THE ROUTLEDGE HANDBOOK OF PRIVACY AND SOCIAL MEDIA

Edited by Sabine Trepte and Philipp K. Masur

First published 2023

ISBN: 978-1-032-11161-2 (hbk)
ISBN: 978-1-003-24467-7 (ebk)

30
THE ROLE OF PARTICIPANTS IN ONLINE PRIVACY RESEARCH

Ethical and Practical Considerations

Johannes Breuer[1,2], Katrin Weller[1,2], and Katharina Kinder-Kurlanda[3]

[1]GESIS – LEIBNIZ INSTITUTE FOR THE SOCIAL SCIENCES, COLOGNE, GERMANY
[2]CENTER FOR ADVANCED INTERNET STUDIES (CAIS), BOCHUM, GERMANY
[3]DIGITAL AGE RESEARCH CENTER, UNIVERSITY OF KLAGENFURT, AUSTRIA

DOI: 10.4324/9781003244677-35

The funder for this chapter is Center for Advanced Internet Studies (CAIS).

Routledge
Taylor & Francis Group
NEW YORK AND LONDON

30

THE ROLE OF PARTICIPANTS IN ONLINE PRIVACY RESEARCH

Ethical and Practical Considerations

Johannes Breuer[1,2]*, Katrin Weller*[1,2]*, and Katharina Kinder-Kurlanda*[3]

[1]GESIS – LEIBNIZ INSTITUTE FOR THE SOCIAL SCIENCES, COLOGNE, GERMANY
[2]CENTER FOR ADVANCED INTERNET STUDIES (CAIS), BOCHUM, GERMANY
[3]DIGITAL AGE RESEARCH CENTER, UNIVERSITY OF KLAGENFURT, AUSTRIA

Introduction

When using online platforms, we generate vast amounts of data. Platform providers, thus, often have access to detailed and personal information about their users and can employ this information for a variety of purposes. Popular platforms for networking and communicating, such as Facebook, Twitter, or YouTube, search engines, such as Google, or shopping portals, such as Amazon and eBay, require or request users to disclose many different kinds of personal information (see chapter 28 by Eichenhofer & Gusy on regulating privacy on online social networks). This disclosure is at the core of online privacy research and has led to the development of various perspectives on motivations and consequences of information disclosure. There are complex reasons and decisions involved when revealing information to internet platforms, and users often report wanting to protect their privacy, while at the same time being forced to disclose information to be able to participate (Lamla & Ochs, 2019; Willson & Kinder-Kurlanda, 2021). The (potential) discrepancy between attitudes and actual behavior regarding privacy has been described as the privacy paradox. Whether or in what form the privacy paradox exists and what it entails is an ongoing debate and a widely studied topic (see, e.g., Dienlin & Sun, 2021; Yu et al., 2020).

Researchers of online privacy often face a similarly paradoxical challenge in their research design: In order to study online privacy, they may collect personal or even sensitive information. Most research designs in online privacy research require participants to disclose information, such as personal attributes, beliefs and attitudes, their usage of digital technology, and other privacy-related behaviors. Much of this information can be sensitive – and may be identical to the information collected by online platforms. This creates conflicts for researchers in the field, who may find themselves facing the key question of how they can study online privacy in an ethical manner when their own studies intrude on people's privacy. Of course, not every study investigating privacy also invades participants' privacy, but the questions that most studies ask and the data they collect are typically personal and potentially sensitive. For that reason, it is important to consider the role of participants in online privacy research. Given the methodological

DOI: 10.4324/9781003244677-35

heterogeneity of studies in this field, "participation" can refer to very different settings from the perspective of those who are being asked to contribute to the study.

The role of study participants in the research process is a key ethical issue within the social and behavioral sciences (Chalmers et al., 1999). It has been a topic of research ethics for some time with a gradual change of terminology in research publications. Since the 1990s, in psychology and other disciplines (Boynton, 1998), there has been a shift from the term "(study) subjects" to "(research) participants" evidencing a move towards bias-free language to address power imbalances in the research process (American Psychological Association, 2020, p. 141). More descriptive terms, such as "college students," "children," or "respondents," are often preferred. A central question in these discussions is to what degree participation is voluntary and active in different research designs. People can be involved in research in various ways that can be more or less active. For example, focus group discussions or ethnographic studies typically require more active involvement of participants than surveys. In most research designs, voluntariness is ensured through the recruitment procedure in which informed consent is obtained from participants prior to data collection. Asking for informed consent establishes transparency about the purpose of the study and the way(s) in which the data will be used. Given that research on online privacy often is perceived as a sensitive endeavor, it is important for researchers in this field to pay special attention to questions related to informed consent. In practice, however, properly obtaining informed consent is not always straightforward and can be especially challenging for some research designs. In experimental studies, for example, participants often are not informed or even misled about the research purposes, to ensure that experimental treatments can work as intended. In such cases, participants need to be debriefed at the end. Ideally, they should also be able to decide whether their data should be used once they are informed about the true purpose of the study (or be given the opportunity to reconfirm or change a previous decision from before the debriefing). Obtaining informed consent can also be difficult when using particular types of data. The personal and often sensitive nature of the data in online privacy research, together with the general research trend of using new digital types of data and of combining different methods and data, requires that we reconsider how research can be performed in an ethical manner and what role(s) study participants (can) play in this.

In this chapter, we discuss the ethical challenges in online privacy research and how these may be addressed. Our main approach is to broaden the perspective on participants' involvement. Doing so allows us to explore various facets of research ethics connected to different ways of studying online privacy. More specifically, we illustrate different paths that may help researchers of online privacy to re-think research contexts and to find innovative approaches, including using new types of data sources and new modes of participation.

From Online Privacy to Research Ethics

In online privacy research, "privacy" is both the object of study and an ethical issue to be considered. Part of the perceived paradox as outlined above may be due to the conflation of these two functions. In a first step, to untangle this, we will try to broaden the perspective by illustrating how privacy is embedded into research ethics. Importantly, research ethics are not a binary concept. While we may say that something is ethical or unethical, in reality, ethical questions and their answers exist on a continuum. Also, when it comes to research ethics, different values and goals may be in conflict with each other, requiring a careful weighing of alternative choices, e.g., with regard to collecting, processing, and sharing research data. Researchers need to consider risks to participants in various stages of their study, starting with the planning phase. Ethical research design includes reflecting on standards and practices intended to ensure research integrity, meaningful results, and avoidance of misrepresentation. It also requires to continuously reconsider

collaboration practices and how to enable responsible reuse of results and data. Adapting to changes in the topic that is being studied to the continuous development of digital media platforms and other technology as well as to the changing landscape of online data and digital methods requires innovative concepts. In the following, we want to discuss some key areas where such innovative concepts may help to address ethical questions in online privacy research: (1) research design and data, (2) participation and informed consent, and (3) data privacy and transparency.

Research Design and Data

The majority of studies on online privacy are based on self-report measures (see, e.g., Gerber et al., 2018; Yu et al., 2020). These are, for example, employed in interviews, surveys, or experimental designs to assess various kinds of privacy behavior, beliefs, attitudes, and types of media use. However, self-report measures have repeatedly been shown to be unreliable as they can be influenced by social desirability or problems with recollection (Parry et al., 2021). One option for measuring media use as well as certain privacy-related behaviors is the use of so-called digital trace data, i.e., "records of activity (…) undertaken through an online information system" (Howison et al., 2011, p. 769). Digital traces are increasingly used as research data in various disciplines and are at the heart of new research areas, such as computational social science or computational communication science. Researchers see various valuable characteristics in these kinds of data (Kinder-Kurlanda & Weller, 2014), such as the potential to assess immediate reactions to events. Digital traces are often generated as a byproduct of daily activities, and are not produced in response to a specific research study design, thus allowing a glimpse into otherwise hidden everyday practices.

For privacy research, the value of digital trace data can be further increased when they are combined with other types of data, e.g., from interviews or surveys (Stier et al., 2020). A combination of digital trace data with self-report data can be particularly interesting for investigating the already mentioned privacy paradox. Given the issues of social desirability and recall errors, there is a considerable risk that people may – intentionally or unintentionally – misreport their engagement in privacy behaviors. Using digital trace data in combination with self-report data can help to uncover and understand such potential biases.

While digital trace can help to find innovative research designs, also in online privacy research, they fundamentally challenge research ethics. Specific challenges in data collection, processing, and publication include questions related to informed consent or perceptions of what are public and what are private spaces of communication and interaction. Another issue is the tradeoff between data protection and privacy concerns on the one hand and transparency and openness of the research methods and data on the other hand. For example, it is often difficult to decide how digital trace data should be aggregated or otherwise processed and reduced before they are shared with other researchers. Currently, only a small, albeit growing, number of resources exists as guidance on ethics for researchers using digital trace data (franzke et al., 2020; Zimmer & Kinder-Kurlanda, 2017). The lack of specific guidance is in large part due to the variety of data summarized under the term "digital traces." Ethical challenges largely depend on the type of data as well as the ways in which they are collected. As the aim usually is to capture digital traces as they occur in everyday situations, doing so implies that those who leave the traces in digital environments are unaware of their data being collected. This situation diverges from the concept of study participants, as people are not consciously participating in a specific study, and also seems to render traditional standards for obtaining informed consent almost impossible.

Fiesler and Proferes (2018) have shown that Twitter users are typically not aware of the possibility that their data might be accessed by researchers and would often only support specific research settings. Hence, it is legitimate to ask whether it is appropriate to speak of "participants" in cases of digital trace data, as the individuals whose data are being collected are in many (or even most) cases

not aware of this and never volunteered to participate in a research study. Most authors of respective research papers, indeed, do not apply the term participants but instead use "platform users" (Halavais, 2019) or similar phrases. While this term may be technically more accurate, thinking of platform users as a specific type of (unconscious or even involuntary) participants may help to shift the focus back to the challenge of successfully finding answers to questions related to research ethics in these new research designs. New ethical research practices are needed in general for new data types, but, as a starting point, it is necessary to look at the concepts of participation and informed consent more closely.

Concepts for Participation and Informed Consent

As stated before, when researchers collect digital trace data (using APIs or other automated approaches like web scraping) it can be very difficult or even impossible to obtain informed consent from the people whose data are being used. Asking for consent may be feasible for small samples if researchers have ways of contacting those whose data are collected (e.g., if a platform allows sending direct messages), but becomes infeasible with large samples or when researchers cannot directly contact all individuals whose data are being collected. For example, on Twitter, the default setting is that users need to follow each other to be able to send direct messages. Another advantage of combining digital trace data with data from interviews or surveys is that the latter can be used to also obtain informed consent for collecting/using people's digital trace data (Breuer et al., 2021). However, self-selection into a study is one of several potential sources of bias in the collection of digital trace data (Sen et al., 2021).

Digital trace data may be accessed for research purposes in different ways (Breuer et al., 2020). A commonly used method is the collection of data via Application Programming Interfaces (APIs) offered by platforms. While some major platforms have substantially reduced or essentially closed off data access via their APIs, most notably Facebook and Instagram, others, such as Twitter, have been offering researchers access to a wide array of data that is also relevant for privacy research. Importantly, both the technical limitations of the APIs as well as the Terms of Services for how they may be used may change over time. This has led to ongoing investigations of data quality and challenges regarding the representativeness and reproducibility of digital-trace-data-based research (see, e.g., Olteanu et al., 2016). It has also sparked discussions about data access and alternatives to APIs for accessing platform data (Bruns, 2019; Freelon, 2018; Halavais, 2019; Puschmann, 2019). These also entail different models of participation and user involvement. In the following, we will introduce three possible approaches to re-thinking participation in the context of online privacy research that makes use of digital trace data: (1) data donation, (2) data exploration and citizen science, and (3) debriefing and opting out.

Data Donation

Studies that combine surveys with digital trace data often ask participants to share user names or links to social media profiles. A related but more involved approach is to directly ask platform users to provide their full data history from a specific platform. This method for accessing digital trace data by collaborating with platform users is currently gaining momentum and is commonly described as data donation (Araujo et al., 2022; Boeschoten et al., 2022). Many platforms and services that are relevant for research on (online) privacy, such as social media platforms or fitness tracking devices, offer their users functionalities for exporting their own data which they can then share with researchers.

Obviously, this method, again, poses questions related to privacy that researchers need to address. For example, solutions may need to be found for secure data upload, anonymization, or

pseudonymization. Digital trace data shared via data donation may also contain information about other – unaware – individuals, such as social media contacts of the donating person, or friends whom they mentioned or tagged in a post or comment. Nevertheless, data donation is a promising approach that offers advantages and opportunities especially interesting for privacy research, particularly when a user's exact privacy settings can be considered in the research design, in addition to the usage data.

Data donations give back to platform users the status of participants who consent to being part of a research study. Through data donation, they can be more actively involved in the research. This begins with the act of downloading and sharing (donating) their data which requires more involvement than, for example, providing a user name. While donation increases participant burden, it also increases transparency, as participants can explore their data before sharing, and may use the insights to decide what parts of these data to donate. It enables participants to potentially act on an eye-to-eye level with researchers with respect to data awareness. Providing users with insights into their own data can be made use of in different ways for the research project itself; from participating by improving data quality, to having a role in generating results from data, or gaining new insights into privacy risks. As these examples show, data donation and options for exploring the donated data can facilitate new forms of participant involvement.

Data Exploration and Citizen Science

Studies using interventions, such as prompts and nudges (see, e.g., Ioannou et al., 2021; Schäwel, 2019; also see chapter 25 by Wang & Acquisiti on nudges for privacy and security), are common in the field of privacy research. Offering participants the opportunity to explore their own digital trace data – as described in the previous section – can be a powerful intervention to make users aware of potential privacy risks. It can also improve participants' data and privacy literacy (see chapter 11 by Hagendorff, Masur & Trepte on privacy literacy). Donating digital trace data allows an even more active involvement. By exploring their data, participants can contribute to identifying and answering research questions and, thus, engage in so-called citizen science (Majumder & McGuire, 2020), which is already common in the natural sciences but not (yet) in the social and behavioral sciences. Citizen science approaches enable that "people should participate in, not be subjects of, research" as laid out by the Standing Advisory Group on Consumer Involvement in the NHS Research and Development Programme (cited by Boynton, 1998). The cited recommendation is that consumers are involved "not as 'subjects' of research, but as active participants in the process of deciding what research should take place, commissioning research, interpreting the results, and disseminating the findings" (Boynton, 1998, p. 1521). As many research projects in online privacy research follow the purpose of identifying means for making users of digital technology responsible and empowered consumers, such a vision is especially attractive for this research area. And while their use also poses privacy challenges (as we will discuss in the following section), digital trace data, especially if obtained through data donation, can help achieve such empowerment.

Debriefing and Opting Out

In most research based on digital trace data, platform users are unaware that their data has been included in a research data collection. A strategy to raise awareness in this context is to set up infrastructures for "debriefing." Debriefing can be adapted for digital trace data, so that people are notified that their data has been included in a collection, and are offered the chance to "opt out." Currently, this is rarely applied in practice – and also viewed as impractical by the research

community (Vitak et al., 2017). The lack of debriefing has been criticized in prominent cases, such as in the reflections by Grimmelmann (2015) on the "emotional contagion" study on Facebook (Kramer et al., 2014). An explorative solution is the open-source system "Bartleby" (Zong & Matias, 2022) that supports notifying platform users who have become "participants" of digital trace studies. Such solutions for debriefing and opting out are especially relevant for social media privacy research, given the sensitivity of its topics and its aim of increasing privacy literacy.

Concepts for Data Privacy and Transparency

Two general directives for research ethics are avoiding harm and maximizing benefits for participants, scholarship, and society. Increasing data privacy is especially important, but on the other hand, the principle of maximizing benefits also entails that data should be used as effectively as possible. Sharing research data so that they can be reused by others increases their value and, thus, maximizes the benefits of the respective research. Accordingly, there always is a tradeoff between (maximizing) privacy (of participants) and transparency (of the research) that researchers need to deal with. For example, if data are reduced or aggregated, they lose some of their reuse value. While this is true for all areas of research that involve data from humans, the importance of this tradeoff is particularly pronounced for online privacy research, given its subject and aims.

Measures for Protecting Data Privacy

Three common ways of increasing data privacy are anonymization, pseudonymization, and data reduction. Anonymization means that the data are processed in such a way that the identification of individuals becomes impossible. This can, e.g., be achieved through aggregation. Pseudonymization describes the process of removing direct identifiers, such as names, and potentially problematic cases or combinations of indirect identifiers from the data with the goal of ensuring that data can only be attributed to individuals through the use of additional data. Identification, thus, does not become impossible but requires considerable effort and depends on the availability of additional (linkable) data. Data reduction is a more general principle that can be part of anonymization as well as pseudonymization. Put simply, data reduction means that specific parts of the data are removed to improve data privacy. These can be variables or cases as well as certain values, for example, turning a numeric variable into a categorical variable.

Applying anonymization or pseudonymization strategies to digital trace data poses new challenges. Even after removing personal information, it may be possible to re-identify someone if additional context information is explored. As Zimmer (2010) has illustrated for a de-identified Facebook dataset, it is not only relevant whether the digital trace data originates from publicly accessible platforms. Instead, it is always important to identify effective measures for protecting privacy across contexts. This leads to the key challenge of assuring privacy when sharing data. Developing better solutions for data access and secondary data use that build on secure data environments and external safeguards can contribute to research transparency while also protecting privacy according to legal and ethical standards.

Data Sharing

Finding a good balance between privacy and transparency can be challenging. Privacy considerations but also practical aspects, such as the size and format of the data, require novel approaches for archiving and publishing them (Van Atteveldt et al., 2020) and many data archives are currently working on developing and implementing these (see, e.g., Breuer et al., 2021; Hemphill, 2019).

Despite the challenges, researchers working with digital traces appear to be principally interested and willing to share research data (Weller & Kinder-Kurlanda, 2017), also in order to increase reproducibility and to make research results more transparent for the scientific community. Archiving and sharing this kind of research data might also increase awareness about the underlying research in the general public, thereby enabling a more solid foundation for data donation approaches and consent forms. Some researchers also feel they need to give back something to the user community of a platform they have been studying and see this as a key motivation for sharing research data (Weller & Kinder-Kurlanda, 2017). Given the lack of standardized archiving solutions, individual initiatives from researchers may be far less effective in contributing to transparency. Future work is required to establish standards for archiving and sharing new types of research data (Weller & Kinder-Kurlanda, 2016). Access to archived datasets may also contribute to the principle of sparsity, i.e., avoiding having to collect the same kind of information multiple times, thus, reducing the number of people that are asked to disclose private information for research purposes.

Summary and Future Directions

Studying online privacy is a research field that in its very nature is connected to sensitive information about the people being studied. Hence, studies in this area need to address various ethical questions in all phases of the research process. Many of those are the same as for other research areas. For example, researchers should always obtain informed consent, if this is possible, and participants in experimental studies should be debriefed if the studies involve experimental deception. While these aspects relate to the collection of the data, other ethical obligations concern the processing and sharing of the data. A key maxim in this is the protection of participants' privacy. Hence, data minimization should be a guiding principle. If the data contain personal information and especially if they are potentially sensitive, researchers should apply anonymization, pseudonymization, or data reduction before making the data available to others. On the other hand, the ideals of open science are also relevant for research ethics considerations. Data sharing not only increases transparency, reproducibility, and replicability, but also the effectiveness of research and the value of the data. This can also benefit study participants as it increases the value of their data and can reduce the risk of unnecessary repeated data collections. Accordingly, if possible, research data from online privacy research should be made available via suitable data archives (with appropriate data protection measures in place). Notably, however, when researchers in only privacy research want to employ innovative research designs and make use of digital trace data, there are a few additional ethical dimensions to be considered. These are summarized in Figure 30.1.

With regard to future directions, in addition to enabling a more active involvement of study participants, one major potential of digital trace data for online privacy research is that they can also be used in intervention studies by allowing participants to explore and learn about the data they generate. When it comes to processing and sharing data in online privacy research, especially when using digital trace data, innovative solutions have to be employed that strike a balance between the privacy of participants and the transparency of the research. The development of such solutions, as well as the implementation of new ways of involving study participants, can not only help to find and answer novel questions in online privacy research but may also improve transparency and foster critical reflection of privacy and research ethics more generally. Innovation in research designs, participation, and transparency can, thus, help resolve the paradox of online privacy research that makes use of personal, private, and often also sensitive data to study what defines and affects people's online privacy experiences.

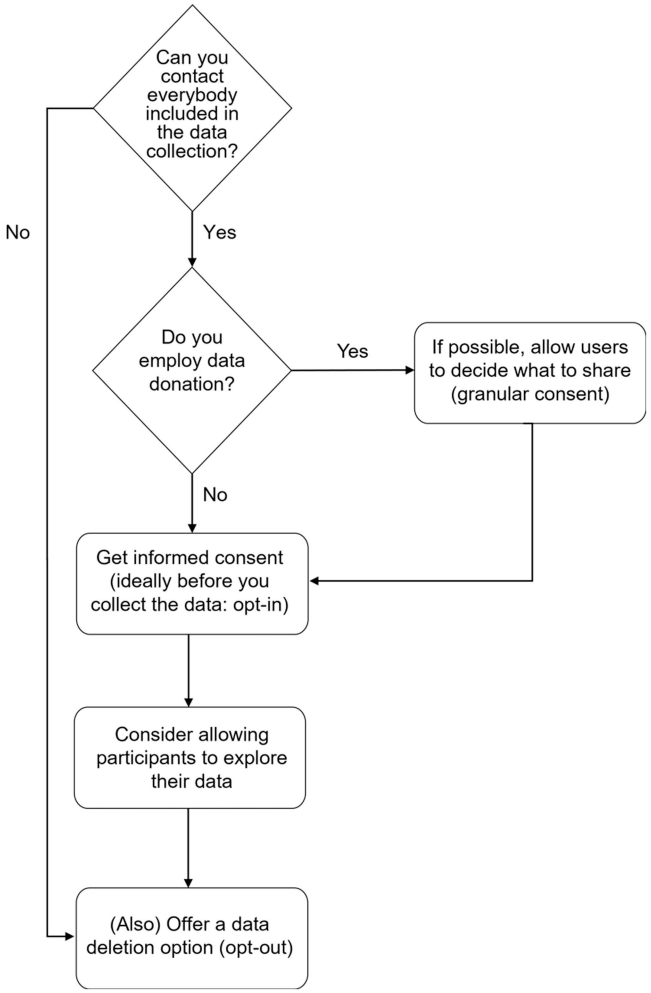

Figure 30.1 Flowchart Illustrating Ethical Considerations to Be Made in Online Privacy Research Employing Digital Trace Data

References

American Psychological Association. (2020). *Publication manual of the American Psychological Association (7th ed.)*. American Psychological Association. https://doi.org/10.1037/0000165-000

Araujo, T., Ausloos, J., van Atteveldt, W., Loecherbach, F., Moeller, J., Ohme, J., Trilling, D., van de Velde, B., de Vreese, C., & Welbers, K. (2022). OSD2F: An Open-Source Data Donation Framework. Computational Communication Research, 4(2) 372–387. https://doi.org/10.5117/CCR2022.2.001.ARAU

Boeschoten, L., Ausloos, J., Moeller, J., Araujo, T., & Oberski, D. L. (2020). Digital trace data collection through data donation. ArXiv:2011.09851 [Cs, Stat]. http://arxiv.org/abs/2011.09851

Boynton, P. M. (1998). People should participate in, not be subjects of, research. *BMJ*, *317*(7171), 1521–1521. https://doi.org/10.1136/bmj.317.7171.1521a

Breuer, J., Al Baghal, T., Sloan, L., Bishop, L., Kondyli, D., & Linardis, A. (2021). Informed consent for linking survey and social media data – Differences between platforms and data types. *IASSIST Quarterly*, *45*(1), 1–27. https://doi.org/10.29173/iq988

Breuer, J., Bishop, L., & Kinder-Kurlanda, K. (2020). The practical and ethical challenges in acquiring and sharing digital trace data: Negotiating public-private partnerships. *New Media & Society*, *22*(11), 2058–2080. https://doi.org/10.1177/1461444820924622

Breuer, J., Borschewski, K., Bishop, L., Vávra, M., Štebe, J., Strapcova, K., & Hegedűs, P. (2021). Archiving social media data: A guide for archivists and researchers. https://doi.org/10.5281/ZENODO.5041072

Bruns, A. (2019). After the 'APIcalypse': Social media platforms and their fight against critical scholarly research. *Information, Communication & Society*, *22*(11), 1544–1566. https://doi.org/10.1080/1369118x.2019.1637447

Chalmers, I., Jackson, W., & Carvel, D. (1999). People are "participants" in research. *BMJ*, *318*(7191), 1141–1141. https://doi.org/10.1136/bmj.318.7191.1141a

Dienlin, T., & Sun, Y. (2021). Does the privacy paradox exist? Comment on Yu et al.'s (2020) meta-analysis. *Meta-Psychology*, *5*. https://doi.org/10.15626/mp.2020.2711

Fiesler, C., & Proferes, N. (2018). "Participant" of Twitter research. *Social Media + Society*, *4*(1), Online publication. https://doi.org/10.1177/2056305118763366

franzke, a. s., Bechmann, A., Zimmer, M., Ess, C. & the Association of Internet Researchers (2020). Internet Research: Ethical Guidelines 3.0. https://aoir.org/reports/ethics3.pdf

Freelon, D. (2018). Computational research in the post-API age. *Political Communication*, *35*(4), 665–668. https://doi.org/10.1080/10584609.2018.1477506

Gerber, N., Gerber, P., & Volkamer, M. (2018). Explaining the privacy paradox: A systematic review of literature investigating privacy attitude and behavior. *Computers & Security*, 77, 226–261. https://doi.org/10.1016/j.cose.2018.04.002

Grimmelmann, J. (2015). The law and ethics of experiments on social media users. *Colorado Technology Law Journal*, *13*, Article 219. https://osf.io/cdt7y

Halavais, A. (2019). Overcoming terms of service: A proposal for ethical distributed research. *Information, Communication & Society*, *22*(11), 1567–1581. https://doi.org/10.1080/1369118X.2019.1627386

Hemphill, L. (2019). Updates on ICPSR's Social Media Archive (SOMAR). https://doi.org/10.5281/ZENODO.3612677

Howison, J., Wiggins, A., & Crowston, K. (2011). Validity issues in the use of social network analysis with digital trace data. *Journal of the Association for Information Systems*, *12*(12), 767–797. https://doi.org/10.17705/1jais.00282

Ioannou, A., Tussyadiah, I., Miller, G., Li, S., & Weick, M. (2021). Privacy nudges for disclosure of personal information: A systematic literature review and meta-analysis. *PLOS ONE*, *16*(8), e0256822. https://doi.org/10.1371/journal.pone.0256822

Kinder-Kurlanda, K., & Weller, K. (2014). "I always feel it must be great to be a hacker!": The role of interdisciplinary work in social media research. *Proceedings of the 2014 ACM conference on web science*, 91–98. https://doi.org/10.1145/2615569.2615685

Kramer, A. D. I., Guillory, J. E., & Hancock, J. T. (2014). Experimental evidence of massive-scale emotional contagion through social networks. *Proceedings of the National Academy of Sciences*, *111*(24), 8788–8790. https://doi.org/10.1073/pnas.1320040111

Lamla, J., & Ochs, C. (2019). Selbstbestimmungspraktiken in der Datenökonomie: Gesellschaftlicher Widerspruch oder 'privates' Paradox? In B. Blättel-Mink, & P. Kenning (Eds.), *Paradoxien des Verbraucherverhaltens. Dokumentation der Jahreskonferenz 2017 des Netzwerks Verbraucherforschung* (pp. 25–39). Springer Gabler: 25-39.

Majumder, M. A., & McGuire, A. L. (2020). Data sharing in the context of health-related citizen science. *Journal of Law, Medicine & Ethics*, *48*(S1), 167–177. https://doi.org/10.1177/1073110520917044

Nissenbaum, H. (2009). *Privacy in context*. Stanford University Press.

Olteanu, A., Castillo, C., Diaz, F., & Kiciman, E. (2016). Social data: Biases, methodological pitfalls, and ethical boundaries. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.2886526

Parry, D. A., Davidson, B. I., Sewall, C. J. R., Fisher, J. T., Mieczkowski, H., & Quintana, D. S. (2021). A systematic review and meta-analysis of discrepancies between logged and self-reported digital media use. *Nature Human Behaviour*. https://doi.org/10.1038/s41562-021-01117-5

Puschmann, C. (2019). An end to the wild west of social media research: A response to Axel Bruns. *Information, Communication & Society*, *22*(11), 1582–1589. https://doi.org/10.1080/1369118X.2019.1646300

Schäwel, J. (2019). How to Raise Users' Awareness of Online Privacy: An Empirical and Theoretical Approach for Examining the Impact of Persuasive Privacy Support Measures on Users' Self-Disclosure on Online Social Networking Sites [Ph.D. thesis, University of Duisburg-Essen]. https://duepublico2.uni-due.de/receive/duepublico_mods_00070691

Sen, I., Flöck, F., Weller, K., Weiß, B., & Wagner, C. (2021). A total error framework for digital traces of human behavior on online platforms. *Public Opinion Quarterly*, *85*(S1), 399–422. https://doi.org/10.1093/poq/nfab018

Stier, S., Breuer, J., Siegers, P., & Thorson, K. (2020). Integrating survey data and digital trace data: Key issues in developing an emerging field. *Social Science Computer Review*, *38*(5), 503–516. https://doi.org/10.1177/0894439319843669

Van Atteveldt, W., Althaus, S., & Wessler, H. (2020). The trouble with sharing your privates: Pursuing ethical open science and collaborative research across national jurisdictions using sensitive data. *Political Communication*, 1–7. https://doi.org/10.1080/10584609.2020.1744780

Vitak, J., Proferes, N., Shilton, K., & Ashktorab, Z. (2017). Ethics regulation in social computing research: Examining the role of Institutional Review Boards. *Journal of Empirical Research on Human Research Ethics*, *12*(5), 372–382. https://doi.org/10.1177/1556264617725200

Weller, K., & Kinder-Kurlanda, K. E. (2016). A manifesto for data sharing in social media research. Proceedings of the 8th ACM Conference on Web Science. 166–172. https://doi.org/10.1145/2908131.2908172

Weller, K., & Kinder-Kurlanda, K. (2017). To share or not to share? Ethical challenges in sharing social media-based research data. In M. Zimmer & K. E. Kinder-Kurlanda (Eds.), *Internet Research Ethics for the Social Age* (pp. 115–129). Peter Lang.

Willson, M., & Kinder-Kurlanda, K. (2021). Social gamers' everyday (in)visibility tactics: Playing within programmed constraints. *Information, Communication & Society*, *24*(1), 134–149. https://doi.org/10.1080/1369118X.2019.1635187

Yu, L., Li, H., He, W., Wang, F.-K., & Jiao, S. (2020). A meta-analysis to explore privacy cognition and information disclosure of internet users. *International Journal of Information Management*, *51*, 102015. https://doi.org/10.1016/j.ijinfomgt.2019.09.011

Zimmer, M. (2010). "But the data is already public": On the ethics of research in Facebook. *Ethics and Information Technology*, *12*(4), 313–325. https://doi.org/10.1007/s10676-010-9227-5

Zimmer, M., & Kinder-Kurlanda, K. (Eds.). (2017). *Internet research ethics for the social age: New challenges, cases, and contexts*. Peter Lang.

Zong, J., & Matias, J. N. (2022). Bartleby: Procedural and substantive ethics in the design of research ethics systems. *Social Media + Society*, *8*(1), Online publication. https://doi.org/10.1177/20563051221077021