

Ethics & Equity in AI

APS360—March 28th, 2019

Jointly prepared by Prof. Lisa Zhang, Dr. Cindy Rottmann & Prof. Patricia Sheridan



UNIVERSITY OF TORONTO
FACULTY OF APPLIED SCIENCE & ENGINEERING
Institute for Leadership Education in Engineering

Acknowledgement of Traditional Territory

We wish to acknowledge this land on which the University of Toronto operates.

For thousands of years it has been the traditional land of the Huron-Wendat, the Seneca, and most recently, the Mississaugas of the Credit River.

Today, this meeting place is still home to many Indigenous people from across Turtle Island and we are grateful to have the opportunity to work on this land.

Roadmap for the class

1. Mindful listening exercise
2. Defining terms
3. Activity 1: AI Ethics in the news
4. Activity 2: Model cards
5. Debrief

Mindful Listening Exercise

- Pair up within your teams, assign A & B
- A answers question, B listens (2 min)
 - What excites you about AI? Share an experience of belonging/non-belonging in the AI community.
- B answers question, A listens (2 min)
- Discuss what you learned from each other (2 min)

Summary of student responses

- Use of data (Origin of training data? Consent?)
- Do all AI projects have ethical implications?
- Framing ethics as commercial utility
- Copyright
- Uneven social implications of false positives/negatives
- Surveillance

Defining terms

Ethics: Moral principles that govern a person's or a group's behaviour.

Diversity: Recognizing the demographic and experiential heterogeneity of a group.

Equity: Creating opportunities and removing barriers to address historic and current disadvantages faced by marginalized groups

Inclusion: Building a welcoming culture in which people feel they belong.

Reflexive Principlism (Prof. R. Irish)

- One of many ethical frameworks used by engineers
- Interdisciplinary collaboration between an engineer and philosopher (Brightman & Beever)
- Objective: To determine a pragmatic ethical approach for engineers

4 principles of Reflexive Principlism

1. Autonomy

- Respecting people as decision-makers

2. Beneficence

- Providing benefits

3. Justice

- Fair distribution of benefits, risks and costs

4. Non-maleficence

- Preventing harm

Professional Codes of Ethics

PEO (Professional Engineers of Ontario)

- Code of Ethics: see sections 77 & 72 of Regulation 941

IEEE (Institute of Electrical and Electronic Engineers)

- Code of Ethics: see IEEE policies

Activity 1: AI ethics in the news (D. Raji)

- 1) Pick one article & apply RP framework
 - AI in health care
 - Self-driving car
 - Project Maven
 - Facial recognition

- 1) Who is impacted/responsible?
(client/user, community, designer)

Debrief activity 1

1) Most salient ethical consideration?

- autonomy,
- beneficence,
- Justice,
- nonmaleficence

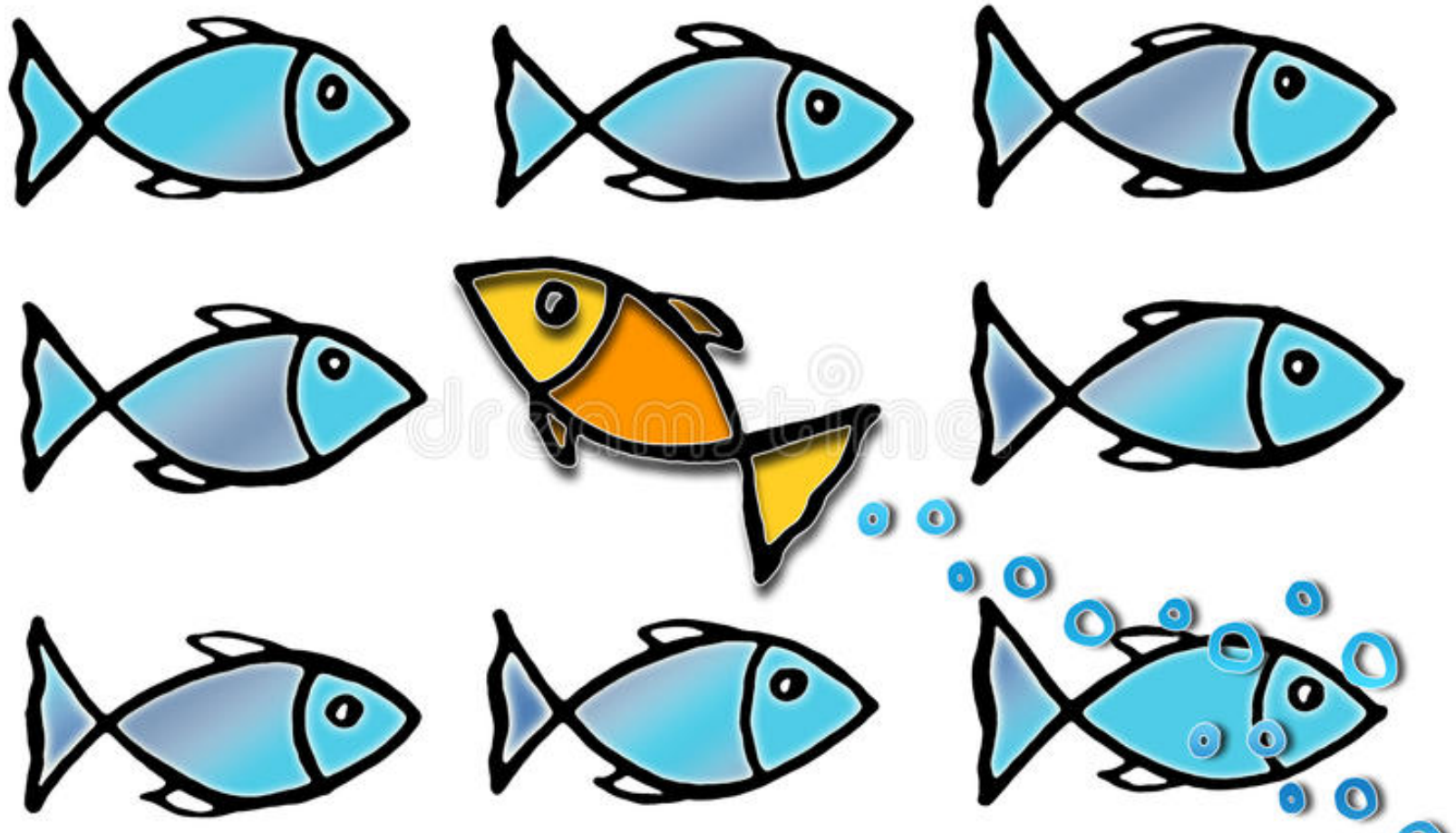
Activity 2: Model cards (Mitchel et al.)

1. Purposes: transparency & forethought
2. You do!
 - Model details
 - **Intended use**
 - **Factors**
 - Metrics
 - Evaluation data
 - Training data
 - **Ethical considerations**
 - Caveats & recommendations

Debrief activity 2

- 2-3 volunteers to share with the class
- Accessibility panel
 - AI (Zhang)
 - Engineer (Sheridan)
 - Citizen (Rottmann)

It can be difficult to swim against the tide of engineering culture



And sometimes, a little dangerous...



Especially when you're marginalized

Sexism

- 5/7 women; 1/8 men

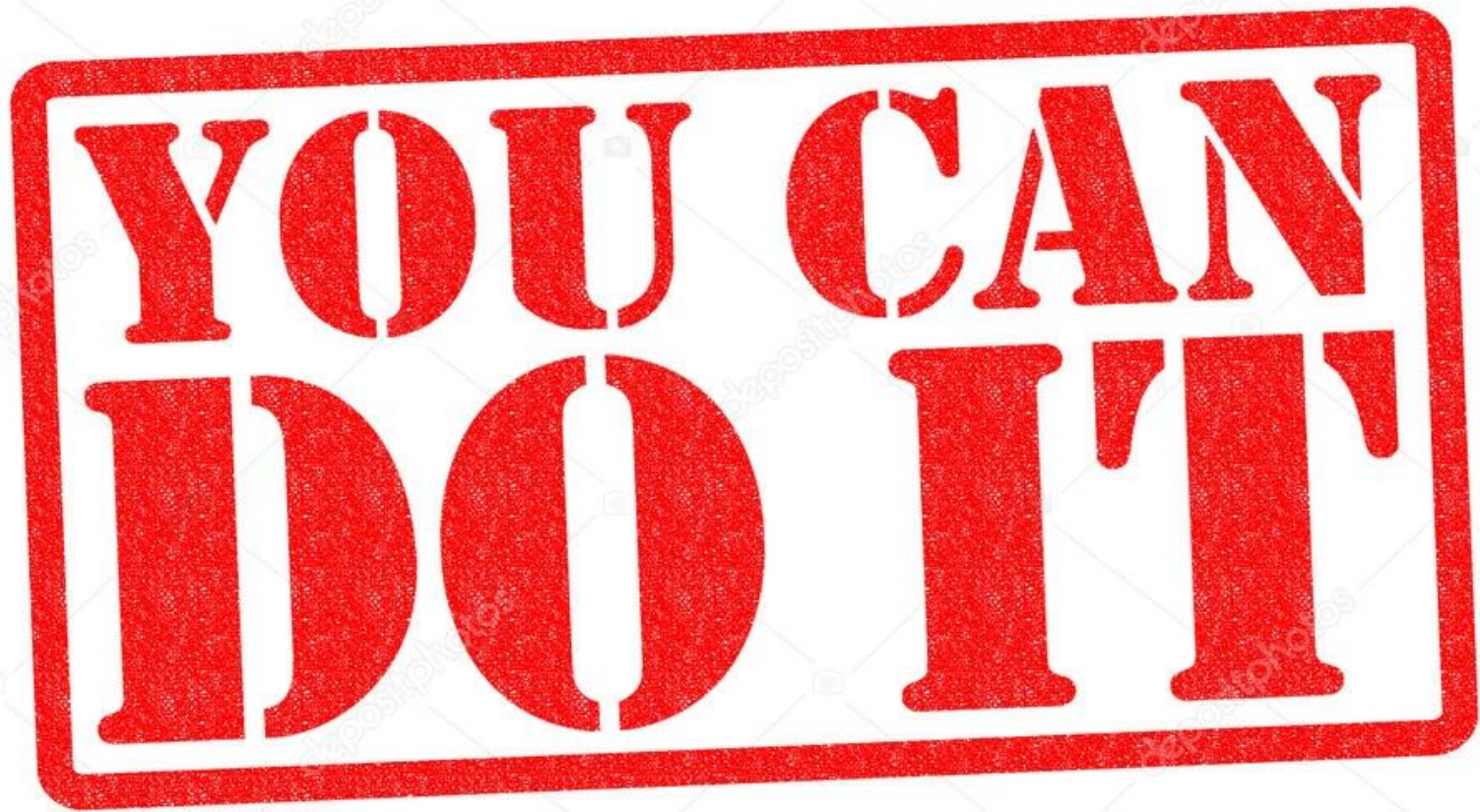
Racism

- 2/4 URM; 0/11 white

Homophobia

- 3/3 LGBTQ; 0/12 heterosexual

But...



Reflections

- Why is it important for you as an engineer to embody ethics and equity?
- Particular challenges/concerns in AI?
- How can you convince peers that it is important?

For more information:

cindy.rottmann@utoronto.ca

patricia.sheridan@utoronto.ca

Words of wisdom

“No one can be in the world, with the world, and with others and maintain a posture of neutrality. I cannot be in the world decontextualized, simply observing life.”

“Critical reflection on practice is a requirement of the relationship between theory and practice. Otherwise theory becomes simply "blah, blah, blah," and practice, pure activism”.

- Paolo Freire