



MSDS640 Week 1: Introduction

MSDS 640: Ethics, Privacy and Social Justice Issues

- A writing- and analysis-focused course as opposed to technical
- Clear communication of ideas is important for everyone
- We'll learn about ethical concerns within data science and sometimes how to avoid them

FTE and readings each week

- At a minimum, read the FTE each week, and try to skim some of the linked articles.
- There are lots of extra optional readings and videos. You do not have to read and watch everything linked in the FTE.

Breakout Group (10 min)

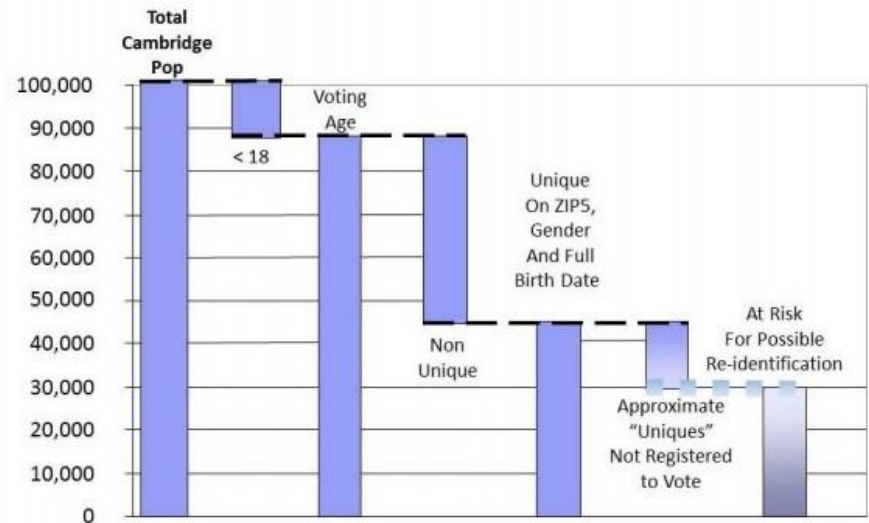
We want to get to know you!

In a breakout group share with your classmates the following:
Then we'll introduce ourselves to the whole class.

1. Your contact information: name, email address, location, and time zone (phone number optional).
2. Your job title.
3. Hobbies, interests, and any other relevant information about yourself for the course.
4. Include a brief reflection on one or two of the [Ignatian Jesuit Values](#) (ethics, social justice, men/women for others, God in all things, and/or global awareness) and reflect on what Jesuit Values means to you as they relate to this course.

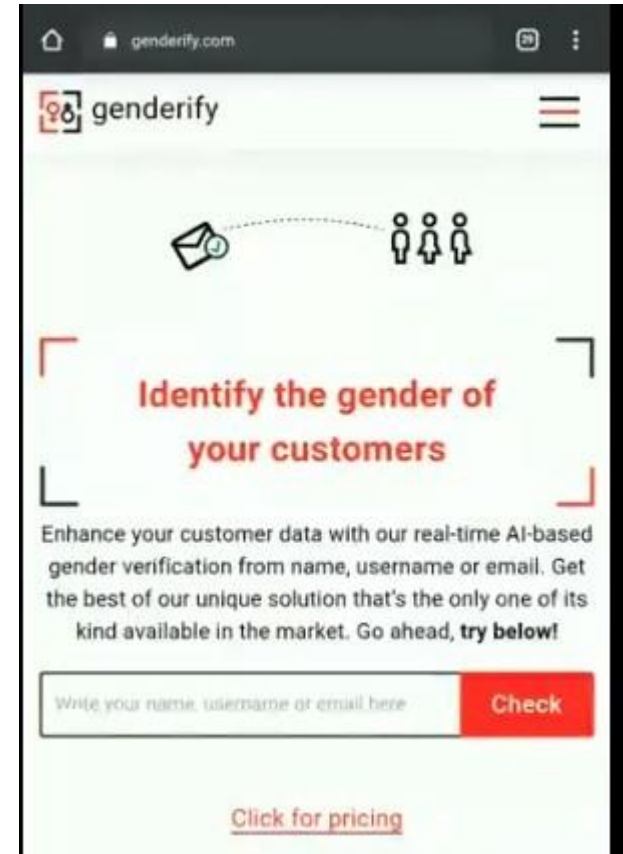
Example of some ethical concerns

- De-identifying anonymous data
 - e.g. identifying MA governor from “anonymous” health insurance records
 - k-anonymity and l-diversity



Example of some ethical concerns

- Algorithm bias
 - e.g. facial recognition, police algorithms, gender bias
- Data collection and privacy



Why so much ethics curriculum?

- Our idea is ethical people make for a better world
- Recent examples of ethical problems with police, finance, government policy, corporate policy, data science, etc
 - Wirecard fraud; stock buybacks for CEOs to get bonuses
 - Check out for more finance fraud like [Wirecard](#)
 - Corporate Bailouts, NSA (Snowden leaks)
 - Lack of sick leave (e.g. many fast food restaurants before the pandemic)
 - [Police algorithms](#), [facial recognition](#), [NSA's SKYNET](#)
 - [ACLU pizza ordering video](#)
 - [Florida data scientist and COVID data](#)

Social justice definition

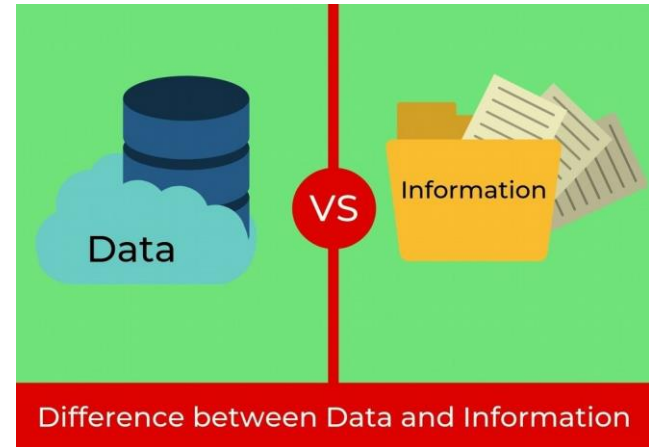


- justice in terms of the distribution of wealth, opportunities, and rights within a society (other definitions [here](#) and [here](#))
- For example, the wealth gap between white and black families in the US could be equalized by giving each black family \$800k, based on total wealth of white vs black families ([source](#))
- Also see the Minnesota paradox (e.g. discussed on the NPR Indicator podcast)
 - Redlining still has impacts for blacks today ([source](#))

[Image source](#)

Data vs Information

- In our context, data are the numbers and text we collect; information is distilling the data into “actionable insights”, or something we can use to understand the situation better and/or make decisions.



[Differ.com explanation](#)

[Stechies.com explanation](#)

[Image source](#)

Conversation Theme

Amazon has become one of the largest companies in the world. It employs about 1.3 million people worldwide in both blue-collar and white-collar jobs. One of the hotly contested issues concerning Amazon is its tax avoidance by reinvesting any profits in other business ventures. Further criticisms include profiting over smaller sellers, collecting fees on each item it sells, and copying successful sellers' products, and driving prices further down. See article [here](#).

- Discuss an ethical issue related to data collection or analysis a data scientist may face.
- Provide an example, and discuss how the ethical issue can be mitigated.

Assignment

ACM Code of Ethics and Professional Conduct

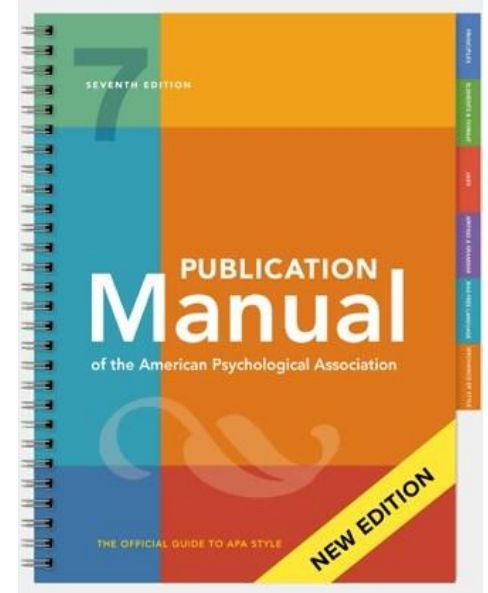
- Overview, description, explanation/example, classification of each ethical principle (1.1 through 1.7) from the [code of ethics](#)
- No page limit or requirement, but aim to write something substantial and meaningful in as short a length as possible
- “I would have written a shorter letter, but I did not have the time.” – Blaise Pascal

Association for Computing in Machinery (ACM) Code of Ethics Grading Rubric

Criteria	Points
Grammar, formatting, organization, writing quality and fluency	50
Classification and description of each ethical principle	50
Total Points	100

APA format

- Papers should be in APA format
 - Templates can be found here:
 - apastyle.apa.org
 - owl.purdue.org
 - overleaf.com
- A [citation manager](#) is recommended



Next week (week 2)

- We'll write on design and ethical considerations for an AI judicial assistant.
- We'll cover the data science lifecycle