



# Introduction to AI Ethics

**Daniel Z Buchman PhD RSW**

Bioethicist and Independent Scientist, Centre for Addiction and Mental Health

Assistant Professor, Dalla Lana School of Public Health, University of Toronto

University of Toronto Joint Centre for Bioethics

@DanielZBuchman

# Disclosures

- Associated Medical Services (AMS) Fellowship in Compassion and Artificial Intelligence
- University of Toronto Dalla Lana School of Public Health Interdisciplinary Data Science Cluster Seed Funding



# CAMH Land Acknowledgement

---

CAMH is situated on lands that have been occupied by First Nations for millennia; lands rich in civilizations with knowledge of medicine, architecture, technology, and extensive trade routes throughout the Americas. In 1860, the site of CAMH appeared in the Colonial Records Office of British Crown as the council grounds of the Mississaugas of the New Credit, as they were known at the time.

Today, Toronto is covered by the Toronto Purchase, treaty No. 13 of 1805 with the Mississaugas of the Credit.

Toronto is now home to a vast diversity of First Nations, Inuit, and Métis who enrich this city.

CAMH is committed to reconciliation. We will honour the land through programs and places that reflect and respect its heritage. We will embrace the healing traditions of the Ancestors, and weave them into our caring practices. We will create new relationships and partnerships with First Nations, Inuit, and Métis and share the land and protect it for future generations.



Krembil Centre for  
Neuroinformatics



**camh**

Shkaabe Makwa

# Code of Conduct

The KCNI Summer School is dedicated to providing a harassment-free learning experience for everyone, regardless of gender, gender identity and expression, sexual orientation, disability, physical appearance, body size, race, age or religion. We do not tolerate harassment of event participants in any form

Harassment includes, but is not limited to:

- Verbal comments that reinforce social structures of domination related to gender, gender identity and expression, sexual orientation, disability, physical appearance, body size, race, age or religion.
- Sexual images in public spaces
- Deliberate intimidation, stalking, or following
- Harassing photography or recording
- Sustained disruption of talks or other events
- Inappropriate physical contact
- Unwelcome sexual attention
- Advocating for, or encouraging, any of the above behaviour

# Respectful Language

When discussing mental health, always remember that **the person comes first** and the diagnosis come second.

## *Please don't say*

depressives

schizophrenics

addict

drug abusers

## *Instead say*

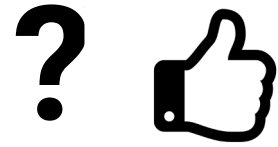
people with major depressive disorder

people with schizophrenia

person with substance use disorder

people who use drugs

# Remember - many ways to engage



(during sessions)  
Use the chat or  
the ask question!



You can always return to the  
session and re-watch the videos  
after the session ends



come chat with us in KCNI  
Summer School Slack :)



virtually meet with us  
in gather.town



Tell us how the session went (post session survey):

<https://forms.gle/ji18qLMZEZ9L16Ln6>



[KCNISchool@camh.ca](mailto:KCNISchool@camh.ca)

# Objectives

- **Define** bioethics
- **Identify** key ethical issues related to AI and machine learning
- **Describe** the importance of ethics to the design of ML systems and its potential research and clinical applications



**ETHICS AND GOVERNANCE  
OF ARTIFICIAL INTELLIGENCE  
FOR HEALTH**

WHO GUIDANCE



World Health  
Organization



A decorative graphic consisting of several overlapping, wavy, teal-colored bands that flow from left to right across the middle of the page. The bands have a slight gradient and a soft, ethereal appearance.

# Defining Bioethics

# How and Why Enterprises Must Tackle Ethical AI

05-27-21 | TELUS INTERNATIONAL

## The ethical use of data in artificial intelligence

Even with the best of intentions, individuals may not recognize their own biases. Here's what must be done to make AI effective—and fair—to all.

Strategy

### A Practical Guide to Building Ethical AI

by Reid Blackman

Vox

Google says it's committed to ethical AI research. Its ethical AI team isn't so sure.

Forbes

Jun 10, 2021, 11:54am EDT | 556 views

### Protecting The Human: Ethics In AI

MENU

The Sydney Morning Herald

SUBSCRIBE

Sponsored Business Workplace The University of Queensland Business School

### New research suggests businesses can combat data distrust with ethical approach

### SOFIC NEWS: Microsoft, SOCOM Highlight Need for Ethical Use of Artificial Intelligence

### Ethical Frameworks for AI Aren't Enough

by Andrew Burt

### Why a data-driven world should consider an oath of ethics

The Hippocratic Oath is perhaps the most widely known of Greek medical texts – and such an

### Ethics of AI: Benefits and risks of artificial intelligence

The increasing scale of AI -- in terms of the size of neural networks, their energy use, the size of data sets, and the prevalence of the technology in society -- is raising the stakes for major ethical questions.

Artificial intelligence

### Stop talking about AI ethics. It's time to talk about power.

### The term 'ethical AI' is finally starting to mean something



# Ethical Principles Related to AI/ Machine Learning

# Respecting and Protecting Human Autonomy

A purple padlock with a gold heart-shaped metal plate on its front, set against a blurred background of other padlocks.

# Privacy, Confidentiality, and Informed Consent

A **lot** of information can be obtained from personal data.

Many argue that people should have the **right** to keep their data private—opting out from sharing should be default.

Privacy breaches have a **disproportionate** negative impact on marginalized communities.

# Social Good, Safety, and the Public Interest



AI technologies should not **harm** individuals and communities.



Measures of quality control and **quality** improvement in the use of AI over time should be available.



The use of AI should not result in harm that could be avoided using an **alternate** practice or approach.

# Transparent, Explainable, and Interpretable

- **Transparent:** sufficient information be documented before the design or deployment of an AI system to facilitate meaningful public consultation and debate.
- **Explainable:** AI should have “explainable architecture” and interpretable according to the capacity of those to whom they are explained.
- **Interpretable:** the transparency of the main prediction algorithm to ‘explain’ the computations of the first.

# Corporate Influences on Healthcare

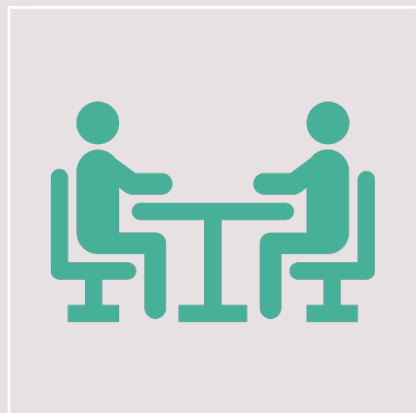
- Individual **financial** interests (e.g., payment for services, ownership of stocks) may influence organizational-level decisions.
- Operational interests of the organization that might be in tension with the **duty** of healthcare professionals to advance the best interest of each patient.
- **Disclosure** practices are (likely) not sufficient.

Jaremko, et al (2019)





# Responsibility and Accountability



Stakeholders should ensure that AI can perform specific tasks, used under **appropriate** conditions, and by appropriately trained people.



Appropriate mechanisms should be available for questioning and for redress for individuals and groups **adversely** affected by automated algorithms.

# Inclusiveness and Equity

- AI should be designed to encourage **equitable** use and access irrespective of age, sex, gender, income, racialized status, ethnicity, sexual orientation, ability, or other characteristics.



# Responsivity and Sustainability

## Responsiveness

- Designers, developers, and users should continuously, systematically, and transparently assess AI applications during use.
- Determines whether AI responds adequately and appropriately according to expectations and requirements

## Sustainability

- Should be consistent with global efforts to reduce the impact of human beings on the Earth
- Governments and companies to address anticipated disruptions to the workplace

# “The Digital Medicine Panopticon”

- “Smart pills” and **surveillance** tracking technologies (e.g., Abilify MyCite™).
- Continuous surveillance may threaten the **trust** in therapeutic relationships, undermine autonomy
- We ought to consider how these technologies are **embedded** in larger systems of surveillance, governance, and control.

Foucault (1977), Klugman, et al (2018); Guta et al (2018); Swartz (2018)

# What Do You Consider the Most Important Ethics Principle related to AI—and Why?

- Inclusiveness and Equity
- Privacy, Confidentiality, and Informed Consent
- Respecting and Promoting Human Autonomy
- Social Good, Safety, and the Public Interest
- The 'Digital Panopticon'
- Transparent, Explainable, and Interpretable

A silver key and a green house-shaped tag are placed on a wooden surface. The key is on the left, and the tag is on the right. The tag has a heart-shaped hole in the center. The text is overlaid on the image.

# Take Home Message

Ethics is **part** of AI from design to implementation.

# References

Char, D.S., Abràmoff, M.D., Feudtner, C. (2020). Identifying ethical considerations for machine learning healthcare applications. *The American Journal of Bioethics*, 20(11), 7-17.

Guta, A., Voronka, J., Gagnon, M. (2018). Resisting the digital medicine panopticon: Toward a bioethics of the oppressed. *The American Journal of Bioethics*, 18(9), 62-64.

Jaremko, J. L., Azar, M., Bromwich, R. et al. (2019). Canadian Association of Radiologists white paper on ethical and legal issues related to artificial intelligence in radiology. *Canadian Association of Radiologists Journal*, 70(2), 107-118.

Klugman, C. M., Dunn, L.B., Schwartz, J., Cohen, I.G. (2018) .The ethics of smart pills and self-acting devices: Autonomy, truth-telling, and trust at the dawn of digital medicine. *American Journal of Bioethics*, 18(9), 38-47.

Murphy, K., Di Ruggiero, E.D., Upshur, R., Willison, D.J., Malhotra, N., Cai, J.C., ... Gibson, J. (2021). Artificial intelligence for good health. A scoping review of the ethics literature. *BMC Medical Ethics*, 22, 14.

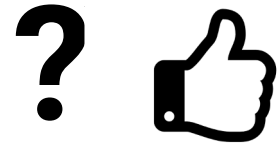
Swartz, A.K. (2018). Smart pills for psychosis: The tricky ethical challenges of digital medicine for serious mental illness. *The American Journal of Bioethics*, 18(9), 65-67.

Whittlestone, J. Nyrup, R. Alexandrova, A. Dihal, K. Cave, S. (2019) Ethical and societal implications of algorithms, data, and artificial intelligence: a roadmap for research. London: Nuffield Foundation.

World Health Organization (2021). *Ethics and governance of artificial intelligence for health: WHO Guidance*. Geneva: World Health Organization.

Yuste, R., Goering, S. (2017). Four ethical priorities for neurotechnologies and AI. *Nature*, 555, 159-163.

# Remember - many ways to engage



(during sessions)  
Use the chat or  
the ask question!



You can always return to the  
session and re-watch the videos  
after the session ends



come chat with us in KCNI  
Summer School Slack :)



virtually meet with us  
in gather.town



Tell us how the session went (post session survey):

<https://forms.gle/ji18qLMZEZ9L16Ln6>



[KCNISchool@camh.ca](mailto:KCNISchool@camh.ca)



# Thank You

[daniel.buchman@camh.ca](mailto:daniel.buchman@camh.ca) | @DanielZBuchman