Data Ethics In-Class Activity (May 18)

Today we will discuss two of the Ethics sub-topics, 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?
2. Which article(s) did you read?
3. 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?
   2. What are the main points of the article?
   3. What are the strong points of the article?
   4. What are the weak points (if any)?
   5. What did you learn from it or take away from it?
   6. Do you recommend that your teammates read this article?
4. All others in the group then should discuss and ask questions about the article.
5. 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?
      1. **Arturo: Targeted ads in social media.**
      2. **Michael: Targeted ads in social media.**
      3. **Al: Targeted ads in social media.**
   2. Who do you think should “own” data about you?
      1. **Arturo & Al & Michael: It depends on what kind of data we are talking about.**
   3. Can you propose an ethical yet profitable solution to the data ownership issue?
      1. **Al: Surveys**
      2. **Arturo : I don't know**
      3. **Mickael: companies should ask consent to use other peoples data**
   4. 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)
      1. **Government records.(5)**
      2. **Shopping list. (1)**
      3. **Entertainment preferences.(2)**
      4. **Social media interests. (4)**
      5. **Medical records (6)**
      6. **Traffic record (3)**

## 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?
2. Which article(s) did you read?
3. 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?
   2. What are the main points of the article?
   3. What are the strong points of the article?
   4. What are the weak points (if any)?
   5. What did you learn from it or take away from it?
   6. Do you recommend that your teammates read this article?
4. All others in the group then should discuss and ask questions about the article.
5. 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?
      1. **Al: Social media biases**
      2. **Arturo: Xbox connect facial recognition.**
      3. **Michael: social media**
   2. What can a Data Engineer do to prevent or lessen the impact of unfairness?
      1. **Al: Consider all the people from different backgrounds**
      2. **Arturo: Consider all the people from different backgrounds.**
      3. **Michael: Consider all the people from different backgrounds.**
   3. How would you actually find/seek fairness issues earlier?
      1. **Social Testing it before releasing it to the community**
   4. When is appropriate/inappropriate to use “surrogates” in data analysis?
      1. **It is appropriate when it eliminates biases**
   5. Was Pokeman Go racist?
      1. **Yes**
   6. 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?
      1. **By seeing how many stops in each area.**
   7. Can a Data Engineer produce/provide bias-free data?
      1. **no**
   8. Should a Data-related team include Social Scientists?
      1. **Yes**