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. 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?
   Zhengmao Zhang
2. Which article(s) did you read?
   The story of data, Part 3: Who owns it?
3. Each person who read any of the articles should take a few minutes to give an overview of the article, specifically:
   a. Who wrote the article, what is their role or point of view?
      Kurzer, Robin
   b. What are the main points of the article?
      Discuss who actually has the right to process the data when the user is using the software.（Users, service providers, or other subjects.）
   c. What are the strong points of the article?
      i. Once a consumer shares data, is the genie out of the bottle? - No matter what happens, for most users, the data they generate has ownership rights, including the right to deletion. 不论发生什么事，对于大多数用户来说，他们产生的数据拥有所有权，包括删除的权利。
      ii. Who owns data within the companies that are collecting and processing it? - The agency is in fact using this internally generated information, but legally it belongs to the user. 机构事实上上在使用这些内部产生的信息，但是按照法理上而言，这属于用户。
      iii. Still, there's no consensus
      Multiple parties are fighting for ownership of user data. 多方都在争取得到用户数据的拥有权。
   d. What are the weak points (if any)?

   e. What did you learn from it or take away from it?
      In a nutshell, Users have all rights to their data.

    f. Do you recommend that your teammates read this article?
4. All others in the group then should discuss and ask questions about the article.
    a. Which third party platform may also own the data? Is there any example?
        i. AWS, MongoDB
5. Discuss the following questions:
    a. For each person on the team, have you ever encountered instances where your personal data was being used in a surprising way?
        i. Application spy on me
    b. Who do you think should “own” data about you?
        i. The application itself can own the part I provided to them
    c. Can you propose an ethical yet profitable solution to the data ownership issue?
        i. Countries with strong governments can set up corresponding institutions to store data, and national legislatures should promulgate corresponding laws.
    d. 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)
        i. General information - name, age
        ii. Browsing history

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:

6. Who read any of the articles about Fairness?
    a. Yixuan
7. Which article(s) did you read?
    a. Widely-used healthcare algorithm racially biased
8. For each article read by at least one person in the group:
    a. Who wrote the article, what is their role or point of view?
        i. There is racially bias in algorithm
    b. What are the main points of the article?
        i. The main point of the article is to explain why and how the algorithm can lead to a racially bias when they finding the sickest patient
    c. What are the strong points of the article?
        i. The strong point of this article is that it explain very clearly why it happens
    d. What are the weak points (if any)?

        i. Make more clear about how the algorithm works/lead to the situation
        ii. Maybe provide a possible solution
    e. What did you learn from it or take away from it?
        i. I learned that there is bias in racial when different races tend to act different when they deal with disease
    f. Do you recommend that your teammates read this article?
        i. Yes, it's a article that explain thing very clear and shows a great point
9. All others in the group then should discuss and ask questions about the article.
    a. How do they solve it?
        i. No solution in the article.
10. Discuss the following questions:
    a. Are you aware of any other issues/reports of unfairness related to data analysis? Bias? Other types of incorrect conclusions?
        i. No.
    b. What can a Data Engineer do to prevent or lessen the impact of unfairness?
        i. Consider multiple data stead of only one data
    c. How would you actually find/seek fairness issues earlier?
        i. Consider more of how the real world works
    d. When is appropriate/inappropriate to use "surrogates" in data analysis?
        i.
    e. Was Pokeman Go racist?
        i. (?)
    f. 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. Visualization
    g. Can a Data Engineer produce/provide bias-free data?
        i. Possibly? It could be hard
    h. Should a Data-related team include Social Scientists?
        i. Yes!