Ethics Inf Technol (2010) 12:313–325
DOI 10.1007/s10676-010-9227-5

# "But the data is already public": on the ethics of research in Facebook

**Michael Zimmer**

Published online: 4 June 2010


**Abstract** In 2008, a group of researchers publicly released profile data collected from the Facebook accounts of an entire cohort of college students from a US university. While good-faith attempts were made to hide the identity of the institution and protect the privacy of the data subjects, the source of the data was quickly identified, placing the privacy of the students at risk. Using this incident as a case study, this paper articulates a set of ethical concerns that must be addressed before embarking on future research in social networking sites, including the nature of consent, properly identifying and respecting expectations of privacy on social network sites, strategies for data anonymization prior to public release, and the relative expertise of institutional review boards when confronted with research projects based on data gleaned from social media.

**Keywords** Research ethics · Social networks · Facebook · Privacy · Anonymity

## Introduction

In September 2008, a group of researchers publicly released data collected from the Facebook accounts of an entire cohort of college students. Titled "Tastes, Ties, and Time" (T3), the announcement accompanying the release noted the uniqueness of the data:

> The dataset comprises machine-readable files of virtually all the information posted on approximately 1,700 [Facebook] profiles by an entire cohort of students at an anonymous, northeastern American university. Profiles were sampled at 1-year intervals, beginning in 2006. This first wave covers first-year profiles, and three additional waves of data will be added over time, one for each year of the cohort's college career.
>
> Though friendships outside the cohort are not part of the data, this snapshot of an entire class over its 4 years in college, including supplementary information about where students lived on campus, makes it possible to pose diverse questions about the relationships between social networks, online and offline. (N.A. 2008)

Recognizing the privacy concerns inherent with the collection and release of social networking data, the T3 research team took various steps in an attempt to protect the identity of the subjects, including the removal of student names and identification numbers from the dataset, a delay in the release of the cultural interests of the subjects, and requiring other researchers to agree to a "terms and conditions for use," prohibiting various uses of the data that might compromise student privacy, and undergoing review by their institutional review board (Lewis 2008, pp. 28–29).

Despite these steps, and claims by the T3 researchers that "all identifying information was deleted or encoded" (Lewis 2008, p. 30), the identity of the source of the dataset was quickly discovered. Using only the publicly available codebook for the dataset and other public comments made about the research project, the identity of the "anonymous, northeastern American university" from which the data

M. Zimmer (✉)
School of Information Studies, University of Wisconsin-Milwaukee, 656 Bolton Hall, 3210 N. Maryland Ave, Milwaukee, WI 53211, USA
e-mail: zimmerm@uwm.edu

was drawn was quickly narrowed down to 13 possible universities (Zimmer 2008b), and then surmised to be Harvard College (Zimmer 2008a). Reminiscent of the ease at which AOL users were re-identified when the search engine thought the release of individuals' search history data was sufficiently anonymized (see Barbaro and Zeller Jr 2006), this re-identification of the source institution of the T3 dataset reveals the fragility of the presumed privacy of the subjects under study.[1]

Using the T3 data release and its aftermath as a case study, this paper will reveal numerous conceptual gaps in the researchers' understanding of the privacy risks related to their project, and will articulate a set of ethical concerns that must be addressed before embarking on future research similarly utilizing social network data. These include challenges to the traditional nature of consent, properly identifying and respecting expectations of privacy on social network sites, developing sufficient strategies for data anonymization prior to the public release of personal data, and the relative expertise of institutional review boards when confronted with research projects based on data gleaned from social media.

## The "Tastes, Ties, and Time" project

Research in social networks has spanned decades, from Georg Simmel's foundational work in sociology (Simmel and Wolff 1964), to Barry Wellman's analyses of social networks in the emerging networked society of the late twentieth century (Wellman and Berkowitz 1988), to the deep ethnographies of contemporary online social networks by boyd (2008b). Indeed, the explosive popularity of online social networking sites such as MySpace, Twitter, and Facebook has attracted attention from a variety of researchers and disciplines (see boyd and Ellison 2008).[2] A primary challenge to fully understanding the nature and dynamic of social networks is obtaining sufficient data. Most existing studies rely on external surveys of social networking participants, ethnographies of smaller subsets of subjects, or the analysis of limited profile information extracted from what subjects chose to make visible. As a result, the available data can often be tainted due to self-reporting biases and errors, have minimal representativeness of the entire population, or fail to reflect the true depth and complexity of the information users submit (and create) on social networking sites.

Recognizing the data limitations faced by typical sociological studies of online social network dynamics, a group of researchers from Harvard University and the University of California—Los Angeles set out to construct a more robust dataset that would fully leverage the rich data available on social networking websites.[3] Given its popularity, the researchers chose the social network site Facebook as their data source, and located a university that allowed them to download the Facebook profiles of every member of the freshman class:

> With permission from Facebook and the university in question, we first accessed Facebook on March 10 and 11, 2006 and downloaded the profile and network data provided by one cohort of college students. This population, the freshman class of 2009 at a diverse private college in the Northeast U.S., has an exceptionally high participation rate on Facebook: of the 1640 freshmen students enrolled at the college, 97.4% maintained Facebook profiles at the time of download and 59.2% of these students had last updated their profile within 5 days. (Lewis et al. 2008, p. 331)

This first wave of data collection took place in 2006, during the spring of the cohort's freshman year, and data collection was repeated annually until 2009, when the vast majority of the study population will have graduated, providing 4 years of data about this collegiate social network. Each student's official housing records were also obtained from the university, allowing the researchers to "connect Internet space to real space" (Kaufman 2008a).

The uniqueness of this dataset is of obvious value for sociologists and Internet researchers. The data was extracted directly from Facebook without direct interaction with the subjects or reliance on self-reporting instruments, either of which could taint the data collected. The dataset includes demographic, relational, and cultural information on each subject, allowing broad analyses beyond more simple profile scraping methods. The inclusion of housing data for each of the 4 years of the study for analysis of any connection between "physical proximity, emerging roommate and friendship groups in the real world and the presence of these two types of relationships in their Facebook space" (Kaufman 2008a). Most importantly, the dataset represents nearly a complete cohort of college students, allowing the unique analysis of "complete social universe" (Kaufman 2008a), and it is longitudinal,

[1] While no individuals within the T3 dataset were positively identified (indeed, the author did not attempt to re-identify individuals), discovering the source institution makes individual re-identification much easier, perhaps even trivial, as discussed below.

[2] See also bibliography maintained by danah boyd at http://www.danah.org/SNSResearch.html.

[3] The research team includes Harvard University professors Jason Kaufman and Nicholas Christakis, UCLA professor Andreas Wimmer, and Harvard sociology graduate students Kevin Lewis and Marco Gonzalez.

providing the ability to study how the social network changes over time.

As a result of its uniqueness, the dataset can be employed for a number of research projects that have heretofore been difficult or impossible to pursue. As one of the "Tastes, Ties, and Time" researchers noted, "We're on the cusp of a new way of doing social science… Our predecessors could only dream of the kind of data we now have" (Nicholas Christakis, qtd in Rosenbloom 2007).

## The dataset release

The "Tastes, Ties, and Time" project has been funded, in part, by a grant from the National Science Foundation,[4] who mandates certain levels of data sharing as a condition of its grants.[5] As a result, the Facebook dataset is being made available for public use in phases, roughly matching the annual frequency of data collection: wave 1 in September 2008, wave 2 in the fall of 2009, wave 3 in the fall of 2010, and wave 4 in the fall of 2011 (Lewis 2008, p. 3).

The first wave of data, comprising of "machine-readable files of virtually all the information posted on approximately 1700 FB profiles by an entire cohort of students at an anonymous, northeastern American university," was publicly released on September 25, 2008 (N.A. 2008).[6] Prospective users of the dataset are required to submit a brief statement detailing how the data will be used, and access is granted at the discretion of the T3 research team. Researchers are also required to agree to a "Terms and Conditions of Use" statement in order to gain access to the dataset, consenting to various licensing, use, and attribution provisions.

A comprehensive codebook was downloadable without the need to submit an application, which included detailed descriptions and frequencies of the various data elements (see Lewis 2008), including gender, race, ethnicity, home state, political views, and college major. For example, the codebook revealed that the dataset included 819 male and 821 female subjects, and that there were 1 self-identified Albanian, 2 Armenians, 3 Bulgarians, 9 Canadians, and so on.

The codebook also included an account of the steps taken by the T3 researchers in an attempt to protect subject privacy:

> All data were collected with the permission of the college being studied, the college's Committee on the Use of Human Subjects, as well as Facebook.com. Pursuant to the authors' agreement with the Committee on the Use of Human Subjects, a number of precautionary steps were taken to ensure that the identity and privacy of students in this study remain protected. Only those data that were accessible by default by each RA were collected, and no students were contacted for additional information. All identifying information was deleted or encoded immediately after the data were downloaded. The roster of student names and identification numbers is maintained on a secure local server accessible only by the authors of this study. This roster will be destroyed immediately after the last wave of data is processed. The complete set of cultural taste labels provides a kind of "cultural fingerprint" for many students, and so these labels will be released only after a substantial delay in order to ensure that students' identities remain anonymous. Finally, in order to access any part of the dataset, prospective users must read and electronically sign [a] user agreement… (Lewis 2008, p. 29)

These steps taken by the T3 researchers to remove identifying information reveal an acknowledgment of—and sensitivity to—the privacy concerns that will necessarily arise given the public release of such a rich and complete set of Facebook data. Their intent, as expressed by the project's principle investigator, Jason Kaufman, was to ensure that "all the data is cleaned so you can not connect anyone to an identity" (Kaufman 2008a). Unfortunately, the T3 researchers were overly optimistic.

## Partial re-identification and withdrawal of dataset

Cognizant of the privacy concerns related to collecting and releasing detailed Facebook profile data from a cohort of college students, the T3 research team—in good faith—took a number of steps in an attempt to protect subject privacy, including review by their institutional review board, the removal of student names and identification numbers from the dataset, a delay in the release of the cultural interests of the subjects, and requiring other researchers to agree to a "terms and conditions for use" that prohibited any attempts to re-identify subjects, to disclose any identities that might be inadvertently re-identified, or otherwise to compromise the privacy of the subjects.

[4] See "Social Networks and Online Spaces: A Cohort Study of American College Students", Award #0819400, http://www.nsf.gov/awardsearch/showAward.do?AwardNumber=0819400.

[5] See relevant National Science Foundation Grant General Conditions (GC-1), section 38. Sharing of Findings, Data, and Other Research Products (http://www.nsf.gov/publications/pub_summ.jsp?ods_key=gc109).

[6] The dataset is archived at the IQSS Dataverse Network at Harvard University (http://dvn.iq.harvard.edu/dvn/).

However, despite these efforts, the team's desire to ensure "all the data is cleaned so you can not connect anyone to an identity" fell short. On September 29, 2008, only 4 days after the initial data release, Fred Stutzman, a Ph.D. student at the University of North Carolina at Chapel Hill's School of Information and Library Science, questioned the T3 researchers' faith in the non-identifiability of the dataset:

> The "non-identifiability" of such a dataset is up for debate. A friend network can be thought of as a fingerprint; it is likely that no two networks will be exactly similar, meaning individuals may be able to be identified in the dataset post-hoc… Further, the authors of the dataset plan to release student "Favorite" data in 2011, which will provide further information that may lead to identification. (Stutzman 2008)

Commenting on Stutzman's blog post on the subject, Eszter Hargittai, an Associate Professor of Communication Studies at Northwestern University, sounded similar concerns:

> I think it's hard to imagine that some of this anonymity wouldn't be breached with some of the participants in the sample. For one thing, some nationalities are only represented by one person. Another issue is that the particular list of majors makes it quite easy to guess which specific school was used to draw the sample. Put those two pieces of information together and I can imagine all sorts of identities becoming rather obvious to at least some people. (Hargittai 2008)

Stutzman and Hargittai share a fear of the possible re-identification of the presumed anonymous Facebook dataset that has been made available to the public. Stutzman's concern over the ability to exploit the uniqueness of one's social graph to identify an individual within a large dataset has proven true in numerous cases (see, for example, Narayanan and Shmatikov 2008, 2009). Hargittai suggests that the uniqueness of the some of the data elements makes identifying the source of the data—and therefore some of the individual subjects—quite trivial. Hargittai's fears were correct.

## Partial re-identification

Within days of its public release, the source of the T3 dataset was identified as Harvard College (see Zimmer 2008a, b). Most striking about this revelation was that the identification of the source of the Facebook data did not require access to the full dataset itself.

Using only the freely available codebook and referencing various public comments about the research, the source of the data was quickly narrowed down from over 2000 possible colleges and universities to a list of only seven (Zimmer 2008b). An examination of the codebook revealed the source was a private, co-educational institution, whose class of 2009 initially had 1640 students in it. Elsewhere, the source was identified as a "New England" school. A search through an online college database[7] revealed only seven private, co-ed colleges in New England states (CT, ME, MA, NH, RI, VT) with total undergraduate populations between 5000 and 7500 students (a likely range if there were 1640 in the 2006 freshman class): Tufts University, Suffolk University, Yale University, University of Hartford, Quinnipiac University, Brown University, and Harvard College.

Upon the public announcement of this initial discovery, and general criticism of the research team's 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" (Kaufman 2008b). He 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 people's 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.) (Kaufman 2008b)

And then:

> 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). (Kaufman 2008c)

However, little "extreme effort" was needed to further "crack" the dataset; it was accomplished a day later, again without ever looking at the data itself (Zimmer 2008a). As Hargittai recognized, the unique majors listed in the codebook allowed for the ultimate identification of the source university. Only Harvard College offers the specific variety of the subjects' majors that are listed in the codebook, such as Near Eastern Languages and Civilizations,

[7] College Board, http://www.collegeboard.com.

Studies of Women, Gender and Sexuality, and Organismic and Evolutionary Biology. The identification of Harvard College was further confirmed after analysis of a June 2008 video presentation by Kaufman, where he noted that "midway through the freshman year, students have to pick between 1 and 7 best friends" that they will essentially live with for the rest of their undergraduate career (Kaufman 2008a). This describes the unique method for determining undergraduate housing at Harvard: all freshman who complete the fall term enter into a lottery, where they can designate a "blocking group" of between 2 and 8 students with whom they would like be housed in close proximity.[8]

In summary, the source of the T3 dataset was established with reasonable certainly in a relatively short period of time, without needing to download or access the dataset itself. While individual subjects were not identified in this process, the ease of identification of the source places their privacy in jeopardy given that the dataset contains a relatively small population with many unique individuals. The hopes by the T3 research team that "extreme effort" would be necessary to "crack" the dataset were, unfortunately, overly optimistic.

### Withdrawal of the dataset

The announcement of this likely identification of the source of the Facebook dataset did not prompt a public reply by the T3 research team, but within 1 week of the discovery, the access page for the "Tastes, Ties, and Time" dataset displayed the following message, indicating that the dataset was, at least for the moment, no longer publicly available:

> Note: As of 10/8/08, prospective users may still submit requests and research statements, but the approval process will be delayed until further notice. We apologize for the inconvenience, and thank you for your patience.[9]

Then, in March 2009, the page was updated with a new message acknowledging the removal was in response to concerns over student privacy:

> UPDATE (3/19/09): Internal revisions are almost complete, and we expect to begin distributing again in the next 2–3 weeks. In the meantime, please DO NOT submit new dataset requests; but please check back frequently at this website for a final release notice. We again apologize for any inconvenience, and thank you for your patience and understanding as we work to ensure that our dataset maintains the highest standards for protecting student privacy.[10]

A full year after the initial release, the dataset remains unavailable, with the following message greeting interested researchers:

> UPDATE (10/2/09): The T3 dataset is still offline as we take further steps to ensure the privacy of students in the dataset. Please check back later at this site for additional updates- a notice will be posted when the distribution process has resumed.[11]

These messages noting the restricted access to the Facebook dataset to "ensure that our dataset maintains the highest standards for protecting student privacy" suggest that the re-identification of the source as Harvard College was correct, and that the T3 research team is re-evaluating their processes and procedures in reaction.

## The insufficiency of privacy protections in the T3 project

The changing nature—and expectations—of privacy in online social networks are being increasingly debated and explored (see, for example, Gross and Acquisti 2005; Barnes 2006; Lenhart and Madden 2007; Nussbaum 2007; Solove 2007; Albrechtslund 2008; Grimmelmann 2009). The events surrounding the release of the Facebook data in the "Tastes, Ties, and Time" reveals many of the fault lines within these debates. Critically examining the methods of the T3 research project, and the public release of the dataset, reveals numerous conceptual gaps in the understanding the nature of privacy and anonymity in the context of social networking sites.

The primary steps taken by the T3 research team to protect subject privacy (quoted above), can be summarized as follows:

1. Only those data that were accessible by default by each RA were collected, and no students were contacted for additional information.
2. All identifying information was deleted or encoded immediately after the data were downloaded.
3. The complete set of cultural taste labels provides a kind of "cultural fingerprint" for many students, and so these labels will be released only after a substantial

[8] This process is described at the Harvard College Office of Residential Life website: http://www.orl.fas.harvard.edu/icb/icb.do?keyword=k11447&tabgroupid=icb.tabgroup17715.

[9] Screenshot of http://dvn.iq.harvard.edu/dvn/dv/t3 taken on October 22, 2008, on file with author.

[10] Screenshot of http://dvn.iq.harvard.edu/dvn/dv/t3 taken on March 27, 2009, on file with author. Webpage remains unchanged as of April 29, 2009.

[11] Screenshot of http://dvn.iq.harvard.edu/dvn/dv/t3 taken on November 1, 2009, on file with author. As of May 29, 2010, this message remains in place.

delay in order to ensure that students' identities remain anonymous.
4. In order to access any part of the dataset, prospective researchers must agree to a "terms and conditions for use" that prohibits any attempts to re-identify subjects, to disclose any identities that might be inadvertently re-identified, or otherwise to compromise the privacy of the subjects.
5. The entire research project, including the above steps, were reviewed and approved by Harvard's Committee on the Use of Human Subjects.

While each of these steps reveal good-faith efforts to protect the privacy of the subjects, each has serious limitations that expose a failures by the researchers to fully understand the nature of privacy in online social network spaces, and to design their research methodology accordingly. Each will be considered below, followed by a brief discussion of some of the public comments made by the T3 research team in defense of their methods and the public release of the dataset.

## Use of in-network RAs to access subject data

In his defense of releasing subjects' Facebook profile data, Jason Kaufmann, the principle investigator of the T3 project, has stated that "our dataset contains almost no information that isn't on Facebook" and that "We have not accessed any information not otherwise available on Facebook" (Kaufman 2008c). Access to this information was granted by Facebook, but only through a manual process. Thus, research assistants (RA) from the source institution (presumably Harvard) were employed to perform the labor-intensive task of search for each first year student's Facebook page and saving the profile information. The dataset's codebook confirms that "Only those data that were accessible by default by each RA were collected, and no students were contacted for additional information" (Lewis 2008, p. 29).

The T3 codebook notes that of the 1,640 students in the cohort, 1,446 were found on Facebook with viewable profiles, 152 had a Facebook profile that was discoverable but not viewable by the RA, and 42 were undiscoverable (either not on Facebook or invisible to those not within their "friend" network) (Lewis 2008, p. 6).[12] Importantly, the codebook notes a peculiarity inherent with using in-network RAs to access the Facebook profile data:

> It is important to note that both undergraduate and graduate student RAs were employed for downloading data, and that each type of RA may have had a different level of default access based on individual students' privacy settings. In other words, a given student's information should not be considered objectively "public" or "private" (or even "not on Facebook")—it should be considered "public" or "private" (or "not on Facebook") from the perspective of the particular RA that downloaded the given student's data. (Lewis 2008, p. 6)

The T3 researchers concede that one RA might have different access to a student's profile than a different RA, and being "public" or "private" on Facebook is merely relative to that particular RAs level of access.

What appears to be lost on the researchers is that a subject might have set her privacy settings to be viewable to only to other users within her network, but to be inaccessible to those outside that sphere. For example, a Facebook user might decide to share her profile information only with other Harvard students, but wants to remain private to the rest of the world. The RAs employed for the project, being from the same network as the subject, would be able to view and download a subject's profile data that was otherwise restricted from outside view. Thus, her profile data—originally meant for only those within the Harvard network—is now included in a dataset released to the public. As a result, it is likely that profile information that a subject explicitly restricted to only "in network" participants in Facebook has been accessed from within that network, but then extracted and shared outside those explicit boundaries.

Given this likelihood, the justification that "we have not accessed any information not otherwise available on Facebook" is true only to a point. While the information was indeed available to the RA, it might have been accessible only due to the fact that the RA was within the same "network" as the subject, and that a privacy setting was explicitly set with the intent to keep that data within the boundaries of that network. Instead, it was included in a dataset released to the general public. This gap in the project's fundamental methodology reveals a troublesome lack of understanding of how users might be using the privacy settings within Facebook to control the flow of their personal information across different spheres, and puts the privacy of those subjects at risk.

## Removal or encoding of "identifying" information

In an effort to protect the identity of the subjects, researchers note that "All identifying information was deleted or encoded immediately after the data were downloaded"

[12] Facebook allows users to control access to their profiles based on variables such as "Friends only", or those in their "Network" (such as the Harvard network), or to "Everyone". Thus, a profile might not be discoverable or viewable to someone outside the boundaries of the access setting.

(Lewis 2008, p. 29), and that "all the data is cleaned so you can not connect anyone to an identity" (Kaufman 2008a). Student names were replaced by "unique identification numbers" and any e-mail addresses or phone numbers that appeared in the Facebook profile data were excluded from the published dataset.

Yet, as the AOL search data release revealed, even if one feels that "all identifying information" has been removed from a dataset, it is often trivial to piece together random bits of information to deduce one's identity (Barbaro and Zeller Jr 2006). The fact that the dataset includes each subjects' gender, race, ethnicity, hometown state, and major makes it increasingly possible that individuals could be identified, especially those with a unique set of characteristics. Repeating Hargittai's concern: "I think it's hard to imagine that some of this anonymity would not be breached with some of the participants in the sample" (Hargittai 2008).

For example, the codebook reveals that each of these states has only a single student represented in the dataset: Delaware, Louisiana, Mississippi, Montana, and Wyoming. Similarly, there are only single instances of students identified as Albanian, Hungarian, Iranian, Malaysian, Nepali, Philippino, and Romanian. Their uniqueness very well might have resulted in publicity: it is possible that local media featured their enrollment at Harvard, or that the local alumni organization listed their name in a publicly-accessible newsletter, and so on. If such unique individuals can be personally identified using external sources, and then located within the dataset, one might also learn his/her stated political views or sexual preference, resulting in a significant privacy breach.

This reveals that even when researchers believe they have removed or encoded "all identifying information," there often remains information that could just as easily be used to re-identify individuals.[13] The T3 researchers' belief that stripping names alone is sufficient resembles the typical definition of "personally identifiable information" (PII) within the United States legal framework. As defined in California law, for example, PII is typically limited to an individual's name or other personally identifiable elements such as a social security number, a driver's license number, or a credit card number.[14] So long as these identifiers are removed from a dataset, it is presumed to be sufficiently anonymous.

However, others take a much broader stance in what constitutes personally identifiable information. The European Union, for example, defines PII much more broadly to include:

> [A]ny information relating to an identified or identifiable natural person…; an identifiable person is one who can be identified, directly or indirectly, in particular by reference to an identification number or to one or more factors specific to his physical, physiological, mental, economic, cultural or social identity.[15]

Thus, while the T3 researchers might have felt simply removing or coding the subjects' names or other specific identifiers from the dataset was sufficient, had they followed the European Union's guidance, they would have recognized that many of the subjects' "physical, physiological, mental, economic, cultural or social identity" could also be used for re-identification. Even after removing the names of the subjects, since the dataset still includes race, ethnicity, and geographic data, re-identification of individual subjects remains a real possibility.

## Delay in release of cultural taste data

Despite the apparent lack of use of the EU's more stringent definition of "personally identifiable information," the T3 researchers do recognize the unique nature of the cultural taste labels they have collected, referring to them as a kind of "cultural fingerprint". To protect subject privacy, the cultural tastes identified by the researchers have been assigned a unique number, and only the numbers will be associated with students for the initial data releases. The entire set of the actual taste labels will only be released in the fall of 2011, corresponding with the release of the wave 4 data.

The T3 researchers are right to recognize how a person's unique set of cultural tastes could easily identify her. Yet, merely instituting a "substantial delay" before releasing this personal data does little to mitigate the privacy fears. Rather, it only delays them, and only by 3 years. Researchers routinely rely on datasets for years after their initial collection: some influential studies of search engine behavior rely on nearly 10-year-old data (see, for example, Jansen and Resnick 2005; Jansen and Spink 2005), and these subjects' privacy needs do not suddenly disappear when they graduate from college in 2011.

Most surprisingly, despite the T3 researchers' recognition of the sensitive nature of the cultural data, they will

[13] Simply stripping names from records is rarely a sufficient means to keep a dataset anonymous. For example, Latanya Sweeny has shown that 87 percent of Americans could be identified by records listing solely their birth date, gender and ZIP code (Sweeney 2002).

[14] See, for example, the California Senate Bill 1386, http://info.sen.ca.gov/pub/01-02/bill/sen/sb_1351-1400/sb_1386_bill_20020926_chaptered.html.

[15] European Union Data Protection Directive 95/46/EC, http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:31995L0046:EN:HTML.

provide immediate access to it on a case-by case basis. As the codebook reveals:

> In the meantime, if prospective users wish to access some subset of the taste labels, special arrangements may be made on a case-by-case basis at the discretion of the authors (send request and detailed justification to t3dataset@gmail.com). (Lewis 2008, p. 20)

No further guidance is provided as to what kinds of arrangements are made and what justifications are needed to make such an exception. If the T3 research team felt strongly enough that it was necessary to encode and delay the release of the subjects' "cultural fingerprints", it does not seem appropriate to announce that exceptions can be made for its release to selected researchers prior to the 3-year delay. If it is potentially privacy invading content, it simply should not be released.

### Terms of use statement

Researchers wanting access to the T3 dataset must (electronically) sign a Terms and Conditions of Use statement. The statement includes various covenants related to protecting the privacy of the subjects in the dataset, including (as numbered in the original):

3. I will use the dataset solely for statistical analysis and reporting of aggregated information, and not for investigation of specific individuals or organizations, except when identification is authorized in writing by the Authors.
4. I will produce no links among the Authors datasets or among the Authors data and other datasets that could identify individuals or organizations.
5. I represent that neither I, nor anyone I know, has any prior knowledge of the possible identities of any study participants in any dataset that I am being licensed to use.
6. I will not knowingly divulge any information that could be used to identify individual participants in the study, nor will I attempt to identify or contact any study participant, and I agree to use any precautions necessary to prevent such identification.
7. I will make no use of the identity of any person or establishment discovered inadvertently. If I suspect that I might recognize or know a study participant, I will immediately inform the Authors, and I will not use or retain a copy of data regarding that study participant. If these measures to resolve an identity disclosure are not sufficient, the Authors may terminate my use of the dataset. (reproduced at Lewis 2008, p. 30)

The language within this statement clearly acknowledges the privacy implications of the T3 dataset, and might prove effective in raising awareness among potential researchers. However, studies have shown that users frequently simply "click through" such agreements without fully reading them or recognizing they are entering into a legally binding contract (Gatt 2002), and it is unclear how the T3 researchers specifically intend to monitor or enforce compliance with these terms. While requiring a terms of use is certainly a positive step, without enforcement it might have limited success in deterring any potential privacy-invasive use of the data.

### IRB approval

As required of any research project involving human interaction, clearance for the research project and data release was provided by Harvard's intuitional review board (IRB), known as the Committee on the Use of Human Subjects in Research.[16] As Kaufman commented: "Our IRB helped quite a bit as well. It is their job to insure that subjects' rights are respected, and we think we have accomplished this" (Kaufman 2008c). Elsewhere he has noted that "The university in question allowed us to do this and Harvard was on board because we don't actually talk to students, we just accessed their Facebook information" (Kaufman 2008a).

Just as we can question whether the T3 researchers full understood the privacy implications of the research, we must critically examine whether Harvard's IRB—a panel of experts in research ethics—also sufficiently understood how the privacy of the subjects in the dataset could be compromised. For example, did the IRB recognize, as noted above, that using an in-network research assistant to pull data could circumvent privacy settings intended to keep that data visible to only other people at Harvard? Or did the IRB understand that individuals with unique characteristics could easily be extracted from the dataset, and perhaps identified? It is unclear whether these concerns were considered and discarded, or whether the IRB did not fully comprehend the complex privacy implications of this particular research project.[17] In either case, the potential privacy-invading consequences of the T3 data release suggest a possible lapse of oversight at some point of the IRB review process.

### Other public comments

Beyond the shortcomings of the documented efforts to protect the privacy of the T3 dataset subjects, the researchers have made various public comments that reveal

[16] http://www.fas.harvard.edu/~research/hum_sub/.

[17] Attempts to obtain information about the IRB deliberations with regard to the T3 project have been unsuccessful.

additional conceptual gaps in their understanding of the privacy implications of the T3 research project.[18]

For example, when confronted with the potential re-identifiability of the dataset, Kaufman responded by pondering "What might hackers want to do with this information, assuming they could crack the data and 'see' these people's Facebook info?" and later acknowledging "Nonetheless, seeing your thought process—how you would attack this dataset—is extremely useful to us" (Kaufman 2008b). Kaufman's mention of "hackers", "attacking" the dataset, and focusing on what someone might "do" with this information exposes a *harm*-based theory of privacy protection. Such a position supposes that so long as the data can be protected from attack by hackers or others wishing to "do" something harmful once gaining access, the privacy of the subjects can be maintained. Such a position ignores the broader *dignity*-based theory of privacy (Bloustein 1964). Such a stance recognizes that one does not need to be a victim of hacking, or have a tangible harm take place, in order for there to be concerns over the privacy of one's personal information. Rather, merely having one's personal information stripped from the intended sphere of the social networking profile, and amassed into a database for external review becomes an affront to the subjects' human dignity and their ability to control the flow of their personal information.

The distinction between harm- and dignity-based theories of privacy are understood—and often debated—among privacy scholars, but when asked if they conferred with privacy experts over the course of the research and data release, Kaufman admits that "we did not consult [with] privacy experts on how to do this, but we did think long and hard about what and how this should be done" (Kaufman 2008c). Given the apparent focus on data security as a solution to privacy, it appears the T3 research team would have benefited from broader discussions on the nature of privacy in these environments.[19]

The T3 researchers also claim that there should be little concern over the ethics of this research since the Facebook data gathered was already publicly available. As Kaufman argues:

> On the issue of the ethics of this kind of research—Would you require that someone sitting in a public square, observing individuals and taking notes on their behavior, would have to ask those individuals' consent in advance? 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… (Kaufman 2008c)

This justification presents a false comparison. The "public square" example depends on random encounters of people who happen to be in the square at the precise time as the researcher. Further, the researchers cannot observe everyone simultaneously, and instead must select which individuals to focus their attention, leaving some subjects out of the dataset. Finally, the data gathered is imprecise, and limited to the researchers ability to discern gender, age, ethnicity, and other physically-observable characteristics.

By contrast, the T3 researchers utilized an in-network research assistant to systematically access and download an entire cohort of college students' Facebook profile pages, each year for 4 years. They successfully targeted a specific and known group of students, obtaining a list of names and e-mail addresses of the students from the source university to improve their ability to gather data on the entire population. The data acquired included not only the subjects' self-reported gender and ethnicity, but also their home state, nation of origin, political views, sexual interests, college major, relational data, and cultural interests—data which would be considerably more difficult to obtain through observations in a public square. Suggesting that the two projects are similar and carry similar (and minimal) ethical dilemmas reveals a worrisome gap in the T3 research team's understanding of the privacy and ethical implications of their project.

## The ethics of the "Tastes, Ties, and Time" project

The above discussion of the unsatisfactory attempts by the T3 researchers to protect subject privacy illuminates two central ethical concerns with the "Tastes, Ties, and Time" project: the failure to properly mitigate what amounts to violations of the subjects' privacy, and, thus, the failure to adhere to ethical research standards.

### Privacy violations

The proceeding discussion notes numerous failures of the T3 researchers to properly understand the privacy implications of the research study. To help concretize these concerns, we can gather them into the following four

[18] This section is intended as an informal analysis of the discourse used when talking about the T3 project. It is meant to reveal gaps in broader understanding of the issues at hand, and not necessarily directed against a particular speaker.

[19] After the T3 research project was funded and well underway, Kaufman became a fellow at the Berkman Center for Internet & Society at Harvard University, an organization dedicated to studying a number of Internet-related issues, including privacy. While Kaufman presented preliminary results of his research to the Berkman community prior to joining the center (Kaufman 2008a), there is no evidence that others at Berkman were consulted prior to the release of the T3 dataset.

salient dimensions of privacy violations, as organized by Smith et al. (1996) and based on thorough review of privacy literature: the amount of personal information collected, improper access to personal information, unauthorized secondary use of personal information, and errors in personal information.[20] Viewing the circumstances of the T3 data release through the lens of this privacy violation framework helps to focus the ethical deficiencies of the overall project.

*Amount of personal information collected*

Privacy violations can occur when "extensive amounts of personally identifiable data are being collected and stored in databases" Smith et al. (1996, p. 172). Notably, the "Tastes, Ties, and Time" project's very existence is dependent on the extensive collection of personal data. The T3 project systematically, and regularly over a 4-year period, collected a vast amount of personal information on over 1,500 college students. Individual bits of data that might have been added and modified on a subject's Facebook profile page over time were harvested and aggregated into a single database, co-mingled with housing data from an outside source, and then compared across datafiles.

*Improper access to personal information*

Privacy violations might occur when information about individuals might be readily available to persons not properly or specifically authorized to have access the data. As described above, subjects within the T3 dataset might have used technological means to restrict access to their profile information to only members of the Harvard community, thus making their data inaccessible to the rest of the world. By using research assistants from within the Harvard community, the T3 researchers—whether intentional or not—would be able to circumvent those access controls, thereby including these subjects' information among those with more liberal restrictions.

Further, no specific consent was sought or received from the subjects in the study; their profile information was simply considered freely accessible for collection and research, regardless of what the subject might have intended or desired regarding its accessibility to be harvested for research purposes. Combined, these two factors reveal how a privacy violation based on improper access has occurred due to the T3 project.

*Unauthorized secondary use*

Unauthorized secondary use of personal information is the concern that information collected from individuals for one purpose might be used for another secondary purpose without authorization form the individual, thus the subject loses control over their information. Within Smith et al.'s. (1996) framework, this loss of control over one's personal information is considered a privacy violation. At least two incidences of unauthorized secondary use of personal information can be identified in the T3 project. First, the students' housing information and personal email addresses were provided to the T3 researchers to aid in their data collection and processing. These pieces of information were initially collected by the university to facilitate various administrative functions, and not for secondary use to assist researchers looking for students' profiles on Facebook. Second, the very nature of collecting Facebook profile information, aggregating it, and releasing it for others to download invites a multitude of secondary uses of the data not authorized by the students. The data was made available on Facebook for the purpose of social networking among friends and colleagues, not to be used as fodder for academic research. Without specific consent, the collection and release of Facebook data invariably brings about unauthorized secondary uses.

*Errors in personal information*

Finally, privacy concerns arise due to the impact of possible errors within datasets, which has lead to various policies ensuring individuals are granted the ability to view and edit data collected about them to minimize any potential privacy violations.[21] In the T3 project, subjects were not aware of the data collection nor provided any access to view the data to correct for errors or unwanted information.

## Ethical research standards

Viewing the privacy concerns of the T3 data release through the lens of Smith et al.'s (1996) privacy violation framework helps to focus the ethical deficiencies of the overall project. In turn, our critique of the T3 project exposes various breeches in ethical research standards that, if followed, might have mitigated many of the privacy threats.

[20] I thank an anonymous reviewer for suggesting this organizing framework.

[21] See, for example, the United States Federal Trade Commission's Fair Information Practice Principles (http://www.ftc.gov/reports/privacy3/fairinfo.shtm), which include "Access" as a key provision, providing data subjects the ability to view and contesting inaccurate or incomplete data.

Ethical issues in human subjects research receive considerable attention, culminating in the scrutiny of research projects by Institutional Review Boards for the Protection of Human Subjects (IRB's) to review research according to federal regulations.[22] These regulations focus on research ethics issues such as subject safety, informed consent, and privacy and confidentiality. Others have then these broad standards and applied them specifically to Internet-based research and data collection. For example, the Association of Internet Researchers have issued a set of recommendations for engaging in ethical research online (see Ess and AoIR ethics working committee 2002), which places considerable focus on informed consent and respecting the ethical expectations within the venue under study.

As noted above, the T3 researchers did not obtain any informed consent by the subjects within the dataset (nor were they asked to do so by their Institutional Review Board). Further, as described in detail, the researchers failed to respect the expectations likely held by the subjects regarding the relative accessibility and purpose of their Facebook profile information. By failing to recognize that users might maintain strong expectations that information shared on Facebook is meant to stay on Facebook, or that only members of the Harvard network would ever have access to the data, the T3 researchers have failed in their duty to engage in ethically-based research.

## Conclusion

The events surrounding the release of the Facebook data in the "Tastes, Ties, and Time" project –including its methodology, its IRB approval, the way in which the data was released, and the viewpoints publicly expressed by the researchers—reveals considerable conceptual gaps in the understanding of the privacy implications of research in social networking spaces. As a result, threats to the privacy of the subjects under study persist, despite the good faith efforts of the T3 research team.

The purpose of this critical analysis of the T3 project is not to place blame or single out these researchers for condemnation, but to use it as a case study to help expose the emerging challenges of engaging in research within online social network settings. These include challenges to the traditional nature of consent, properly identifying and respecting expectations of privacy on social network sites, developing sufficient strategies for data anonymization prior to the public release of personal data, and the relative expertise of institutional review boards when confronted with research projects based on data gleaned from social media.

As made apparent to the position of some of the T3 research team that their data collection methods were unproblematic since the "information was already on Facebook", future researchers must gain a better understanding of the contextual nature of privacy in these spheres (Nissenbaum 1998, 2004, 2009), recognizing that just because personal information is made available in some fashion on a social network, does not mean it is fair game for capture and release to all (see, generally, Stutzman 2006; Zimmer 2006; McGeveran 2007; boyd 2008a). Similarly, the notion of what constitutes "consent" within the context of divulging personal information in social networking spaces must be further explored, especially in light of this contextual understanding of norms of information flow within specific spheres. The case of the T3 data release also reveals that we still have not learned the lessons of the AOL data release and similar instances where presumed anonymous datasets have been re-identified. Perhaps most significantly, this case study has uncovered possible shortcomings in the oversight functions of institutional review boards, the very bodies bestowed with the responsibility of protecting the rights of data subjects.

Overcoming these challenges and conceptual muddles is no easy task, but three steps can be taken immediately to guide future research in social media spaces. One, scholars engaging in research similar to the T3 project must recognize their own gaps in understanding the changing nature of privacy and the challenges of anonymizing datasets, and should strive to bring together an interdisciplinary team of collaborators to help ensure the shortcomings of the T3 data release are not repeated. Two, we must evaluate and educate IRBs and related policy makers as to the complexities of engaging in research on social networks.[23] And three, we must ensure that our research methods courses, codes of best practices, and research protocols recognize the unique challenges of engaging in research on Internet and social media spaces.[24]

The "Tastes, Ties, and Time" research project might very well be ushering in "a new way of doing social

[22] See Part 46 Protection of Human Subjects of Title 45 Public Welfare of the Code of Federal Regulations at http://www.hhs.gov/ohrp/humansubjects/guidance/45cfr46.htm.

[23] See, for example, the "Internet Research Ethics: Discourse, Inquiry, and Policy" research project directed by Elizabeth Buchanan and Charles Ess (http://www.nsf.gov/awardsearch/showAward.do?AwardNumber=0646591).

[24] An important movement in this direction is the recently funded "Internet Research and Ethics 2.0: The Internet Research Ethics Digital Library, Interactive Resource Center, and Online Ethics Advisory Board" project, also directly by Elizabeth Buchanan and Charles Ess (http://www.nsf.gov/awardsearch/showAward.do?AwardNumber=0924604 and http://www.internetresearchethics.org/).

science", but it is our responsibility scholars to ensure our research methods and processes remain rooted in long-standing ethical practices. Concerns over consent, privacy and anonymity do not disappear simply because subjects participate in online social networks; rather, they become even more important.

**Acknowledgments** The author thanks the participants at the International Conference of Computer Ethics: Philosophical Enquiry in Corfu, Greece, as well as the Internet Research 10: Internet Critical conference in Milwaukee, Wisconsin, for their helpful comments and feedback. Additional thanks to Elizabeth Buchanan, Charles Ess, Alex Halavais, Anthony Hoffmann, Jon Pincus, Adam Shostack, and Fred Stutzman for their valuable insights and conversations, both online and off. The author also thanks the anonymous reviewers for their helpful suggestions and criticisms. This article would not have been possible without the research assistance of Wyatt Ditzler and Renea Drews. Finally, I would like to thank Jason Kaufman and Colin McKay at the Berkman Center for Internet & Society, for their valued and continued feedback regarding this work.

## References

Albrechtslund, A. (2008). Online social networking as participatory surveillance. *First Monday* Retrieved 2008, March 3, from http://www.uic.edu/htbin/cgiwrap/bin/ojs/index.php/fm/article/view/2142/1949.

Barbaro, M., & Zeller Jr, T. (2006). A face is exposed for AOL searcher no. 4417749. *The New York Times*, p. A1.

Barnes, S. (2006). A privacy paradox: Social networking in the United States. *First Monday* Retrieved October 12, 2007, from http://www.firstmonday.org/ISSUES/issue11_9/barnes/.

Bloustein, E. (1964). Privacy as an aspect of human dignity: An answer to Dean Prosser. *New York University Law Review, 39*, 962–1007.

boyd, D. (2008a). Putting privacy settings in the context of use (in Facebook and elsewhere). *Apophenia* Retrieved October 22, 2008, from http://www.zephoria.org/thoughts/archives/2008/10/22/putting_privacy.html.

boyd, D. (2008b). Taken out of context: American teen sociality in networked publics. Unpublished Dissertation, University of California-Berkeley.

boyd, D., & Ellison, N. (2008). Social network sites: Definition, history, and scholarship. *Journal of Computer-Mediated Communication, 13*(1), 210–230.

Ess, C., & AoIR ethics working committee. (2002). Ethical decision-making and Internet research. Retrieved March 12, 2010, from http://www.aoir.org/reports/ethics.pdf.

Gatt, A. (2002). Click-wrap agreements the enforceability of click-wrap agreements. *Computer Law & Security Report, 18*(6), 404–410.

Grimmelmann, J. (2009). Facebook and the social dynamics of privacy. *Iowa Law Review, 95*, 4.

Gross, R., & Acquisti, A. (2005). Information revelation and privacy in online social networks. Paper presented at the 2005 ACM workshop on Privacy in the electronic society, Alexandria, VA.

Jansen, B. J., & Resnick, M. (2005). Examining searcher perceptions of and interactions with sponsored results. Paper presented at the Workshop on Sponsored Search Auctions at ACM Conference on Electronic Commerce, Vancouver, BC.

Jansen, B. J., & Spink, A. (2005). How are we searching the world wide web? A comparison of nine search engine transaction logs. *Information Processing & Management, 42*(1), 248–263.

Kaufman, J. (2008a). Considering the sociology of Facebook: Harvard Research on Collegiate Social Networking [Video].: Berkman Center for Internet & Society.

Kaufman, J. (2008b). I am the Principal Investigator… [Blog comment]. *On the "Anonymity" of the Facebook dataset* Retrieved September 30, 2008, from http://michaelzimmer.org/2008/09/30/on-the-anonymity-of-the-facebook-dataset/.

Kaufman, J. (2008c). Michael—We did not consult… [Blog comment]. *michaelzimmer.org* Retrieved September 30, 2008, from http://michaelzimmer.org/2008/09/30/on-the-anonymity-of-the-facebook-dataset/.

Lenhart, A., & Madden, M. (2007). Teens, privacy & online social networks. *Pew internet & American life project* Retrieved April 20, 2007, from http://www.pewinternet.org/pdfs/PIP_Teens_Privacy_SNS_Report_Final.pdf.

Lewis, K. (2008). Tastes, Ties, and Time: Cumulative codebook. Retrieved September 30, 2008, from http://dvn.iq.harvard.edu/dvn/dv/t3.

Lewis, 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.

McGeveran, W. (2007). Facebook, context, and privacy. *Info/Law* Retrieved October 3, 2008, from http://blogs.law.harvard.edu/infolaw/2007/09/17/facebook-context/.

N.A. (2008). Tastes, Ties, and Time: Facebook data release. *Berkman Center for Internet & Society* Retrieved September 30, 2008, from http://cyber.law.harvard.edu/node/4682.

Narayanan, A., & Shmatikov, V. (2008). Robust de-anonymization of large sparse datasets. Paper presented at the IEEE Symposium on Security and Privacy, 2008.

Narayanan, A., & Shmatikov, V. (2009). De-anonymizing social networks. Paper presented at the 30th IEEE Symposium on Security and Privacy.

Nissenbaum, H. (1998). Protecting privacy in an information age: The problem of privacy in public. *Law and Philosophy, 17*(5), 559–596.

Nissenbaum, H. (2004). Privacy as contextual integrity. *Washington Law Review, 79*(1), 119–157.

Nissenbaum, H. (2009). *Privacy in context: Technology, policy, and the integrity of social life*. Stanford, CA: Stanford University Press.

Nussbaum, E. (2007). Kids, the Internet, and the end of privacy. *New York Magazine* Retrieved February 13, 2007, from http://nymag.com/news/features/27341/.

Rosenbloom, S. (2007). On Facebook, scholars link up with data. *New York Times* Retrieved September 30, 2008, from http://www.nytimes.com/2007/12/17/style/17facebook.html?ref=us.

Simmel, G., & Wolff, K. H. (1964). *The sociology of Georg Simmel*. Glencoe, Ill: Free Press.

Smith, H. J., Milberg, S. J., & Burke, S. J. (1996). Information privacy: Measuring individuals' concerns about organizational practices. *MIS Quarterly, 20*(2), 167–196.

Solove, D. (2007). *The future of reputation: Gossip, rumor, and privacy on the internet*. New Haven, CT: Yale University Press.

Stutzman, F. (2006). How Facebook broke its culture. *Unit Structures* Retrieved 2008, October 3, from http://chimprawk.blogspot.com/2006/09/how-facebook-broke-its-culture.html.

Stutzman, F. (2008). Facebook datasets and private chrome. *Unit Structures* Retrieved 2008, September 30, from http://fstutzman.com/2008/09/29/facebook-datasets-and-private-chrome/.

Sweeney, L. (2002). k-anonymity: A model for protecting privacy. *International Journal of Uncertainty Fuzziness and Knowledge-Based Systems, 10*(5), 557–570.

Wellman, B., & Berkowitz, S. D. (1988). *Social structures: A network approach*. Cambridge: University Press Cambridge.

Zimmer, M. (2006). More on Facebook and the contextual integrity of personal information flows. *michaelzimmer.org* Retrieved 2008, October 3, from http://michaelzimmer.org/2006/09/08/more-on-facebook-and-the-contextual-integrity-of-personal-information-flows/.

Zimmer, M. (2008a). More on the "Anonymity" of the Facebook dataset—It's Harvard College. *michaelzimmer.org* Retrieved October 3, 2008, from http://michaelzimmer.org/2008/10/03/more-on-the-anonymity-of-the-facebook-dataset-its-harvard-college/.

Zimmer, M. (2008b). On the "Anonymity" of the Facebook dataset. *michaelzimmer.org* Retrieved September 30, 2008, from http://michaelzimmer.org/2008/09/30/on-the-anonymity-of-the-facebook-dataset/.