

Short Paper #1: Ethics of "Taste, Ties, and Time"

SHEN HAN

October 9, 2016

In 2008, some descriptive findings of the “tastes, ties, and time” project¹, which based on a social network dataset introduced by the project team that characterizing a cohort of “anonymous” college students by their gender, race/ethnicity, and socioeconomic status, were published. Nevertheless, debates have been widely aroused regarding its ethical propriety and intended usage. In my point of view, this dataset is not in compliance with most of the ethnic principles of research and should not be used for further research without necessary corrections and protection.

1 Comments on Research Ethics of the Experiment

1.1 Respect for Persons

The principle of Respect for Persons from the Belmont Report² argued that (1) individuals should be treated as autonomous and (2) individuals with diminished autonomy should be entitled to additional protections. In practice, this principle was also interpreted as “informed consent”, which means research participants should be presented with relevant information in a comprehensible format and then should voluntarily agree to participate³. Unfortunately, the “tastes, ties, and time” project went in the opposite direction of this principle at the beginning.

The research violated its participants’ autonomy by carrying the project without their consent or even awareness. In the case that may take advantage of users’ sensitive information, it is far from enough that just lean on the “terms of service”, which almost nobody will actually have a look.

¹Lewis, K., Kaufman, J., Gonzalez, M., Wimmer, A., & Christakis, N. (2008). Tastes, ties, and time: A new social network dataset using Facebook.com. *Social networks*, 30(4), 330-342.

²Belmont Report. 1979. The Belmont Report: Ethical Principles and Guidelines for the Protection of Human Subjects of Research. US Department of Health, Education, and Welfare.

³Salganik, Matthew J. *Bit by Bit: Social Research in the Digital Age*, Princeton University Press, Open review edition.

Although it is true that the project’s violation of Respect for Persons principle does not necessarily mark the research down as ethically impermissible, it was blamed for its lack of fundamental protection and remediation for people’s autonomy. It not only failed to let its participants know their sensitive information was used elsewhere than their original intention, but also left them in a vulnerable position that anyone else may access their data easily. In addition to taking further steps to ensure the privacy of students in the dataset, the research team should also get some consent form from participants before or at least after the study.

1.2 Beneficence

Unlike the Respect for Persons principle, the principle of Beneficence is an obligation that researchers must respect, according to the Belmont Report, and it involves two parts: (1) do not harm and (2) maximize possible benefits and minimize possible harms. In practice, researchers are required to undertake a risk/benefit analysis and a comprehensive decision based on the analysis. Although we know about neither the project team conducted a risk/benefit analysis in advance or not, nor the details and the conclusion of their analysis, here we can simply have a similar analysis and see whether our decision go in the same way.

The principal contradiction between its benefit and risk is about the efficiency of exploring new possibilities for social network research at the expense of potentially harming the autonomy and privacy of the participants. Due to the problems in the research design, as I will discuss in the last section, the academic advancement of the project seems limited and totally overwhelmed by its drawbacks. While the information is potentially identifiable and sensitive, as some researchers proved the project data actually are, easily by some de-anonymization⁴, the decision-making process of the research should really be assessed with prudence and care.

Hence, in consideration of the unfavorable balance between the risks and benefits of the project, it is better to suspend the experiment before further discussions are carried out and safeguards are placed. From my perspective, it maybe the case that the researchers did not know the “bugs” in their “anonymization” process, or they thought the consequences of the leakage of students’ sensitive information would not carry a lot of weight.

1.3 Justice

The principle of Justice from the The Belmont Report argued that the distribution of the burdens and benefits of research should be equal. However, it does not mean absolute equality in practice. It means disadvantaged and vulnerable

⁴Zimmer, M. (2008a). More on the “Anonymity” of the Facebook dataset—It’s Harvard College. michaelzimmer.org Retrieved October 9, 2016, from <http://michaelzimmer.org/2008/10/03/more-on-the-anonymity-of-the-facebook-dataset-its-harvard-college/>.

people should be protected in the research, instead of intentionally preyed. By this mean, the “tastes, ties, and time” project did not do a good job as well.

Let alone the credibility and validity of their research outcome, the distribution of the potential risks and benefits of the project failed to be reasonable and balanced. A cohort of students take the whole burden of the research, while the rest of the society may enjoy its potential benefit is not the way justice works. Although those students may not belong to disadvantaged population, they still needed to be protected with special safeguards against informational risk according to their vulnerable position.

1.4 Respect for Law and Public Interest

In online research, the principle of Respect for Law and Public Interest from Menlo Report⁵ has its unique meaning of guidance for social science research, which “encourages researchers to take a wider view and include law in their considerations.” The “tastes, ties, and time” project seems adhere to this principle, but at the expense of breaking others.

In consideration of its approval from Harvard IRB and Facebook’s Terms of Service, the project was technically complied with relevant laws and terms. But it is still questionable that whether the Terms of Service of Facebook are complied with other laws and regulations or not. Regarding to its transparency-based accountability, it indeed made its research data available to other researchers for a short time, but later they made it offline as they “take further steps to ensure the privacy of students in the dataset.”⁶ Till now, eight years have passed, and the dataset is still offline, which aroused more doubts about their ability and willingness to protecting the privacy of research participants.

2 My Attitude towards this Dataset

From my perspective, I would not use this data for my own research due to its technical and ethical problems.

Technically, I am concerned about the external validity and research degrees of freedom problems regarding to this dataset. The data are too concentrated and its gathering process may be vulnerable to other problems such as selection bias and omitted variables as well. Also, some of the data were generated by “using names or official college photographs to identify the students’ race and ethnicity”, which seems to have too much research degrees of freedom.

Ethically, I am more in favor of the framework of deontology than consequentialism in this case, which means the correctness of means is more important than its consequences. As I analyzed above, the effort of the research team to

⁵Dittrich, David, Erin Kenneally, et al. 2011. “The Menlo Report: Ethical Principles Guiding Information and Communication Technology Research.” US Department of Homeland Security.

⁶Lewis, K., et al. Tastes, Ties, and Time Dataverse. Retrieved October 9, 2016, from <https://dataverse.harvard.edu/dataverse/t3>

adhere with most of the ethnical principles is far from satisfactory. Ethnic principles are the floor instead of the ceiling. The justification of this dataset will emit a dangerous signal that regardless of the principles, researchers can carry out all kinds of social experiments without having a comprehensive discussion in advance, as long as the outcome is favorable.

Therefore, as the paper said in the abstract, it will “provide a starting point for future analyses”, and I have to admit that their novel data source is undoubtedly promising, though its generating process may break several ethical principles. However, if we just focus on the end instead of the justice of the mean that leads to the end, I am afraid that it may not be the starting point that we really want in this digital age.