Michael Campoverde

Antonio Johnson

Dmytro Lukinchuk

**Q1)** Is this data still considered private even though it does not include the names of individuals? Why or why not?

I believe this data can still be considered private even though it does not include the names of individuals because the data being displayed is shown as an anonymous user. If it had the person’s name, email, or IP address then it would not be considered private.

**Q2)** Which columns in the dataset might reveal identifiable information about individuals, potentially violating data privacy? List and briefly explain why.

I would consider waist cm or body mass index being the only two that might reveal identifiable information about individuals. The main reason I talk about this is because these measurements can pinpoint and help provide people with what you might look like and find ways to easily identify someone using these numbers. Like with what police do when chasing a person as they identify the person with a rough estimate of their height or weight so the other police can be able to identify them easily too.

**Q3)** This data is being used to help predict if an individual has type-2 diabetes or not. Is this an appropriate use of the data? Why or why not?

Yeah, I would say this is appropriate use of the data because it lists most of the common routes that people with type 2 diabetes might go. It then sets them into categories with age, gender, weight, and waist circumference and this helps us see the common trends with people who do have diabetes or do not have it.

**Q4)** Consider the following scenarios:

* Scenario A: A researcher wants to use the dataset to study the correlation between weight and hypertension in middle-aged adults without obtaining consent from the individuals.

Is this an appropriate use of the data? Why or why not?

I would consider this appropriate use of the data because it does not use it for malicious purposes or money purposes like advertisement and is only using it to gather data and for research purposes.

* Scenario B: A pharmaceutical company requests access to the dataset, including personal identifiers, to market diabetes medication directly to individuals.

Is this an appropriate use of the data? Why or why not?

It is not appropriate because instead of using this data for research purposes the pharmaceutical company is using it to gain economic value and the specific individuals never asked to get advertised for diabetes medication from the pharmaceutical company and they breached access just to get money.

**Q5)** Suppose a new column, "Full Name," is added to the dataset.

1. In what contexts might it be appropriate to share the full dataset, including Full Name, and with whom?

I believe it would be most appropriate to share the full dataset with your doctor or therapist whom you trust your personal information to get proper medical care. With a full dataset your doctor will be able to collect data about you and get you the best treatment.

1. In what contexts would it be inappropriate to share this information? Discuss the potential risks to individual privacy.

It would be inappropriate to share this information if there is no limit to who it is being accessed. It would be preferred that professionals such as doctors could get access to information on the names of individuals.

**Q6)** Using the **ACM Digital Library**, find a technical article related to privacy concerns in diabetes data or medical data analysis

**Instructions:**

* Visit CCNY’s library website: [CCNY Libraries](https://library.ccny.cuny.edu/az.php).
* Navigate to the link for "**ACM Digital Library**."
* Click on “Advanced Search” next to the search bar.
* Change the search criteria to something like the below (Feel free to play around with this and use your own terms):

1. Choose an article and read the Abstract, Introduction, Motivation/Problem Statement, and Conclusion.

Wang, Yinzongxiao, and Huaru Wang. *Application of Telemedicine in Disease and Related Problems*. Vol. 41, 20 Oct. 2023, pp. 756–759, <https://doi.org/10.1145/3644116.3644243>. Accessed 4 Nov. 2024.

1. Summarize the key privacy insights discussed in the article.

Focused on the use of privacy during patient data transmission and they importance of equipment that doctors use to help patients.

1. How do the findings relate to the dataset we are examining?

They both share valuable information that doctors may use to treat their patients

1. What recommendations do the authors make regarding privacy preservation in medical datasets?

They promote remote equipment to regulate privacy preservation and more laws that china should use to guarantee the patients treatment.