«Human-Centered Data Science»

Exercise 6

Lars Sipos

Human-Centered Computing, Institute of Computer Science

Freie Universität Berlin

31.05.2022

Datasheets Exercise

https://github.com/FUB-HCC/hcds-summer-2022/wiki/06_exercise

Protected Attributes

List of Protected Attributes:

Fairness through unawareness

Creating a dataset

https://box.fu-berlin.de/s/CgsPnNLANQZFZ3d

Password: On the board

Datasheet Creation

First, get into groups. After that, we will divide the topics between all groups:

- Motivation:

- Composition (huge):

- Collection process (huge):

- Preprocessing / Cleaning / Labeling:

- Uses:

- Distribution:

- Maintenance:

Datasheet Creation

- Read the instructions on Github

- Follow the link on Github

- Answer the questions in your section(s) as best as you can

- If you don't know how to answer feel free to ask.

You have time until 11:20 AM

Datasheets Discussion

Next Time

you will have …

1. actively participated in the lecture

2. submitted the third programming assignment

Have fun!