

## Ethics of "Taste, Ties, and Time"

In October 2008, the journal *Social Networks* published "Taste, ties and time: A new social network dataset using Facebook.com". This article immediately sparked debate over whether or not the newly introduced dataset's existence and use is ethical. When facing ethical dilemmas of data use, researchers can be guided by the four principles laid out in the Belmont Report and the Menlo Report. These four principles are respect for persons, beneficence, justice, and respect for law and public interest (Salganik).

The first principle, respect for persons, is a reminder that human beings are autonomous and that their free will and wishes need to be respected (Salganik). In the case of "Taste, ties and time", the Facebook data of 1,640 freshmen students from a "diverse, private college in the Northeast U.S." was collected (Lewis et al.). Principal investigator Jason Kaufman and his team downloaded the data manually after receiving consent from Harvard University and Facebook (Lewis et al.). The issue of consent becomes murky when dealing with social network data. No student consent was acquired and students were not even informed of the researchers collecting, analyzing, and, ultimately, publishing their information as part of an open dataset (Zimmer, "Anonymity"). Researchers assumed that since the data had been published by the subject on Facebook, it was now free to use for purposes outside of the social network.

What must be considered is the intent of the student in making a Facebook profile. The intent, one can safely assume, is to network with others. It is a stretch to imagine that posting a photo on Facebook is an invitation to be studied, especially when viewed in the light of contextual integrity. Within the concept of contextual integrity exists two norms: appropriateness and distribution (Zimmer, "Privacy"). The norms of appropriateness describe the kind of information "within a given context" that is "allowable, expected, or even demanded" (Zimmer, "Privacy"). In the case of Facebook, it is allowable and strongly expected for individuals to share their names and at least one photo. Information such as hometown, school, and other interests are also expected and freely shared within the Facebook community. Given that an individual is sharing this information with the expectation that it will be viewed by friends that the individual must approve of, many users feel free to share very personal information as they are in the virtual company of friends.

The norms of distribution would then generally constrain friends who have personal details about an individual from sharing that information freely to a third party. However, these norms were violated as individuals had no way to opt out of the study and no knowledge that information they were sharing in the relaxed and informal setting of Facebook was downloaded for public consumption. Researchers did not collect data on private Facebook accounts (Lewis et al.), meaning that the only way to (unknowingly) opt out of the study was for the student to make his/her Facebook profile as private as possible, a decision that a student generally makes for reasons other than to be precluded from becoming an unwitting research subject. However,

84.6% of the private profiles still had network data available and this data was collected by the researchers (Lewis et al.). These students were given no autonomy to decide whether or not to participate in the study and therefore their free will was not respected.

Beneficence, meaning to maximize possible benefits and minimize possible harm (Salagnik), is the second principle. Kaufman and his team recognized the harm that could be done to individuals whose identities were matched with the dataset and did take measures to avoid this. Names, phone numbers, emails, and addresses were encoded or removed from the dataset and the labels for the students' favorite books, movies, and music were temporarily removed with the intention of releasing them at a later date (Lewis et al.). These safeguards proved to be insufficient as the identity of the college was quickly uncovered (Zimmer, "Anonymity"). Facing scrutiny, Kaufman defended his opinion that the benefits of the study outweighed the risk by questioning what harm could be done by hackers who could piece together the identities of the students. However, as Michael Zimmer points out, "privacy is about dignity as much as about informational harm by some evil agent" (Zimmer, "Anonymity"). Once this data was made public, the 1,640 students whose data was collected learned that the information they provided for personal reasons on Facebook was taken public and their identities and opinions were publicly analyzed. It can be seen how this study might negatively impact the feelings of security and privacy that these students once enjoyed.

The principle of justice demands that vulnerable people be protected and that risks and benefits be evenly distributed among all people (Salagnik). The most easily identifiable individuals from the dataset are those from poorly represented states, nations, and ethnic backgrounds. For example, in the dataset there exists only one student from Albania and two from Armenia (Zimmer, "But the data is already public"). These students can be easily identified for anyone familiar with a minority student or with other knowledge about the identities of Harvard's class of 2009. Therefore, the risk of identification is much greater for students with unique backgrounds. A unique background very often coincides with a vulnerable population and it is exactly that vulnerable population that should be protected by researchers.

Finally, researchers are expected to demonstrate respect for law and public interest. Being transparent about "goals, methods, and results at all stages" (Salagnik) is a key portion of this principle. The authors of "Taste, ties, and time" published their study and the data openly once the study had concluded, however they may have been subject to the Common Rule, the federal policy meant to protect subjects by making sure that they are aware of the study and the risks involved. If so, Kaufman and his team violated the Common Rule by not ensuring the subjects were aware of study. Facebook's Data Use Policy states that posted information may be used for academic research, however one can question the validity of assuming that a subject is aware of a study when only provisionally notified by a terms of service agreement.

Even though the "Taste, ties, and time" dataset does not strictly adhere to the four principles outlined, I would feel comfortable using the dataset once it was made available. While I do not agree with the way in which the data was collected, I also believe that the dataset can

be used to answer far-ranging questions with no further risk to the subjects. In my work with the dataset, I would be careful not to point out identifiable characteristics of individuals (such as the Albanian student mentioned above) and I would not seek further identifying information about the students. I am concerned that my use of the dataset would justify its collection; however, I hope that, as a researcher, I can collect data ethically moving forward, while not sacrificing the insights to be gained by using past datasets, whose methods we have since improved upon.

## Works Cited

Lewis, Kevin, et al. "Tastes, ties, and time: A new social network dataset using Facebook.com." *Social Networks*, vol. 30, no. 4, July 2008. <http://www.sciencedirect.com.proxy.uchicago.edu/science/article/pii/S0378873308000385>.

Salganik, Matthew J. "Bit By Bit: Social Research in the Digital Age." <http://www.bitbybitbook.com/en/ethics/principles/>.

Zimmer, Michael. "'But the data is already public': on the ethics of research in Facebook." *Ethics and Information Technology*, vol. 12, no. 4, June 2010. <http://link.springer.com.proxy.uchicago.edu/article/10.1007%2Fs10676-010-9227-5>.

Zimmer, Michael. "On the 'Anonymity' of the Facebook Dataset (Updated)." <http://www.michaelzimmer.org/2008/09/30/on-the-anonymity-of-the-facebook-dataset/>.

Zimmer, Michael. "Privacy as Contextual Integrity (Part 2): Norms of Information Flow." <http://www.michaelzimmer.org/2005/04/13/privacy-as-contextual-integrity-part-2-norms-of-information-flow/>.