Lecture 18: Responsibility and Recap

CMSC 25910

Spring 2022

The University of Chicago

Responsibility

Who is responsible for…

- …a bug in code?
    - …a security flaw caused by a bug in code?
        - …the Heartbleed vulnerability?

Who is responsible for…

- …an ML model?
  - …an ML model that makes an automated decision?
    - …an ML model that causes someone to be jailed incorrectly?

Who is responsible for…

- …a self-driving car?
  - …a self-driving car that crashes?
    - …a self-driving car that crashes after being cut off by a human?

Who is responsible for…

- …a robot?
  - …a robot that is offensive?
    - …a robot that bullies someone and causes them harm?

Who is responsible for…

- …automation software?
  - …automation software that saves a company money?
    - …automation software that puts someone out of work?

Consideration of Ethics

Joe Redmon @pjreddie · Feb 20, 2020

"We shouldn't have to think about the societal impact of our work because it's hard and other people can do it for us" is a really bad argument. twitter.com/RogerGrosse/st…

> Roger Grosse @RogerGrosse
> Replying to @kevin_zakka @hardmaru
>
> To be clear, I don't think this is a positive step. Societal impacts of AI is a tough field, and there are researchers and organizations that study it professionally. Most authors do not have expertise in the area and won't do good enough scholarship to say something meaningful.

Joe Redmon
@pjreddie

I stopped doing CV research because I saw the impact my work was having. I loved the work but the military applications and privacy concerns eventually became impossible to ignore.twitter.com/RogerGrosse/st…

What will you do?

- Real-world scenarios rarely have a perfect decision. Instead, you'll make and participate in myriad decisions that, together, will have consequences, good and bad.

Recap

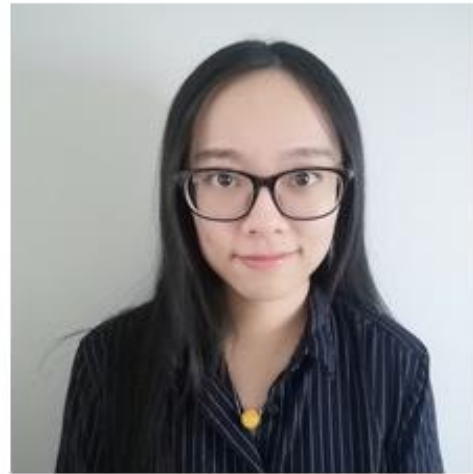# Topics Covered in the Last Nine Weeks

- The influence of UI / dark patterns
- Anonymity
- Privacy law / regulation / philosophy
- Data lifecycles / data rights (erasure, access, portability)
- Ethical experiments
- Data quality, data errors, and statistical pitfalls
- Generalization in ML
- Fairness and bias in ML and NLP
- Documenting, auditing, and explaining models
- Differential privacy
- Privacy engineering techniques
- Data formats, accessibility, and internationalization
- Wasting human's time; wasting energy to prove work
- Tracking and inference, surveillance
- Responsibility and values

THANK YOU!

Course Staff

Blase Ur
Instructor

Weijia He
TA

Kevin Bryson
TA