Name: Esha Sharma

Hard Problem Area: Human Behavior / Policy

Overall Research Goal(s):
What do you ultimately hope to show with your research? This section can involve some jargon, but try to relate these goals to your broad impact section. Start with the larger goal(s) and narrow the scope towards your specific study (and specific goals, below).

To find interest / concern in Privacy topics. To find which privacy topics generate more concern. and find interest level in privacy topics.

Broad Impact:
Why is your research important? This section should avoid any technical jargon and should be meaningful to the general public. Try to keep this down to five sentences. This should be hierarchical: the broad impact decomposed into more specific impacts that connect your overall research goals to your more specific goals.

Because of the Privacy paradox: Behavioral decisions seem at odds with reported preferences. Example: Even though people are aware that they can read Privacy policies they don't do that.

Why does this happen?
We find the level of concern through surveys and survey phrasing may bias the respondents. Also respondents learn about Privacy through surveys. We are trying to find an objective way to find privacy concern.

Specific Research Goals:
Lay out the steps you are going to take to achieve your overall research goal. You can get technical here.

Using twitter to find Privacy concern is the overall goal. We first aim to build a classifier to find twitter privacy tweets. We have used 6 classifiers so far and have to evaluate the performance. Next we aim to cluster tweets based on privacy topics and then find the activity level on each topic and compare these activity levels.

Proposed Data Collection (if applicable):
What data will you collect to answer your research goals? How will you collect it? Will it be an observational study, randomized comparative experiment, or simulation study? Include potential biases and be prepared to explain how the data will achieve your specific research goals.

We aim to collect tweets over a period of a month or longer and apply our finding to this data collected.

Success Criteria:
How will you determine whether you satisfied your specific and overall research goals?

For the classifier the baseline is using tweets classified from twitter privacy filter. We have to yet come up with ways to evaluate the clustering to find the cohesiveness of the clusters made. Also, We have to come up with way to evaluate Recall for the classifiers.

**Anticipated Difficulties, Limitations, and Criticisms:**
What will make the above specific research goals difficult to achieve? How do you plan on dealing with these difficulties if they arise?

How do we find interest levels?
How do we evaluate the clusters we make?
How do we evaluate recall?