Data Ethics In-Class Activity (May 18)

Today we will discuss two of the Ethics subtopics, Data Ownership and Fairness. Refer to the corresponding articles listed in [the previous activity document](https://docs.google.com/document/d/13J_-nHDxBU_2zlobJNP78J2SnImEXvBPPqFicYffYHw/edit#). There is no need to submit any written report about the topics or about the discussion. Just be ready to participate fully in the class discussion.

## Data Ownership

Your job is to discuss the Data Ownership topic first within your breakout group (for approximately 30 minutes) and then join the main room for a full class discussion of the topic.

Within your group:

1. Who read any of the articles about Data Ownership?

Prabhu Marappan

1. Which article(s) did you read?

Story of the data

1. Each person who read any of the articles should take a few minutes to give an overview of the article, specifically:
   1. Who wrote the article, what is their role or point of view?

Robin Kurkez

* 1. What are the main points of the article?

Ownership of the data

* 1. What are the strong points of the article?

The owner of the data has all the rights even though they have signed a consent form

* 1. What are the weak points (if any)?

None of the big companies have rights to share or use the data

* 1. What did you learn from it or take away from it?

Who has the ownership of data

* 1. Do you recommend that your teammates read this article?

Yes

1. All others in the group then should discuss and ask questions about the article.
2. Discuss the following questions:
   1. For each person on the team, have you ever encountered instances where your personal data was being used in a surprising way?

No

* 1. Who do you think should “own” data about you?

Us, our data is always owned by us

* 1. Can you propose an ethical yet profitable solution to the data ownership issue?

The parts of the data can be used in data analysis which is used for experimental learning and other purposes

* 1. List all of the general categories of data about a person that might be collected automatically by computer systems. Rank them by how comfortable you are with somebody else owning that data (and using it for whatever legal purpose they choose).

Address, gps location, sex, name, their day to day habits are recorded, bank accounts

## Fairness

Your job is to discuss the Fairness topic first within your breakout group (for approximately 30 minutes) and then join the main room for a full class discussion of the topic.

Within your group:

1. Who read any of the articles about Fairness?

Ans:

Myself (Manish Manchegowda)

1. Which article(s) did you read?

Is pokemon go a racist?

1. For each article read by at least one person in the group:
   1. Who wrote the article, what is their role or point of view?

Ans:

Allana Akhtar, the author has collected data from multiple sources and giving critics in the author point of view

* 1. What are the main points of the article?

This article is before the release of the game worldwide, which means must be still under development. This article tells how the game was advantage to the white people rather than being equality

* 1. What are the strong points of the article?

It all says that there must be always people of colors in the organisation to balance the things

* 1. What are the weak points (if any)?

No

* 1. What did you learn from it or take away from it?

Nothing, because its a very old article and does not even mean anything

* 1. Do you recommend that your teammates read this article?

No, it is not a worth

1. All others in the group then should discuss and ask questions about the article.

Everyone in the group was asking whether this article is valid or not, I do not think this article is good enough

1. Discuss the following questions:
   1. Are you aware of any other issues/reports of unfairness related to data analysis? Bias? Other types of incorrect conclusions?

Actually people have collected numerous data of the players playing this game and analyzed the data and came to conclusion regarding this game

* 1. What can a Data Engineer do to prevent or lessen the impact of unfairness?

The data must be valid, I do not think the data collected in this article is not valid as the the game was not released world wide during this article

* 1. How would you actually find/seek fairness issues earlier?

By checking the validity of the data, validity and fairness goes hand on hand

* 1. When is appropriate/inappropriate to use “surrogates” in data analysis?

When the data is not valid

* 1. Was Pokeman Go racist?

Can not come to the conclusion, as the data was very old and it was before game was released world wide. Need newer data for this conclusion

* 1. How would you notice/find the type of bias found in Pokeman Go? How could they have noticed this before it turned into a news article?

I do play this game, I do not think there is such a bias currently

* 1. Can a Data Engineer produce/provide bias-free data?

Yes

* 1. Should a Data-related team include Social Scientists?

Yes, must be