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.
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.

<table>
<thead>
<tr>
<th>Please don’t say</th>
<th>Instead say</th>
</tr>
</thead>
<tbody>
<tr>
<td>depressives</td>
<td>people with major depressive disorder</td>
</tr>
<tr>
<td>schizophrenics</td>
<td>people with schizophrenia</td>
</tr>
<tr>
<td>addict</td>
<td>person with substance use disorder</td>
</tr>
<tr>
<td>drug abusers</td>
<td>people who use drugs</td>
</tr>
</tbody>
</table>
Remember - many ways to engage

- 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
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
Defining Bioethics
How and Why Enterprises Must Tackle Ethical AI

A Practical Guide to Building Ethical AI
by Reid Blackman

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

Protecting The Human: Ethics In AI

New research suggests businesses can combat data distrust with ethical approach

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 oath for artificial intelligence.

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.
Ethical Principles
Related to AI/
Machine Learning
Respecting and Protecting Human Autonomy

WHO (2021)
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.

WHO (2021); Yuste, Goering, et al. (2017)
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.

Char, Abràmoff, & Feudtner (2020); Whittlestone, et al. (2019); WHO (2021)
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.

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.

Char, Abràmoff, & Feudtner (2020); Murphy, et al., (2021); WHO (2021)
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.

WHO (2021)
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.

WHO (2021)
“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.

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
Take Home Message

Ethics is part of AI from design to implementation.
References


Remember - many ways to engage

- **crowdcast**
  - (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.

- **Slack**
  - come chat with us in KCNI Summer School Slack :)

- **Gather**
  - virtually meet with us in gather.town

Tell us how the session went (post session survey):
https://forms.gle/ji18qLMZEZ9L16Ln6

KCNISchool@camh.ca
Thank You

daniel.buchman@camh.ca | @DanielZBuchman