Parameters set or tuned to match known statistics



# *Preliminary Results*




Mila

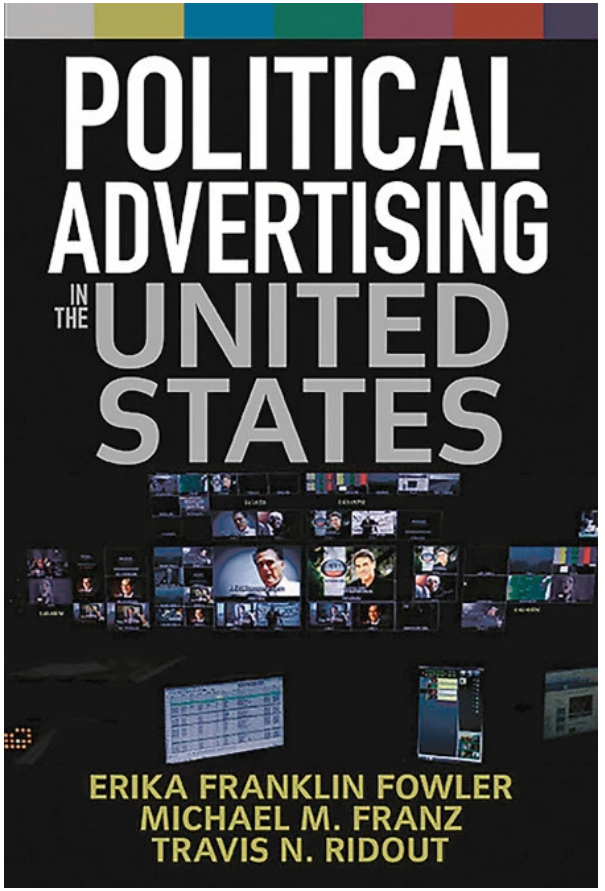
# Dangers and Concerns with AI

- Big Brother and killer robots
- Misery for jobless people, at least in transition
- Manipulation from advertising and social media
- Reinforcement of social biases and discrimination
- Increased inequality and power concentration in few companies


Mila

# Advertising and Social Networks can lead to Psychological Manipulation & Hurt Innovation

- **Moral hazard:**
  - Psychological manipulation: for who's benefit? The advertiser.
  - Targeting children forbidden in Canada because they are vulnerable
  - Political advertising
  - Need to draw a red line
- **Economic drag:**
  - favours large incumbants and thus slows innovation
- **Bad Nash equilibrium**




Mila

# Killer Robots: Moral & Security Threat

- 62% worldwide in favour of a ban to limit LAWS
- UN Secretary General: ’’The prospect of machines with the discretion and power to take human lives is morally repugnant… could trigger new arms races”
- 28 states calling for an international ban
- Leading AI scientists signed letters to governments
- 4000 Google employees named “Arms Control Person of the Year”
- Companies, countries, civil society must rise
- Example of landmines


Mila

# Weak AI but Strong Technology

- Mistakes reveal that current AI is very weak compared to human intelligence
  - Shallow & superficial understanding
  - Poor generalization out of training domain
  - No moral understanding whatsoever, no understanding of human psychology


- But powerful enough to be economically very valuable and politically very dangerous, and power is rising

Mila

# The Wisdom Race

**Collective and individual wisdom** has increased

e.g., the decrease in crimes, wars and extreme poverty

**But not fast enough** to catch up with the rise in power of the tools we are building, which enable power concentration and can be dangerous

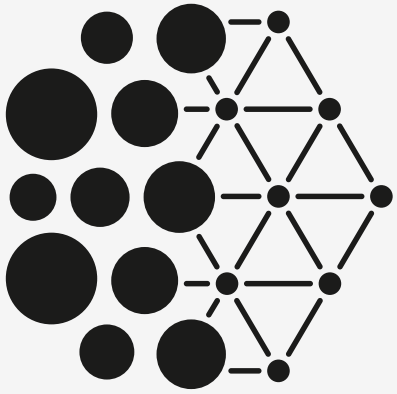
Mila

# Everyone Must Benefit or Else

Think about what happens when any human (e.g. crazy person) can throw off major nuclear weapons on the face of the Earth. This is the threat of unbridled technological development.

**To reap the amazing benefits of technology, we also need to increase the pace of progress of collective wisdom, manage international AI governance, make sure every human is well-fed, well-educated and healthy**

Mila



Mila

Université de Montréal

McGill

Mila

Québec

CIFAR