Keertana V. Chidambaram (12211266)
Assignment #8, MACS 30000, Dr. Evans

Problem 1: Identification risks in anonymized data

The two example articles reviewed in the answer are "A Face Is Exposed for AOL Searcher No. 4417749" and "Unique in the shopping mall: On the re-identifiability of credit card metadata". The first paper examines re-identification of users from an anonymized AOL searches dataset, while the second quantifies the risk of deanonymizing given a dataset containing credit card transactions.

(a) Although the publishing dates of the articles are years apart (one in 2006 and other in 2015) and though they deal with privacy risks of very different datasets, they have a similar approach to re-identification. Both the datasets are at first glance, anonymous. There is no explicit mention of people's names, addresses, or any other unique identification attributes. But the articles assert that the datasets provide a wealth of information about the users which makes re-identification possible.

A person's search history is a 'database of intentions' for the person, it provides a glimpse into the person's thoughts. Considering the ubiquitous nature of internet searches in everyday life, sensitive information like a person's location, interests, concerns, etc. can be easily inferred from it. The article found that a combination of this information obtained from the search history could be used to trace the person conducting searches. Similarly, it was shown that from the credit card dataset, 90% of users could be identified by knowing just 4 of their transactions (shop and time). Thus, even though in both the datasets, direct identifiers were absent, a combination of all the information that can be directly availed or inferred from these datasets by themselves or when merged thoughtfully with some external datasets could easily be used to trace back the person.

(b) Search history and financial transaction data is a trove of sensitive information. For example, a person's search history can be used to narrow down not just the identity of the person, but also provides a good insight into the person's behavior, habits and thoughts. Making this data public not only invites malicious activity like identity theft, but also personal information about a person's private life would be on display for the public. Similarly, financial transaction information is also considered to be very sensitive. For example, it could be used to judge a person's financial status which can lead to say, targeted financial fraud or even robbery.

Problem 2: Describing ethical thinking

i. `Were sociologists, not technologists, so a lot of this is new to us and `Sociologists generally want to know as much as possible about research subjects."'

‘We considered possible ethical breaches of the study from the perspective of sociologists. Following the ethical framework of ‘beneficence’, we tried our best to aggregate and anonymize data to protect the privacy of the research subjects. Additionally, we also put forth a protection strategy for the dataset to minimize the risk to the students. We have also taken ‘respect for public interest into consideration’. As sociologists, we try to find out as much as possible about the research subjects to provide research of the highest quality. We could have also refrained from publishing the study dataset. But keeping in mind the ‘respect for public interest’, we felt obliged to share our dataset in line with the concept of ‘transparence-based accountability’. However, we apologize for not seeking ethical guidelines from experts in computer science ethics.’

ii. *`What might hackers want to do with this information, assuming they could crack the data and `see' these peoples Facebook info? Couldn’t they do this just as easily via Facebook itself? Our dataset contains almost no information that isn’t on Facebook. (Privacy filters obviously aren’t much of an obstacle to those who want to get around them.)"'*

‘Public data on Facebook was already prone to exposure risk. Even if our data is exposed, we feel that the information is not very sensitive that it would lead to any serious repercussions for the participants. Therefore, from a consequentialist point of view, usage of Facebook data does not significantly increase the risk for any participant. Whereas the study results can benefit society as a whole. We understand that this is not ideal considering the ‘Justice’ part of the ethical framework. That is, the students are exposed to the risk, while the benefits are distributed unfairly to the whole population. There are also concerns raised that the study violates ‘respect for people’ because we have not explicitly taken consent from the participating students. Our defense for this point is that when a person provides approval for certain information to be public on Facebook, they have implicitly given permission for the world to access this information.’

iii. *“We have not accessed any information not otherwise available on Facebook. We have not interviewed anyone, nor asked them for any information, nor made information about them public (unless, as you all point out, someone goes to the extreme effort of cracking our dataset, which we hope it will be hard to do).*"

‘We have adhered to ‘respect for law and public interest’. We were careful to seek and gain approval from both the IRB board of the university and from Facebook before collecting the data. We have tried our best to minimize the risk to the students. We have only accessed public information from Facebook, obtained no information from the students themselves, nor publicized our dataset. Therefore, we are convinced that the data collection is ethical even without consent from individual students. Moreover, we have also anonymized our data. We believe this is sufficient to control the risk to the participants. However, in the unlikely event that the dataset can be re-identified, we would withdraw the dataset to ensure that the participant privacy is not risked.’

Problem 3: Ethics of encore

The Encore study proposes a novel mechanism to collect large-scale information about censoring: the time, geographic location and the information being censored on the internet. The data could be used to answer a broad range of research questions like the technical mechanism of censorship implementation, the speed of removal of objectionable online content, how is information modified or removed, the motivation for censorship, impact on freedom of speech etc. The methodology presented was to embed a few lines of codes in popular websites that would automatically make users visiting them to ping possibly censored websites from their IP addresses without any explicit notice to the users.

The suggested novel approach could answer research questions of significant importance to the world as mentioned before. An alternate method to monitor censoring is to recruit volunteers to provide consent to use their computers to perform automated pinging. But the proposed solution is more scalable and can provide more granular information. Moreover, it also solves the issue of censors blocking IP address because of the random nature of people who would visit the web pages. Despite the numerous gains in using this mechanism, it comes with serious ethical concerns concerning its usage. The authors proceed to analyze the study ethics by employing the 'respect for persons', 'beneficence', 'justice', and 'respect for law and public interest' framework from Menlo and Belmont reports.

The authors first make a case for why the Encore study should be considered under the preview of a human experiment. Quoting from the paper, a human experiment's definition under common IRB guidelines is "(1) Data through intervention or interaction with the individual, or (2) Identifiable private information". The study design collects only the IP address and no other user information. It is ambiguous whether IP addresses can qualify as identifiable private information. The study can be modified to generalize the collected IP data. Therefore, following strict guidelines, the study would not qualify as a 'human experiment'. However, the guideline omits the case when the user's information (IP address) can be collected by other parties, like the censor or the websites pinged in this case. Therefore, the authors argue that the study should be held to ethical standards fit for a human experiment.

The study violates the 'respect for persons' clause of the ethical framework. There was no informed consent from the parties involved, partly because of the scale of the study and partly because it would have been impractical to educate all the users about the technical details of the study. The researchers involved in the study further state that the lack of informed consent could help the users "plausibly deny" participating in the study in case of legal complications. But Narayanan and Zevenbergen (2015) note that this would be of little value to the users in oppressive regimes who do not hold fair trials.

The report also delves extensively into evaluating the 'beneficence' and 'justice' criterion of the framework from a consequentialist point of view. That is, the authors try to weight the potential harms and benefits of the study from a purely outcome-oriented perspective. There is no direct harm to the persons involved in the study. The risks to the users participating in the research are very hard to quantify because of the scale of the study and the diversity of the participants. The risk experienced by each user is influenced by the user's social, philosophical views, history,

their location's political climate etc. For example, visiting an extremist religious group from a repressive regime could invite legal actions against the ignorant user. While in more liberal countries, the visit could go unnoticed with no associated repercussions. There are numerous benefits to the study as stated earlier. However, the study complies to the 'justice' section. Study participants are a random sample of the population that might benefit from censorship research. So, the benefits and risks of the study are distributed uniformly.

As for the 'respect for law and public interest', the study took approval from IRB boards of both the involved universities. The study also provides an opt-out option for users, making it comply to the legal laws in the U.S. But the compliance to the laws of the various geographies the study will be conducted in has not been evaluated because of the complexity. Thus, there is a chance that it violates some local laws. Moreover, the team behind the Encore study have not publicly accepted the 'responsibility for their actions and consequences' nor 'have the necessary mitigation strategies in place' according to the case study.

The authors have, unlike many other researchers, evaluated the ethical quality of the study. They have gone over and beyond obtaining just the IRB approval to discuss the case with various experts in the domain. But my personal view is that the study does not pass the ethical bar. The researchers should take the consent of the people involved in the study. Narayanan and Zevenbergen (2015) rightly point out that the users neither consented to participate in the study nor were informed about the researcher's intent. Although obtaining consent improves the ethical quality of the study, it is not enough to shield the participants from risk. The risk is merely transferred from the researchers to the participants. My verdict is that the uncertainty in the risks faced by users outweighs all the benefits that can be reaped from the study. A suggestion is that the researchers could administer the study only in certain safe geographies, even though this undermines the spirit of the research to perform massive data collection.