

Data Ethics In-Class Activity (May 23)

Today we will discuss two of the Ethics subtopics, Data Privacy and Validity. 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. Be sure to stay in your breakout room even if you finish your group discussion early, as there is a class discussion at the end of the session.

Data Privacy

Your job is to discuss the Data Privacy 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 Privacy?
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:
 - a. Who wrote the article, what is their role or point of view?
 - b. What are the main points of the article?
 - c. What are the strong points of the article?
 - d. What are the weak points (if any)?
 - e. What did you learn from it or take away from it?
4. All others in the group then should discuss and ask questions about the article.
5. Be ready to discuss the following questions with the full class
 - a. What is the GDPR?
 - b. GDPR is a European effort, how does it relate to the USA?
 - c. How might a Data Engineer be involved in GDPR compliance?
 - d. Discuss the following questions:
 - i. Popups everywhere. It's annoying and the average internet user has no idea how to control/configure data privacy consent, so they just agree to everything.
 - ii. Companies are scared, so they are spending bajillions protecting themselves. Bajillions that could be spent on things that actually benefit customers.
 - iii. The whole thing is toothless. Only Ireland can bring an actual judgment, and they are in the pocket of big tech. So there have not been many significant cases or judgements so far.
 - iv. It requires private data to be transparent and easily accessible by the users, and that makes it easier for hackers to obtain private data by impersonating users.

Validity

Discuss the Validity topic first within your breakout group (for approximately 30 minutes) and then join the main room for a full class discussion.

Within your group:

6. Who read articles about Validity?
7. Which article(s) did you read?
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?
 - b. What are the main points of the article?
 - c. What are the strong points of the article?
 - d. What are the weak points (if any)?
 - e. What did you learn from it or take away from it?
9. Discuss the following questions:
 - a. The articles list many problems with data validity. Which of these problems could be helped by a Data Engineering approach?
 - b. What specifically could/should a Data Engineer do to address the challenges listed in these articles?

Submit

Create a copy of this document (or create a new document if you prefer), and use it to answer the following question.

For each of the four major areas of Data Ethics, mention a situation that you have experienced that involved the corresponding area of Data Ethics. Say whether or not (in your opinion) the issue was handled satisfactorily. Finally, state how you might improve the handling of Data Ethics in similar situations in the future.

Answer:-

Ownership:- I have personally faced this issue as when I took a gym membership, they used a fingerprint mechanism to check in (this was before COVID 19). So as they just saved my fingerprint and I have to sign there terms and conditions to get a membership, I honestly didn't find a way to tackle this you have to agree their terms and conditions else you won't get the membership so it felt like a deadlock situation to me. For improvement I would suggest that the user should have the ownership and should have the power to remove his fingerprint from the organization which is involved in it.

Fairness:- I personally have not experienced data fairness issue, I know its out there and could be applicable to me as well but I think I am not aware of it.

Privacy:- This is an issue that I believe everyone faces on daily basis and its very vulnerable to leak as well. I face a lot of issues like this as I get calls from all types of businesses trying to sell me their services and its because my information is probably sold to some 3rd party. In order to combat this, I take a lot of measures some of them are using dummy email address, not giving out my personal phone number to everyone, making my social media accounts private and disabling my location on my devices. For improvement I would say use VPN as much as you can but it can act like a double edged sword as well.

Validity:- This is again an issue that is usually faced by everyone about data validity more specifically speaking its biased data, whenever we go to buy something online we look at the reviews and star rating, which are not accurate most of the time. They are being manipulated in in favor of the seller which leads us to buying wrong product. For improvement it's a very difficult situation as either party can be lying about the issue. But nevertheless the review section is open to both the parties and both can explain the situation if any issue arises in the purchase of a product.

Use the in-class assignment submission form to submit your response(s).