Assignment #8

Sanja Miklin
12/03/2018

1. Identification risk in anonymized data

(a) Pick two of the examples in Table 1 and describe in one or two paragraphs how the re-identification attack in both cases has a similar structure.

Both credit card transaction data, as well as search log data could be re-identified in an attack, as the shopping and searching behaviors are unique to specific individuals. The overall structure of an attack is similar in both cases, as specific data-points which are not uniquely identifying of anyone can point to specific individuals when aggregated.

With credit card transaction data, knowing someone's whereabouts on only a few occasions (through personal knowledge, social media check-ins etc.) might be enough to uniquely identify them (serve as their 'unique fingerprint') in a dataset. For example, I might be the only person who used my card at Nella, Whole Foods and Hyde Park Produce mid-day Saturday.

A similar process could be employed on internet-search log data, as a collections of interests could serve as a unique fingerprint (Most people who know me could easily guess I spent much time googling for information on dogs, circus, suicide, R, sustainability and Croatia, and might be the only person in the world doing so). Additionally, especially as individuals might google quite identifiable information themselves (such as local restaurants, information regarding their university etc), their information could emerge from the data as well.

(b) In one or two paragraphs, describe how the data could reveal sensitive information about the people in the dataset for each of your two examples in part (a)

Credit card transaction data, for example, might reveal behavior and interests that the individual wishes to hide: maybe they have purchased alcohol at a strip-club, or a room in a hotel when they were supposed to be 'out of town on business', maybe they charge their medication to their credit card, and a regular monthly trip to a pharmacy will indicate that they are using some sort of a prescription medication.
Similarly, search log data could reveal searches for different kind of stigmatized or even illegal material, reveal political/religious/sexual attitudes and affiliations, and will at least be somewhat embarrassing for most people—these days we ask google some of the weirdest questions we would not ask of anyone else.

2. Describing ethical thinking

In approaching this assignment, I was unclear as to the actual objective, that is if we were meant to more powerfully defend the T3 study within the ethics framework, or to try to word, as shakily as might be necessary, Kauffman's apparent position in a different language. I struggled with the first, so I mainly try the second.

> "Upon the public announcement of this initial discovery, and general criticism of the research teams attempts to protect the privacy of the subjects, Jason Kaufman, the principle investigator of the T3 research project, was quick to react, noting that, perhaps in justification for the amount of details released in the dataset, **'We're 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."'** [Zimmer (2010) citing Kauffman (Sep. 30, 2008b)]

The above comment centers on lack of sociologists' awareness of technology (which I think is difficult to communicate within an ethical framework) as well as the value of detailed data for sociological research (the ends of research, within the consequentialist perspective), which could've been worded as follows:

> REWRITE:
> "Detailed data is central to high-quality sociological research that aims to improve our understanding of how human communities and societies. In assessing our project, we weighed the minimal potential risks to individuals with the extensive benefit this kind of a longitudinal and near-complete dataset would offer to our understanding of social networks. While we are not experts at technology, we followed the Common Rule and ethical practices as specifically established in sociology."

> "[Kauffman] then attempts to diffuse some of the implicit privacy concerns with the following comment:
>
> **'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.)"'** [Zimmer (2010) citing Kauffman (Sep. 30, 2008b)]
>
> **"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)."** [Kauffman (Sep. 30, 2008c)]

In the above comments, Kauffman wishes to establish that the dataset basically does not infringe on the privacy of persons any more than the Facebook data already publicly available. This might be rewritten as follows:

> REWRITE:
> "Individual Facebook users choose the privacy settings of their profiles. Their choice to make any information publicly visible, is in itself a consent for that information to be public. Researchers, hackers and other users alike will all have access to that data. Furthermore, our data-set does not, in itself, contain sensitive information (beyond what individual students have shown to display). In contrast, hackers could access this information irrespective of the existence of our data. Therefore, our data-set does not violate the privacy of the students and follows the principle of Respect for Persons. Our use of data also followed relevant laws and Facebook's terms of Service.

3. Ethics of Encore (3 points)

Read the Encore web censorship study by Burnett and Feamster (2015) and the reply and critique of that study by Narayanan and Zevenbergen (2015).

(a) Write a one-half-to-one-page summary of Narayanan and Zevenbergen's assessment of the Burnett and Feamster (2015) Encore study. Make reference to the consequentialist framework and to the principle of beneficence.

In their assessment of the Encore study, Narayanan and Zevenbergen (N&Z) approach their assessment through four loose frameworks as they seek to determine if 1) If the study involved human subjects (Not under current definition), 2) If the benefits outweigh the risks (Yes), 3) If the researchers should have been more transparent about the work and asked for consent (Yes), and 4) If any US (No) or International (Unclear) laws were violated.

In terms of ethics of the work, the authors argue that computer science research is quite different from usual research, in that it seeks to maximize and not minimize the number of 'participants', in a way that makes identifying stakeholders unfeasible. Additionally, even though the research collects the users IP address, the authors note that not only can the IP addresses be generalized, but that the data collected is not about the individuals but the censorship systems—if it were possibly could've easily been conducted without human participation. As such, the research is not necessarily 'human subject' research. However, looking at the research from the consequentialist perspective, a study like Encore can nevertheless have consequences to humans and can be designated 'human-harming' research, which is why other aspects of the research have to be examined.

For example, looking at the study with the principle of beneficence in mind, authors argue that censorship overall is harmful to humans as it violates human rights. Studies of censorship, therefore, are highly beneficial to society, though the authors do note that this is not a universal perspective. If all the participants in the study. Additionally, authors acknowledge that Encore study does not put individuals at more risk than regular browsing, which would make the risks of the study minimal. Still, they do point

out not only that putting internet-users at risk in general is not an ethical practice (regardless of it being commonly done), but that there are risks that extend beyond individuals—if studies like Encore become more common they might change how the internet and the censors operate.

In terms of informed consent, transparency, and accountability, authors show that although Encore study did not violate any US laws, it is unclear—and possibly impossible to determine—if it would violate any international laws, or if calls to particular websites could be seen as illegal in certain countries. This is additionally important because illegal activity might put participating users themselves at risk. This is why the authors argue that beyond harm mitigating strategies already in place within the study (e.g. limiting the URLS within the Encore script only to common websites), the ethical thing to do would be to ask for consent. While the Encore authors have argued that getting consent from individual users would not be feasible, and would make the study equivalent to some projects already in place, N&Z argue that at the least, transparency could've been improved by providing users of volunteer websites a notice, and explaining the risk and benefits in the FAQ.

Ultimately, the authors show that computer science research, in method and in scale, does not fit easily within the frameworks of ethical research we've developed, but that doesn't mean we cannot or should not approach computer science research with ethics in mind.

(b) In one or two paragraphs, write your assessment of the ethical quality of the Burnett and Feamster (2015) Encore study.

In thinking about the Encore study, I found Salagnink's (2017, Ch. 6) notion of ethics as continuous, rather than discrete really useful. I would not call the Encore study absolutely unethical, especially as it is conducted at a time when we are just thinking through and doing our best to establish the ethics of large scale computer research.

That said, I also believe that it is important to be a bit conservative in research, especially when the waters are unclear and there is potential harm to individuals. Specifically, as this study explicitly focuses on censorship and might be putting individual users at risk (however minimal), this is enough to convince me that consent of some sort should have been used—after all, these days we're all consenting to many a website's use of Cookies. While the authors' reasoning for why consent would be impractical, and would likely not decrease risk for participants are valid, that doesn't mean doing away with consent was the only option—changing the design of the study would've been a different one.

Another thing that would be important to me is the assessment of actual benefits of Encore, especially in thinking about the ethics of further deployment of Encore after the small pilot. If Encore does not really give a better sense of censorship than already existing projects or a similar project that would not put unsuspecting users at risk (e.g. some sort of a VPN?) then it would be impossible to argue that it is ethical.