MACS30000 Assignment 8

Dr. Richard Evans
Submitted by: Junho Choi
Due December 3, 2018 (11:30 AM)

Problem 1

(a) In Zimmer (2010) and de Montjoye et al. (2015), the researchers highlight that certain datasets used in research (in combination, perhaps, with their codebooks and other public information) can be used to reveal individual identities even after the supposed process of anonymization. Both cases use the idea that only just a few points of data are required to de-anonymize an observation, as such will reveal unique characteristics about the individuals behind the data.

For instance, making note of the "Taste, Ties, and Time" (or T3) research by Lewis et al. (2008), Zimmer refers to a comment by Stutzman (2008) that social media network can be thought of as a "fingerprint" of sort for an individual having the said network (Zimmer 2010, p. 316). The author elaborates that a simple investigation into the dataset's codebook (along with some other information) was enough to narrow down which universities or colleges the dataset was based upon (Zimmer 2010, p. 316). Similarly, de Montjoye et al. (2015) uses the concept of unicity, or the "risk of reidentification knowing [a certain number of] pieces of outside information about a user," to exemplify that it takes only around four spatiotemporal data points on credit card usage to identify an individual with sufficient accuracy (p. 537).

(b) As the researchers of T3 have argued, data regarding the study were already "public" in the sense that anyone with sufficient access to Facebook could view them (Lewis et al. 2008). At the same time, however, this would not mean that an outside individual has the capacity (in time and effort, perhaps) to view and analyze all of such individual sub-networks. By providing this data, the T3 researchers have drastically lowered the cost of analyzing individual behaviors on Facebook. While seemingly innocuous, some may inherently oppose to the fact that they are being analyzed or monitored, which would be violating the principle of "respect for persons" (Salganik 2018). For instance, take the case of closeted individuals who do not (and perhaps do not want to) reveal their sexual orientation. A study by McConnell et al. (2018) reveals characteristics of these individuals on social networks (such as more inactivity). If combining this information with data such as that used in T3, one may be able to find out not only one's identity, but also sexual orientation without one's consent.

In the case of credit card information that is dealt with in de Montjoye et al. (2015), it is uncertain whether the identification of one's *spending* could bring forth some harm. However, the very fact that spatiotemporal information is accompanied

in such a dataset could potentially be dangerous to one's privacy. If an individual is identified, but only one's spendings are known (without knowing what they spent on), then a researcher may only be able to impute information such as income or propensity to consume. However, combined with geographical and time-series data, even more serious breach of privacy could be of concern as these individuals' housing or work environment as well as daily movement information can be traced and made public.

Problem 2

When rewriting the comments (or passages) by Kaufman, I will try to make amends to Kaufman's tone in which it seems as if he does not realize what potential transgressions he and his team might have made.

First Comment

"As sociologists, we do want to do know as much as possible about our research subjects. However, we perhaps should have been more careful in that our subjects were not given the chance to respond to informed consent. In this regard, we agree that we have breached the 'respect for persons' criterion delineated in the Belmont and Menlo reports."

Second Comment

"We understand that, while all of the information was publicly disclosed through Facebook, we severely lowered the cost for potential 'hackers' to use such information in a malicious manner. Furthermore, while we maintain the position that the principles such as beneficence, justice, and respect for law and public interest have not been violated, we do recognize that respect for persons (in terms of informed consent) could have been violated. In this sense, even if there are no potentially-malicious users of our data, we acknowledge that our approach may have been problematic."

Third Comment

"In addition to our previous comment, we acknowledge that we may have undermined the possibility that our dataset could be de-anonymized. The fact that we did not interview anyone or asked of their information is not something that we should proudly state, but rather be ashamed of as it may potentially violate, once again, respect for persons and may also strengthen only the consequentialist point-of-view."

Problem 3[1]

(a) The criticism made by Narayanan and Zevenbergen towards the Encore study (Burnett and Feamster 2015) is built on the argument that the study altered the behaviors of study participants' computers, thereby landing an impact on the participants themselves (N&Z 2015, p. 3). While there are multiple dimensions to the authors' criticisms, one of the key themes is beneficence; that is, the authors make analyses of the potential benefits and harms from the Encore study, and argue that the harms may have outweighed benefits (N&Z 2015, pp. 14-20). This can be understood as a consequentialist approach, in which a certain study can be justified when the net benefit is positive (i.e. "ends" are good) (Salganik 2018).[2]

Regarding the potential benefits, the authors borrow an argument from Wright, de Souza, and Brown (2015), that the study could have helped in the attempt to remove potential secrecy surrounding censorship process and devise better ways to circumvent censorship (N&Z 2015, p. 15). Yet this benefit is based on the very idea that censorship is "bad," which is noted as *not* to be agreed by all (N&Z 2015, p. 16). On the other hand, Narayanan and Zevenbergen posit that the harm from this study could have exceeded the "minimal risk" of the everyday use of the Internet (N&Z 2015, p. 17). The authors make this claim in three steps: first, that the users' anticipation of using the Internet does not encompass the potential impact from Encore; second, "probability and magnitude of harm" may differ by Encore accessing different websites; and third, potential harm from the study may impact not only individuals, but also overflow onto others who are not associated with it (N&Z 2015, pp. 16-19).

The authors are not conclusive on whether the harm or benefit from the Encore study outweighed one another. Rather, they return to the basic principle of beneficence, which demands "no harm" and "maximization of potential benefits and minimization of potential harms" (N&Z 2015, p. 20; Salganik 2018). In doing so, they emphasize on how one should consider defining and minimizing the "vulnerabilities" or "harm" when conducting studies akin to Encore (N&Z 2015, p. 20).

(b) I believe that, due to the subtlety in the design of the study that Burnett and Feamster (2015) conducted, there are no serious violations of the four basic principles that Salganik (2018) makes notes of. However, as argued in Narayanan and Zevenbergen (2015), this issue can potentially become trickier as the tests of censorship are made for not just one website in country but practically all over the world. For example, it is mentioned that the so-called "pings" were made not only in the United States

[1] For the sake of convenience, I will abbreviate the authors' last names Narayanan and Zevenbergen as "N&Z" when referencing them in parentheses.

[2] Of course, this is not to say that Narayanan and Zevenbergen would have justified the Encore study if benefits outweighed harms; in fact, the authors clearly mention that there are other ethical concerns one must be mindful of (N&Z 2015, pp. 20-23)

or other democratic nations but also in countries like China, where censorship is much stronger and violation of it can lead to an outcome much serious than in the United States (N&Z 2015, p. 3). Therefore, the Encore study is very likely to violate, at least marginally, the principles of beneficence (in which harm may overly outweigh benefits) and justice (in which, at the cost of some people's harm, others may benefit).

References

Burnett, Sam, and Nick Feamster. 2015. "Encore: Lightweight Measurement of Web Censorship with Cross-Origin Requests."

Lewis, Kevin, Jason Kaufman, Marco Gonzalez, Andreas Wimmer, and Nicholas Christakis. 2008. "Tastes, Ties, and Time: A New Social Network Dataset Using Facebook.com." *Social Networks* 30 (4), 330-342.

McConnell, Elizabeth, Bálint Néray, Bernie Hogan, Aaron Korpak, Antonia Clifford, and Michelle Birkett. 2018. "'Everybody Puts Their Whole Life on Facebook': Identity Management and the Online Social Networks of LGBTQ Youth." *International Journal of Environmental Research and Public Health* 18 (5).

de Montjoye, Yves-Alexandre, Laura Radaelli, Vivek Kumar Singh, and Alex "Sandy" Pentland. 2015. "Unique in the Shopping Mall: On the Reidentifiability of Credit Card Metadata." *Science* 347 (6221), 536-539.

Narayanan, Arvind, and Bendert Zevenbergen. 2015. "No Encore for Encore? Ethical Questions for Web-based Censorship Measurement." *Technology Science*.

Salganik, Matthew J. 2018. *Bit by Bit: Social Research in the Digital Age*. Princeton: Princeton University Press.

Stutzman, Fred. 2008. "Facebook Datasets and Private Chrome." *Unit Structures*.

Zimmer, Michael. 2010. "But the Data is Already Public: On the Ethics of Research in Facebook." *Ethics and Information Technology*, 12 (4), 313-325.