PRIVACY IN PUBLIC? THE ETHICS OF ACADEMIC RESEARCH WITH PUBLICLY AVAILABLE SOCIAL MEDIA DATA

by NATALIYA NEDZHVETSKAYA AND STEVEN LAUTERWASSER

The fact that social media data are public, are known to be public, and sell themselves partly on their public nature has allowed anyone with a purpose —e.g., academic researchers, government agencies, and private firms— to access and collect data, confronting few, if any, formal ethical challenges. Regulations differ between social media platforms and across borders, determined by either majority shareholders, government officials, or some uneasy balance of the two. When companies change hands, the issue of how data also changes hands should lead us to question what it means for that data to be public in the first place.

The publicness of social media data has trivialized the question of consent and the ethical concerns which would normally arise around data collection and processing, resulting in what Shoshanna Zuboff (2019) calls an "uncontract" between platforms, their users, and the third parties they contract with. Users submit to having their data collected by virtue of signing up for and posting on a social media platform without any explicit discussion of the exchange taking place.

Corporations, policymakers, and academic institutions have failed to achieve any kind of consensus around conducting research with social media data, compared with other data types, such as biometric data. Researchers, consequently, face real conflicts between their aim to produce new knowledge and their obligation to not harm those who are, unwittingly, taken as subjects of their work.

This essay attempts to think through these conflicts from the standpoint of researchers' "duty of care" to those they research, grounded in our own grappling with these conflicts in a study of polarization on Twitter, for which we gathered nine months of tweets (November 2019 to July 2020) on the U.S. presidential election. The first section presents an analysis of the challenges posed by public data, while the second offers a summary of existing solutions and their limits. The

third delves into common uses of this data, implications of its mis-use, and best practices for researchers (laid out schematically in Table 1). We conclude with recommendations for addressing these challenges.

Table 1. Common Uses of User-Identifiable Data in Social Media Research		
Research Use of User-Identifiable Data	Research Context	Examples
Individual spotlighting	Qualitative exploration, explanation of anomalies	Understanding the most active users or the most retweeted tweets; qualitative investigations of bot-ness
Inferred Behavior	Modeling, e.g., structural topic models	Connecting certain topics with combinations of individual user level covariates
Representatives Examples	Model explanation, e.g., topic models	Using specific tweets as exemplars for their corresponding topics in a topic model or other inductive analysis

The Problem of "Public" Data

Our purpose as academic researchers is to produce high-quality research with as accurate a dataset as possible. How should we consider ethics in the face of these professional obligations? We make two substantive interventions. First, we consider what "public data" might mean for social media users. Second, we ask whether taking this perspective into account has any implications for researchers' "duty of care."

The Meaning of Public in Public Data

We define "being public" as creating public data about oneself, visible to any other user of the publicly available platform (when online) or inhabitant of a publicly accessible space (when in-person). The fundamental difference between being public in physical space and on the internet is the fact that the former tends to be ephemeral and circumscribed while the latter generates a permanent, searchable, and potentially identifiable record.

Importantly, the records which online public appearances leave behind tend to carry with them detail (i.e. what exactly was said), but not context (i.e. was it said to a friend in jest back when you had three followers?). This is what is known as

the problem of "context collapse" –the fact that social media allows utterances to travel well beyond their initial audiences, thus challenging the specificity and variability of self-presentation that sociologists argue is a fundamental aspect of normal social life (Goffman 1959; Marwick and boyd 2011). Twitter users have been shown to recreate contextual audiences online by strategically switching the privacy setting of their account from "public" to "protected" as a way of regulating their reach (Kekulluoglu 2022).

Pragmatic innovations like this, however, cannot solve these problems, in part because they are grounded in and exacerbated by the fundamental, experiential contradictions of being public. Being public online feels private and anonymous in a way physical space cannot. It is easy to feel anonymous online: the lack of embodiment means you cannot be recognized except (potentially) by whatever pseudonym or avatar you have chosen. One can be "in" a public forum online without engaging and have no trace of having been there aside from the registration of their IP address to a company server or a cookie recorded by their browser– neither of which should, under most circumstances, be publicly accessible. By comparison, simply being physically present in public is inherently less anonymous because you are present and can be perceived. It is impossible to "lurk" in a coffee shop in the way you can on an internet forum.

But these experiential differences are misleading, because your actual personal control over your "being in public" is far higher in the physical world. Because of the ephemeral and circumscribed nature of physical presence, you have a direct, immediate, and generally accurate understanding your audience at any given moment and can modulate your behavior appropriately. Over the course of a single conversation, you can whisper, laugh, shout, etc. as appropriate to your setting and how you wish to be perceived, and whose attention you want. Without a video recording, an awkward interaction or a misspoken phrase will recede into the memories of those present in the room.

Online engagement allows none of that control. Interactions cannot be modulated to the same extent– attuned to an audience's facial expressions and reactions, adjusted instantaneously to feedback, and preserved selectively in the memories of only those present. The success of platforms that deliver ephemeral, small-group content, such as Snapchat and BeReal, suggests individuals and communities perceive a real benefit in these sorts of interactions. The internet may give us a greater feeling of control in some ways because it feels straightforwardly anonymous, but it provides almost none of the control which makes "being public" possible and functional.

We should also consider that much of the internet is *not* anonymous. Platforms such as Facebook, Twitter, and Instagram represented a move towards integrating physical personas and online personas in a way that earlier internet platforms, such as MySpace, did not. To the extent that any kind of fine-grained control over how we are perceived is even possible on the internet, its permanent, searchable, nature makes this completely useless because there is always the possibility of our actions being re-interpreted in later contexts. But our intuitions about what our choices in self-presentation *should* mean are still there, leading us astray.

This is a general problem with social media, but this publicness also poses specific challenges to academic researchers. A growing body of scholarship (e.g. see Fiesler and Proferes 2018; Gilbert et al. 2021) suggests that social media users are generally uncomfortable with undeclared research uses of their social media data. The lack of established norms in the field may fuel a sense of trust violation (Nissenbaum 2009). This discomfort should direct our attention to the tensions between our interests as researchers and our subjects' interests as users. Because there is a potential for real harm here.

These harms are grounded in precisely the ambivalence between the public and private we are discussing. Privacy is not simply a negative space, everything which is not public (Eiermann 2022). Rather, it is an affirmatively constructed response to a "socially created need" (Moore 1984:73). Privacy facilitates social interaction, identity construction, and stable ongoing relationships by letting us leave the "front stage" and giving us temporary respite from grasp of external judgment and obligation (Goffman 1959; Schwartz 1968). This function of privacy is complicated tremendously by what Brubaker has called "digital hyperconnectivity," the infinite and constant (potential) connection between, increasingly, everyone and everything, all the time (Brubaker 2020). But users continue to construct varied, contextual identities, in the face of (and sometimes in opposition to) these complications (van Dijck 2013; Wilson & Peterson 2002). By gathering, and potentially revealing, user data, we intervene in this context, and it is this intervention that, we contend, creates our "duty of care."

The Duty of Care

In the common law doctrine of negligence, "duty of care" describes a special relationship that creates an obligation for one party to take care to avoid or prevent harm to another (Gregory 1951). For instance, a doctor owes a duty of care to their patient in a way they do not to people encountered in everyday life, a duty created when they take that person on as a patient. It is in this sense that we argue

for researchers' duty of care: by collecting our subjects' data, we assume certain obligations towards them; namely, to take care to avoid harms that would not otherwise be our concern. This is our second problem: what is our responsibility to our subjects?

Anonymity

This responsibility comes up in different ways. For instance, a researcher may decide to illustrate a particular point by "spotlighting" a specific user as part of the production of empirical evidence. But how must researchers concretely balance the need for specificity in the example with the protection of an individual's anonymity? Furthermore, making the user's identification harder, for instance by using a pseudonym, may be insufficient if they are also quoted exactly: unless the post has been deleted, searching the internet for the quote itself is a trivial barrier. The usual rule that equates anonymity with name retention does not really apply under conditions of total publicness and searchability (Mancosu and Vegetti 2020). All this might raise the question of whether anonymity is feasible or even desirable. As Jerolmack and Murphy (2019) have pointed out for the case of ethnography, anonymity should not be the default ethical choice: for ethnographic subjects, anonymity is both nearly impossible to protect and often contrary to the subjects' own expressed interests.

But our situation is substantially different: though anonymity is difficult, public social media data generally lack the "small world" problem of ethnographic data. We *can* do much to protect the anonymity of our subjects, and, because we are collecting data without their knowledge, it is important that we attempt to. We do not know our subjects' interests, but we can sometimes infer them: accounts or individual tweets or even profiles that have been made private or deleted in the time since they were originally released can be understood as a kind of implied opt-out.

Data from our own work shows this issue clearly. Returning to a set of tweets collected for our study of the 2020 U.S. presidential election, we drew samples from the beginning (November 8, 2019) and end (July 5, 2020) of the collection period, and attempted to re-access those tweets (1).

1. The tweets were originally collected using Twitter's streaming API (i.e., the "firehose") based on a bag of words relating to the election (e.g., candidate or party names, "US election", etc.). This resulted in a dataset containing hundreds of thousands to millions of tweets per day. From these, we sampled 4200 from the first and last weeks of data collection (as attempting to resolve millions of tweet ids to check availability is prohibitively time intensive). Neither the original data nor the sample were filtered for bot accounts; however, the different classes of tweet (available, deleted, account protected, account suspended) are largely similar in statistical profile. For instance, though suspended accounts tend to be a bit younger, and available accounts a bit older, the distributions of account age are otherwise very similar. Thus, it seems unlikely that bot accounts are truly skewing the numbers presented here.

The results are striking: across both samples, only just more than half (54%) of the tweets were still accessible. Of the remainder, about 22% had been deleted by the user, 20% were from accounts since suspended, and around 4% were from protected accounts (defined here as either private or deactivated by the user) (2).

Now, the figures for these samples are almost certainly higher than they would be for a random sample of all tweets. The 2020 presidential election was a contentious topic and we would expect more missing tweets than for less contentious topics. However, discourse around contentious events is often exactly what we want to study (see, e.g., concerns around social media and echo chambers, misinformation, polarization, etc.). Furthermore, even substantially lower figures would still be significant.

This is one place where the tension between the researcher, user, and public interest becomes especially acute. Researchers want to ensure the most comprehensive and accurate dataset about past events and are usually compelled by journals to make these datasets publicly available to ensure reproducibility. The accurate representation of the historical record of online speech can be, itself, a public good (Bernstein et al. 2021). At the same time, however, users are keen to constantly readjust their self-presentation, following changes in their own trajectories and in public opinion.

This section would not be complete without some discussion of Large Language Models and their potential for privacy violations. In the process of amassing the large and indiscriminate training datasets necessary for building these sorts of models, researchers inadvertently gather personal information as well. Carlini et al. (2021) demonstrate how GPT-2 could be used to reveal personal information such as an individual's name, phone number, and physical address using a training data extraction attack. By taking advantage of the probabilistic statistics that undergird LLMs, attackers can generate queries with a low likelihood of occurring that cause the model to reproduce strings of text that it has effectively memorized. Researchers have suggested that an overhaul of our data governance infrastructure is necessary for such models to be deployed in a way that continues to protect data holders from inadvertent privacy violations (see Jernite et al., 2022).

2. There is some additional uncertainty around this number as the status message Twitter provides

Inferred Behavior

These concerns become more important when researchers generate individual characteristics from *patterns* in the data. Beyond what people willingly reveal, social media data allows anyone with the right skills to *infer* an awful lot by constructing variables that aggregate across posts—for instance, Kosinski, et al. (2013) used records of Facebook likes from 58,000 volunteers to generate models that predicted users' race (95% accuracy), sexual orientation (88% accuracy), and political party (85% accuracy).

While people may implicitly consent to the publicness of their individual tweets, they may not be aware of what their tweeting patterns add up to, nor desire that these derived patterns, in turn, be made "public," or shared widely with other entities (for instance under academic norms of transparency and reproducibility) (see Metcalf and Crawford 2016 on Hauge et al. 2016). Under these circumstances, we must ask whether datasets legitimately assembled *and produced* by researchers can reasonably be shared.

Who gets the rights to stay private?

There is an argument that we needn't concern ourselves with all of this: social media is, arguably, a conscious effort to speak to the world. Social media users are independent actors who make their own decisions within the confines of platforms that they have consented to use. The specificity of *this* historical moment, powered by social media, is that the public/private distinction has finally collapsed, so that everyone today, no matter how unknown, is a potential public figure (Dawson, 2018).

While such considerations may serve to superficially assuage our conscience, this discussion has made clear that these justifications are not as uncomplicated or straightforward as we might like them to be. The publicness of social media data, the problem of implied consent, and our duty of care, all feed into one another. Our duty of care becomes more serious in view of the dangers of publicity and the fact that some users might not have the choice to opt out of the "uncontract" system of most social media platforms. User agency cannot free us from ethical considerations.

Ethical Frameworks for Research on "Public" Social Media Data

Addressing these challenges requires an ethical framework. Such a framework currently comes in two different forms: as *public policy*, typically deployed by states or supranational political entities; or as *internal governance structures* developed by organizations themselves.

Public Policy

As user data has become more widely visible, state and government organizations have stepped up to protect the interests of citizens. While consumers have been the focus of much of this legislation, policymakers have also taken measures to support researchers who may require access to this data for their work. The California Consumer Privacy Act (CCPA) includes a clause that prohibits businesses from removing consumer data if it threatens "public or peer-reviewed scientific, historical, or statistical research in the public interest" provided such research "adheres to all other applicable ethics and privacy laws" and the data was obtained with "informed consent" from the consumer (CCPA 1798.105).

The E.U.'s General Data Protection Regulation (GDPR) provides similar exemptions but goes further to explicitly allow researchers to use consumer data for purposes other than its original intention, to obtain data across registries, and to store data for longer periods of time than originally stipulated– again, provided the data are used for purposes "in the public interest" (GDPR Article 5; Article 21).

Though these regulations give researchers some room to work by explicitly allowing certain uses of data, they are blunt instruments for actually regulating research. In practice, what companies decide to do with the data they control is a far more effective regulation of what research gets done. These choices can regulate quite precisely: beyond simply what data is public and accessible, firms increasingly partner directly with researchers. For example, Fradkin, Grewal, and Holtz (2021) combined public Airbnb reviews with proprietary data on their timing to study the effect of reciprocity on two-sided reputation systems. While clearly of academic and commercial benefit, the question of public interest need not even arise.

Public *harms*, however, can and do, as the egregious case of Cambridge Analytica illustrates (Rosenberg et al. 2018). The for-profit political consulting firm obtained the private records of millions of Facebook users through a psychologist at

Cambridge University, who claimed that he was collecting data for research purposes. Users' data was then applied to political manipulation operations leading up to the 2016 Brexit vote in the United Kingdom and the U.S. Presidential election. For academic researchers to maintain public trust– and also the legal privilege of greater access to data– it is essential to distinguish between acceptable and unacceptable uses and collaborations.

Internal Governance Structures

Second, private and public institutions can create *governance structures* that have the power to monitor research practices. Among those we consider are institutional review boards, professional organizations, and internal corporate auditors.

Institutional review boards (IRBs) are the most prominent examples of these governance structures. In the U.S. any institution conducting federally funded research must establish an institutional review board or research ethics committee to protect the rights of the individuals taking part in that research. More than 80 countries around the world have similar institutions (McCarthy 2008).

In theory, IRBs are well-positioned to consider the ethics of using social media datasets for academic research because they are independent and because their institutional mandate includes protecting research subjects from potential harm. While funded by the institutions they work for, they are overseen by federal law, which has the power to punish institutions that fail to comply by cutting off funding.

In practice, however, IRBs rarely concern themselves with public data source research precisely because it is public. The same federal regulation that mandates IRBs within most academic institutions also exempts publicly available data from their jurisdiction (McCarthy 2008). IRBs are incentivized to adhere to the law but not to push forward new ethical frameworks (Bernstein et al. 2021).

Professional organizations can also oversee the work of researchers, including independent researchers. For instance, the American Sociological Association (2018) has a Code of Ethics for its members. Though non-enforceable, they do provide guidance on how to consider public data sources in an ethical manner. Interestingly, while the Code acknowledges that such data can be considered public, there should nevertheless be efforts to maintain confidentiality and respect the privacy of certain Internet spaces. For example, researchers should not make attempts to "re-identify" data publicly gathered from the Internet (10.1.f). In addition, the Code acknowledges that not all spaces on the Internet can be

considered public, though it does not explain where to draw this line. General professional principles are useful for guiding researchers, but ultimately place the burden of deciding ethics on either the researcher or their institution.

At this point, the main regulation of the use of social media data by researchers comes from the platforms themselves. The recent series of sudden and dramatic changes to data availability from Twitter demonstrate the fundamental unreliability of relying on potentially capricious corporate policy. But even more stable oversight is not without its issues, as Meta's Oversight Board makes clear. The Board consists of eleven to forty members serving three-year terms. Notably, the initial set of Board members was selected by Meta management but future Boards should be appointed by an independent board of trustees (Oversight Board 2021). The strength of such institutions, in theory, is that they should function independently of the larger corporation. But Meta retains the power to disband its Oversight Board or re-write its charter without accountability to any higher power. In addition, the corporation can determine which ethical issues the Board addresses, effectively limiting its scope. Thus, although such structures claim to have the values of the public good in mind, they are in fact vulnerable to conflicts of interest, due to a lack of transparency and external accountability.

Recently, regulation has shifted to primarily a matter of restriction, with a number of platforms moving towards preventing publicly accessible user data from being collected on a large scale, except in carefully negotiated case-by-case agreements. Some of this is about ethics, some of it is about controlling the outcome, and some of it is about financial interest. Following the Cambridge Analytica scandal, Facebook restricted access to its API, ending at least the most easily accessible means of scraping its website (Mancosu and Vegetti 2020). LinkedIn has gone so far as to sue a data analytics company that scraped its website for violating its terms of service (Gatto and Almasi 2021). Critically, these decisions are made at the level of the company and can change quickly and unpredictably, for example, if company ownership changes. Twitter, which had historically been one of the few large platform companies to offer relatively unfettered access to its API, recently moved to ban free access after changes in company ownership (Kupferschmidt 2023).

Implications and Best Practices for Researchers

These are issues we cannot ignore: there are, as we have shown, real ethical concerns around the use of user-identifiable data. Neither, however, can they amount to total prohibitions. Institutions, as we have seen in the previous section, offer

only limited regulation around working with such data. However, working with social media data requires engaging with that data in ways that impinge on these issues. As researchers, we can address at least some of these issues by adopting more ethical research practices. In the following section, we discuss three of these challenges with specific examples from our own work. Table 2, at the end of the section, schematically lays out these common uses and research contexts, and suggests best practices.

Individual Spotlighting

The practice of spotlighting individual tweets or users has already been mentioned above for its obvious ramifications in potentially drawing outsized, unwanted attention to specific users. However, it is also an extremely useful tool both substantively and rhetorically throughout the research process. In our own work, targeted investigation of specific tweets and users was invaluable for building a qualitative understanding of Twitter during the 2020 election: who were the most active tweeters, and what was the most retweeted content? We propose minimizing potential harms by obtaining consent, when possible, from any individuals who are spotlighted.

Inferred Behavior

Part of what is useful about so-called big data is precisely the ability to assemble that data into complex inferences beyond what would be obvious in any given tweet by any given user. We discussed this above in the context of inferring individual characteristics, e.g., political orientation or sexuality. Structural topic modeling is a fairly benign example. Topic modeling is a way of computationally identifying the topics being discussed in a set of documents. Structural topic modeling then looks across documents to establish what topics are discussed by different kinds of authors and how. For our work on Twitter during the 2020 election, we used structural topic models to try and understand the nature of the conversation across political groups. But even here there is room for concern: our assignment to political groups was not based on explicit affiliation, but was itself inferred, and by drawing connections from documents to groups to users, we build associations between users and (given the nature of political discussion on social media) sometimes objectionable content. As a solution, we propose refraining from giving these inferences at the level of the individual user in any public presentation of data or results.

Representative Examples Supporting Broader Patterns

Pointing to individual documents, if not users, is fundamental to both the analysis and presentation of topic models, where examples are necessary for making sense of the computationally generated topics. However, this collides directly with the issue of perfect searchability, raising issues for anonymity even if no other information is provided. In some ways this has been both the easiest and hardest challenge to deal with: although necessary during analysis, the use of individual documents is not literally required to present a model effectively and can often be effectively replaced with synthetic summaries and topic-characteristic words. However, at the same time, individual documents bear substantive qualitative and rhetorical weight. One partial solution in our work has been to only present anonymized examples which are no longer available, generally because the account in question has since been suspended.

The purpose of these examples is not prescriptive but rather demonstrative. The exact tradeoff between our duty to produce good research and our duty of care to our subjects is fuzzy. None of these concerns can be allowed to be prohibitions, a fact amply illustrated in our work on the election. Contentious and historical events are important and, for instance, the implied opt-out of a deleted tweet must be weighed against its importance. However, we still must recognize these concerns and approach the tradeoffs they entail with respect. Put another way, we must begin to accept that our position with respect to our subjects has more in common with the complexities dealt with by ethnographers and interviewers than we have thus far assumed.

Table 2. Suggested Best Practices Around User-Identifiable Data in Social Media Research.

Research Use of User-Identifiable Data	Research Context	Examples
Individual spotlighting	Qualitative exploration, explanation of anomalies	Present the minimum possible information. Receive prior consent when spotlighting when possible, or spotlight public figures if that is not possible
Inferred Behavior	Modeling, e.g., structural topic models	Disconnect individual inferences and individual users in public presentations of the data and results

Table 2. Suggested Best Practices Around User-Identifiable Data in Social Media Research.		
Representatives Examples	Model explanation, e.g., topic models	Provide summaries rather than examples; only use examples which are still available (have not been removed by the user)

Conclusion: Privacy in Public

Social media architectures are optimized to drive their users to widely and publicly *expose* their lives, thoughts, and social relations. This has created an abundance of highly visible information about individuals with tremendous historical and economic value. But neither public policy nor internal governance structures adequately address the possible concerns users might raise about how their publicly accessible data might be studied, packaged, and recycled into derivative products. As researchers, we must think beyond raw regulatory limitations and consider a broader "duty of care" to the people whose data we are using. The "uncontract" nature of social media means that user consent is often implied as opposed to specifically granted (Zuboff 2019). We provide two recommendations going forward.

First, researchers should not make publicly available individual characteristics that are inferred from (but not stated in) a user's social media posts, such as religious affiliation, political orientation, or sexual identity. When users post on social media, few imagine that their posts will be collected and analyzed in this way. To conform to norms of reproducibility and transparency, researchers can instead make available the code that generates the derived characteristics. While not completely preventing others from accessing the derived characteristics, it places a significant hurdle in front of anyone who wants to access them.
To become a part of widespread ethical practice, such recommendations need to be embedded within institutionalized ethical frameworks that codify and enforce ethical guidelines. We see two possible, potentially complementary approaches. The first is to change the way that IRBs treat individual-level public data so that concerns about individual privacy in public data are part of regular human subjects ethical reviews. Although we are wary of mission creep and the bureaucratization of IRBs, we note that the potential harms to individual social media users align well with the types of harms that IRBs are already concerned with preventing. (3)

A second possibility is for the professional organizations of academic researchers to include ethical guidelines in their publication requirements for journals. As we have seen with recent movements for transparency and reproducibility, journal publication requirements provide powerful incentives for compliance at a natural point in the process of publishing scientific research. Researchers should take measures to institutionalize a "duty of care" using transparent enforcement mechanisms. This requires both adapting existing practices and creating new standards that can better protect the rights of social media users.

References

American Sociological Association. 2018. "American Sociological Association Code of Ethics." Washington, DC: American Sociological Association. Retrieved November 26, 2021 (https://www.asanet.org/sites/default/files/asa-code-of-ethics-june2018.pdf).

Bernstein, Michael S., Margaret Levi, David Magnus, et al. 2021. "Ethics and society review: Ethics reflection as a precondition to research funding." *Proceedings of the National Academy of Sciences* 118(52):e2117261118. doi: 10.1073/pnas.2117261118.

Brubaker, Rogers. 2020. "Digital Hyperconnectivity and the Self." *Theory and Society* 49(5–6):771–801. doi: 10.1007/s11186-020-09405-1.

California Consumer Privacy Act, California Civil Code 1798.105 (2018) https://leginfo.legislature.ca.gov/faces/codes displayText.xhtml?division=3.&par=4.&lawCode =CIV&title=1.81.5

Carlini, Nicholas, Florian Tramèr, Eric Wallace, et al. 2021. "Extracting training data from large language models." Pp. 2633-2650 in *30th USENIX Security Symposium* (USENIX Security 21). USENIX Association.

Dawson, Ella. 2018. "We Are All Public Figures Now." Retrieved November 24, 2021 (https://elladawson.com/2018/07/08/we-are-all-public-figures-now/).

Eiermann, Martin. 2022. "American Privacy: Diffusion and Institutionalization of an Emerging Political Logic, 1870-1930." University of California, Berkeley. European Commission. 2016. Regulation 679 (General Data Protection Regulation) https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32016R0679

Fiesler Casey and Nicholas Proferes. 2018. 'Participant' Perceptions of Twitter Research Ethics. Social Media + Society 4(1): 1-14.

Fradkin, Andrey, Elena Grewal, and David Holtz. 2021. "Reciprocity and Unveiling in Two-sided Reputation Systems: Evidence from an Experiment on Airbnb." *Marketing Science* 40(6): 1013–1029.

Gatto, James G. and Pouneh Almasi. 2021. "Supreme Court Scraps LinkedIn Data-Scraping Decision." *Bloomberg Law*, 5 July. Retrieved November 21, 2021 (https://news.bloomberglaw.com/us-law-week/supreme-court scraps-linkedin-data-scraping-decision).

Gilbert, Sarah, Jessica Vitak, and Katie Shilton. 2021. "Measuring Americans' Comfort With Research Uses of Their Social Media Data." *Social Media + Society* 7(3). doi: 10.1177/20563051211033824.

Goffman, Erving. 1959. The Presentation of Self in Everyday Life. New York: Anchor Books.
Gregory, Charles O. 1951. "Gratuitous Undertakings and the Duty of Care." DePaul Law Review 1(1):30–68.

Grinberg, Nir, Kenneth Joseph, Lisa Friedland, Briony Swire-Thompson, and David Lazer. 2019. "Fake news on Twitter during the 2016 U.S. presidential election." Science. 363 (6425):374-378.

Hauge, Michelle V., Mark D. Stevenson, D. Kim Rossmo, et al. 2016. "Tagging Banksy: Using geographic profiling to investigate a modern art mystery." Journal of Spatial Science 61(1):185–190.

Jernite Y, Nguyen H, Biderman S et al. (2022) Data governance in the age of large-scale data-driven language technology. In: Proceedings of 2022 ACM Conference on Fairness, Accountability, and Transparency, Seoul, Republic of Korea, June 21–24 2022, pp. 2206–2222.

Jerolmack, Colin, and Alexandra K. Murphy. 2019. "The Ethical Dilemmas and Social Scientific Trade-Offs of Masking in Ethnography." *Sociological Methods & Research* 48(4):801–27.

Kekulluoglu, Dilara, Kami Vaniea, and Walid Magdy. 2022. "Understanding Privacy Switching Behaviour on Twitter." Pp. 1-14 in *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. New Orleans LA USA: ACM.

Kosinski, Michal, David Stillwell, and Thore Graepel. 2013. "Private traits and attributes are predictable from digital records of human behavior." *Proceedings of the National Academy of Sciences* 110(15): 5802–5805.

Kupferschmidt, Kai. 2023. "Twitter's plan to cut off free data access evokes 'fair amount of panic' among scientists." *Science*. https://www.science.org/content/article/twitters-plan-cut-free-data-access-evokes-fair-amount-panic-among-scientists

McCarthy, Charles R. 2008. "The origins and policies that govern institutional review boards." Pp. 541-551 in *The Oxford Textbook of Clinical Research Ethic*s, edited by E. Emanuel, C. Grady, R. Crouch, et al. New York, NY: Oxford University Press.

Mancosu, Moreno and Federico Vegetti. 2020. "What You Can Scrape and What Is Right to Scrape: A Proposal for a Tool to Collect Public Facebook Data." *Social Media + Society*. 6(3).

Marwick, Alice E. and danah boyd. 2011. "I tweet honestly, I tweet passionately: Twitter users, context collapse, and the imagined audience." *New Media & Society* 13(1): 114–133.

Metcalf, Jacob and Kate Crawford. 2016. "Where are human subjects in big data research? The emerging ethics divide." *Big Data & Society* 3(1): 2053951716650211.

Moore, Barrington. 1984. *Privacy: Studies in Social and Cultural History*. London: Routledge.

Nissenbaum, Helen. 2009. *Privacy in Context: Technology, Policy, and the Integrity of Social Life*. Stanford, Calif: Stanford Law Books.

Oversight Board. 2021. "Governance." https://about.fb.com/news/tag/oversight-board/

Rosenberg, Matthew, Nicholas Confessore, and Carole Cadwalladr. 2018. "How Trump Consultants Exploited the Facebook Data of Millions." *The New York Times*, 17 March. https://www.nytimes.com/2018/03/17/us/politics/cambridge-analytica-trump-campaign.html.
Schwartz, Barry. 1968. "The Social Psychology of Privacy." *American Journal of Sociology* 773(6):741–52.

van Dijck, José. 2013. "'You Have One Identity': Performing the Self on Facebook and LinkedIn." *Media, Culture & Society* 35(2):199–215. doi: 10.1177/0163443712468605.

Wilson, Samuel M., and Leighton C. Peterson. 2002. "The Anthropology of Online Communities." *Annual Review of Anthropology* 31(1):449–67. doi: 10.1146/annurev.anthro.31.040402.085436.

Zuboff, Shoshana. 2019. *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power.* New York: Public Affairs.

NATALIA NEDZHVETSKAYA is a Ph.D. candidate at the University of California, Berkeley in the Department of Sociology studying economic sociology, organizations, and technology. In addition, she holds a Designated Emphasis in the Sociology of Organizations and Markets from Haas School of Business. She uses mixed methods to study corporations as sites of broader changes in the economy, with a particular interest in understanding the tensions between shareholder and stakeholder capitalism. Her dissertation examines these themes further by studying workplace protests (employee activism) as an organizational phenomenon. Her research has been featured in The Guardian, WIRED, MIT Technology Review, NBC News, NPR, The LA Times, and TIME and has been funded by the Jain Family Institute, the Center for Technology, Society, and Policy, and the Berkeley Culture Initiative.

STEVEN LAUTERWASSER is a Postdoctoral Research Associate in the Department of Sociology and Anthropology at Northeastern University. He received his PhD in sociology from the University of California, Berkeley in 2022, where he studied how polarization in the US differs among partisans, not only as a matter of degree, but as a matter of kind. Drawing on diverse quantitative methods, he now works on the production of politicized knowledge more broadly, including the dissemination of feminist knowledge in academic organizations.