Considerations for Using Social Media Data in Learning Design and Technology Research

Spencer P. Greenhalgh
University of Kentucky
spencer.greenhalgh@uky.edu

Matthew J. Koehler
Michigan State University
mkoehler@msu.edu

Joshua M. Rosenberg
University of Tennessee, Knoxville
jmrosenberg@utk.edu

K. Bret Staudt Willet
Michigan State University
staudtwi@msu.edu

Abstract

Social media platforms have firmly established themselves as phenomena of interest for Learning Design and Technology (LDT) researchers, and the data accessible from these platforms provide new methodological possibilities. Given the number, diversity, and constant evolution of both social media platforms and research tools, it would be inappropriate to suggest that there is a *single correct way* to carry out LDT research with social media data. In contrast, this chapter introduces six broad, interconnected steps to social media research: Conducting Ethical Research; Framing the Research; Organizing the Research Process; Collecting Data; Analyzing Data; and Writing, Sharing, and Publicizing Research. Each step is associated with several considerations, careful attention to which will guide researchers toward a *particular correct way* to accomplish their specific research objectives.

Introduction

Since their emergence in the early-to-mid 2000s, social media platforms have quickly established themselves as phenomena of interest for Learning Design and Technology (LDT) researchers. Indeed, in Van Osch and Coursaris's 2015 review of 610 articles on social media published between 2004 and 2011, they found that 13.9% were focused on *Education and Learning*, more than on any other topic. In addition to providing new phenomena for educational researchers to study, the emergence of new technologies also offers new ways for researchers to collect and analyze data (Mishra, Koehler, & Greenhow, 2016). In short, when using social media platforms for teaching and learning, "people leave traces of their identity, their actions, and their social relations" (Welser, Smith, Fisher, & Gleave, 2008, p. 116). This means that social media data can often be unobtrusively collected as an authentic representation of educational practices within these platforms. Our purpose in this chapter is, therefore, to suggest considerations for the use of social media data in LDT research. These considerations are intended to help LDT professionals use social media data to produce research that is rigorous, rich, and—importantly—ethical.

We also recognize limits to the focus and approach we have adopted in this chapter. First, it is entirely possible to study social media *phenomena* without using social media *data*. However, we do not consider these approaches in this chapter, we welcome (and regularly cite) them, as they make contributions that analyzing these data cannot. Second, we refrain from providing a step-by-step technical overview of how projects using social media data could or should be completed. Social media platforms and research tools are numerous, diverse, and continuously changing—even with considerably more space, we would struggle to provide enough detail and ensure enough longevity to make a technical walkthrough useful. Third, although the considerations we suggest are applicable to social media data understood broadly, we are most familiar with Twitter data, which will be evident throughout the chapter. This overrepresentation of Twitter is characteristic of all social media research (Tufekci, 2014), but we nonetheless emphasize that our suggestions should be considered in the context of the specific platform being studied. Finally, as our chat logs and email conversations would attest, there is considerable room for debate and disagreement around the proper way to conduct this research—our goal is not to describe *the right way* to work with social media data so much as it is to scaffold researchers' efforts to find *a proper way* for their context. In short, we avoid giving specific *answers* in this chapter, preferring instead to draw LDT researchers' attention to *questions* they should be asking as they use social media data in their research.

We have organized our discussion of these considerations into six broad steps:

- Conducting Ethical Research;
- Framing the Research;
- Organizing the Research Process;
- Collecting Data;
- Analyzing Data; and
- Writing, Sharing, and Publicizing Research.

Although we present these steps as distinct and in approximate chronological order, it is highly unlikely (and even inadvisable) that researchers would complete them in strict order or as independent decisions. Indeed, we present Conducting Ethical Research as a "first" step only because it should be considered throughout the entire research process. Furthermore, we have regularly collected interesting data before formally framing or organizing a research project (when doing so has not conflicted with ethical considerations). Finally, as Peng and Matsui

(2017) have noted in the context of data science, research is an iterative process, with each step potentially inviting a re-evaluation of the previous ones.

LDT research involving social media data is characterized by a mix of innovation and established tradition. None of the steps listed above is unique to this kind of research, but the considerations we list under each step are likely new—or of additional importance—for scholars exploring the use of social media data in LDT research. On a similar note, readers will notice that we avoid traditional distinctions between "quantitative" and "qualitative" research in this chapter. Methodological developments with obvious applications to social media research—such as computational text analysis and ethnographic approaches to digital trace data (see Selwyn, 2019)—have blurred the distinctions usually associated with these terms. We, therefore, prefer to describe a given study using: (a) the four considerations described in the Framing the Research section, (b) the distinction made between "obtrusive" and "unobtrusive" methods in the Collecting Data section, and (c) the distinction made between "human" and "machine" methods in the Analyzing Data section.

Conducting Ethical Research

Social media data challenge traditional understandings of "human subject research" and the "common rule" (U.S. Department of Health and Human Services, 2018) that have traditionally guided researchers and institutional oversight in the United States. Thus, depending on their institution and their research design, LDT researchers working with social media data may experience different levels of oversight or receive conflicting guidance (see Vitak, Proferes, Shilton, & Ashktorab, 2017).

Even if an institutional body does not consider an analysis of social media data to be "human subjects research," social media data constitute a human phenomenon and thus demand ethical reflection. We argue, therefore, that regardless of oversight, the primary responsibility for the ethical study of social media data lies with the researchers themselves. Furthermore, we join a growing consensus of researchers in suggesting that thorough consideration of ethical principles situated in a specific research phase and context—rather than a one-size-fits-all "checklist" approach—is critical to ensure responsible and ethical digital research (e.g., franzke, Bechmann, Zimmer, Ess, & the Association of Internet Researchers, 2020). Indeed, here we reiterate the emphasis from franzke et al. (2020) and Krutka, Heath, and Staudt Willet (in press) that ethical considerations begin with asking and answering critical questions. In this section, we draw from efforts by the Association of Internet Researchers (franzke et al., 2020); Fiesler and Proferes (2018); Golder, Ahmed, Norman, and Booth (2017); Moreno, Goniu, Moreno, and Diekema (2013); Sloan, Quan-Haase, Kitchin, and Beninger (2017); Suomela, Chee, Berendt, and Rockwell (2019); and Townsend and Wallace (2016) to identify the following, selected ethical principles that should influence LDT researchers' decisions of what data to collect, how to analyze it, and how to present findings:

Public vs. Private: Some social media platforms make users' data fully public, whereas others impose privacy restrictions. However, Proferes's (2017) survey of Twitter users' beliefs about how the platform works demonstrated that many social media users do not fully understand—or truly consent to—who can see their posts.

Harms and Benefits: A researcher must identify potential harms and benefits and adapt research strategies to reduce harms—including those resulting from unexpected publicity. For example, LDT researchers may be interested in social media's role in teacher activism (e.g., Krutka, Asino, & Haselwood, 2018). However, identifying teachers who participate in online activism in research on the subject could indirectly lead to professional or political reprisals against those teachers.

Vulnerability: Considerations of harms and benefits must be made with the understanding that some individuals and groups are especially vulnerable in social media spaces, mainly when research focuses on sensitive topics.

Anonymity: Researchers may try to avoid harm by "anonymizing" data (e.g., quoting a post without identifying its author). However, "anonymized" data often retain enough information to identify individuals—for example, searching the Internet for the text from a public social media post reproduced "anonymously" in research may allow someone to identify the author.

Consent: Researchers must often decide the extent to which participants should provide consent to access social media contexts; collect data from those contexts; and store, share, or publish excerpts from collected data. Proferes and Walker (2020) summarize some of the diverse opinions and practices of social media researchers in a helpful paper focused on this topic.

Legal Considerations: Local, national, and international laws concerning privacy, copyright, and conduct of research—as well as the terms of service agreements for social media platforms (cf. SIGCHI Ethics Committee, 2017)—may impact the collection, storage, and sharing of data.

Framing the Research

Social media data present new opportunities for LDT researchers, but these data do not speak for themselves. Rather, "interpretation [remains] at the center of data analysis" (boyd & Crawford, 2012, p. 668). We therefore suggest that researchers remain attentive to the following considerations that may influence their study and interpretation of social media data:

Paradigms and Assumptions: LDT research is characterized by a plurality of paradigms, including "positivism, interpretivism, critical theory, feminism, post-modernism, and design-based research" (Kimmons & Johnstun, 2019, p. 2). Each represents a different set of accepted understandings, methods, values, and expectations for research. Social media research in LDT can adopt any of these paradigms, and researchers benefit from explicitly examining, exploring, and declaring their paradigmatic assumptions.

Research Design, Methods, and Modes of Inquiry: Researchers must consider both research design (overall structure) and methods (processes, procedures, tools, and analyses). Together, these form *modes of inquiry* that "employ different standards of evidence for ... the validity and reliability of claims, and make different kinds of value commitments" (Penuel & Frank, 2016, p. 16). Social media data can be approached using any established mode of inquiry (e.g., Creswell, 2014; Remler & Van Ryzin, 2015), from a randomized control trial investigating educational outcomes of social media use to an ethnography of a social media learning community. We recommend Sloan and Quan-Haase's (2017) *Handbook of Social Media Research Methods* for a detailed overview of issues and innovations in research design, methods, and modes of inquiry in social media research.

Conceptual Frameworks: Researchers should guide their inquiry with a *conceptual framework* that indicates "what to pay attention to, what difficulties to expect, what questions to ask, and how to approach problems" (Wenger, 1998, p. 9). Ngai, Tao, & Moon (2015) found that 30 different conceptual framings, theories, or models were used in 46 studies of social media. There are also many frameworks unique to education or LDT research that may also be helpful, and the choice of a conceptual framework is therefore not easy—but remains important. Indeed, Staudt Willet, Koehler, and Greenhalgh (2017) demonstrated that researchers employing different conceptual frameworks may be interested in different elements of the same social media dataset and may interpret that data in different ways.

Phenomena and Units of Analysis: A single collection of social media data may include all of the following features (and perhaps more): (a) content of posts, (b) characteristics of posts, (c) characteristics of individuals, (d) activity of individuals, (e) interactions between individuals, (f) networks or communities formed by groups, and (g) connections between groups. Within a given study, LDT researchers may be interested in just one of these *phenomena* or may wish to consider several of them; however, it is unlikely that they will be able to consider all of them within a single research project. Thus, LDT researchers must necessarily delineate one or more target *phenomena*—this delineation will indicate one or more *units of analysis,* each of which has implications for research design, analysis, and conceptual framing. Researchers' choice of a phenomenon may be guided by established theory and research or inspired by new conventions and phenomena that have organically emerged from social media contexts (e.g., boyd, Golder, & Lotan, 2010).

Organizing the Research Process

The novel affordances of social media data and innovations in social media research methods reinforce the importance for LDT researchers in carefully organizing the research process.

Software: Most social media research requires—or is aided by—the use of specialized software for collecting or analyzing data. We suggest evaluating software in terms of *control*, *simplicity*, *cost*, and *training*. As a general (but far from absolute) rule, the more *control* a tool allows over the collection and analysis of data, the less *simple* it is to use—these features tend to exist in a reciprocal tension. Because the tools offering the finest control are often available at low, or no, *cost* (programming languages such as R, Python, and PHP can be used for free), we generally recommend these approaches as an ideal to strive for. However, we recognize that they also require more *training* not often provided to LDT scholars (Kimmons & Veletsianos, 2018) and therefore also recommend considering the use of specialized software like the researcher-focused *Netlytic* (https://netlytic.org; see also Gruzd, Mai, & Kampen, 2017), which allows for advanced analysis despite a relatively low cost and technical threshold. Similarly, we have long found Hawksey's Twitter Archiving Google Sheet (https://tags.hawksey.info), which is free and requires no programming, to be helpful for data collection and some simple analysis.

Storing Data: It is often relatively simple to collect large amounts of social media data, creating a need for careful storage and organization. As they respond to this need, LDT researchers should carefully consider the security and privacy that is (not) offered by a given storage solution. Furthermore, LDT researchers may be interested in storage platforms (e.g., Open Science Framework repositories; https://osf.io) that facilitate sharing data with collaborators, peer reviewers, and even the public, though this also requires careful attention to security and privacy.

Workflows: Social media researchers should establish workflows—common and ongoing understandings of *who* is doing *what*, *when* (and *why*)—even when working alone. Scheduling research tasks is especially important because social media platforms may place restrictions on how much historical data can be collected—or how much contemporary data can be collected within a particular timeframe. Workflows should also involve validating data collection and analysis. In one study (Rosenberg, Greenhalgh, Wolf, & Koehler, 2017), we experienced a data gap because our workflow did not include any supervision of an automated data collection process; human supervision could have noticed an important change in how a learning community was operating that our automated process was incapable of detecting. Multiple researchers should, therefore, regularly review and audit the data, code, and processes associated with a project.

Documentation: Those adopting positivist approaches to LDT inquiry may be interested in documenting their work so that other education researchers can reproduce their studies (van der Zee & Reich, 2018), whether in a different context or to confirm the contributions of the original study. However, even if supporting reproducibility is judged to be inappropriate given chosen ethical principles (e.g., because sharing data from the original study would reveal personal information) or unnecessary given a chosen research paradigm, careful documentation of the research process will be of great value to the social media researcher. The ready availability of social media data may make it necessary for researchers to manage several files, and seemingly-trivial technical details can have a significant impact on how phenomena are operationalized or data are analyzed. Thus, carefully and thoroughly recording decisions (e.g., commenting code, writing memos, etc.) will help researchers accurately represent—and convincingly defend—their findings. Maintaining access to older versions of files (e.g., through version control software like Git, which is most accessible through services like Github; https://github.com) will allow for understanding and auditing previous decisions, and tools like R Markdown (https://rmarkdown.rstudio.com/) allow researchers to document their work (and even write manuscripts) by embedding analyses (e.g., the output from statistical tests) into text documents.

Collecting Data

The use of social media leaves behind many different kinds and scales of data, ranging from meaningful minutiae to staggering amounts of "big data." We have previously discussed the importance of identifying a phenomenon of interest (or unit of analysis), which will provide LDT researchers with initial guidance for what (and how much) social media data to collect. In this section, we identify three additional considerations related to collecting data:

Obtrusive or Unobtrusive: Much social media data can be collected *unobtrusively* (i.e., without alerting participants or thereby influencing the data; Lee, 2015). This has advantages in terms of authenticity and accuracy—particularly since participants do not always provide accurate accounts of their social media activity, as evidenced by Junco's (2013) finding that college students tended to overestimate the amount of time they spent on Facebook. However, unobtrusive methods can only report observable phenomena, and LDT researchers may wish to add obtrusive data collection to their research to better understand their participants. For example, Greenhow, Gleason, Marich, and Staudt Willet (2017) supplemented a collection of doctoral students' tweets with a survey of those students, which provided additional insight into their experience with Twitter. Furthermore, obtrusive methods—such as Gleason's (2018) request for participants' Twitter archives—may provide access to social media data that cannot be obtained through unobtrusive methods.

Process: Social media data may be collected in a number of ways. When using obtrusive methods, researchers will follow more traditional processes of contacting research participants directly and obtaining data from them. When using unobtrusive methods, researchers will most often collect data through a platform's *application programming interface* (API), whether directly (by writing their code) or indirectly (by using third-party software that accesses the API). Platforms sometimes change how their API works or what it provides access to, which may create obstacles for researchers; for more on this subject, we recommend Bruns and Burgess's (2016) now-partial history of researchers' responses to changes in the Twitter API.

Other techniques may also be used to unobtrusively collect data, though they are associated with concerns about reliability (in the case of manually collecting or copying data) or violating terms of service agreements (in the case of *web scraping*—i.e., programmatically collecting data that has been formatted for an end-user). In yet other cases, a third party has already collected social media data, and researchers may ask permission to access that data. For example, Gao and Li (2017) and Xing and Gao (2018) asked to access tweets archived by the organizers of a weekly synchronous chat. Similarly, Carpenter, Tani, Morrison, and Keane (2018) accessed Participate's archive of past Twitter chats from hundreds of different hashtags (see https://archive.participate.com/) to compare teachers' use of 16 different education-related Twitter hashtags.

Quantity: Different quantities of data are useful for answering different research questions. Large-scale studies incorporating "big data" are helpful for observing connections and trends, but small data are better for understanding specificity and motives (Latzko-Toth, Bonneau, & Millette, 2017). Given the sheer amount and availability of social media data, it is not uncommon for researchers to have access to—or even intentionally collect—more data than they actually need; in this case, they will need to decide how to sample, subset, or otherwise limit the data they have collected prior to their analysis. These concerns are not limited to social media research (e.g., Glesne, 2016; Remler & Van Ryzin, 2015), but social media researchers should also make this decision based on the phenomenon or unit of analysis that the researchers have previously decided on. For example, Greenhow, Li, and Mai (2019) were interested in social media use *during a specific educational event* (a conference) and therefore limited their data by a timeframe; in contrast, Carpenter, Kimmons, Short, Clements, and Staples (2019) were focused on *a specific kind of social media user* (teachers) and eliminated data from other users. On social media platforms that allow for reposting others' material, LDT researchers have articulated compelling reasons to both retain duplicate material (e.g., because reposts serve as the amplification of a theme; Greenhalgh, Rosenberg, & Wolf, 2016) and remove it from the data (e.g., because reposts do not add new topics to a discussion; Gao & Li, 2017) while preparing for analysis.

Analyzing Data

In this section, we identify considerations that LDT researchers must make as they analyze social media data. Whatever decisions they make, researchers should both acknowledge that any analysis comes with a degree of uncertainty and take steps to establish the trustworthiness of their research. Naturally, this concern is not unique to social media research, and LDT researchers may consult one of several resources to guide their efforts (e.g., Creswell, 2014; Creswell & Miller, 2000; Maxwell, 2013; Remler & Van Ryzin, 2015). The following considerations are more specific to the analysis of social media data:

Spam: Before analyzing social media data, researchers must determine what to do with spam, which can be understood as "undesirable text, whether repetitive, excessive or interfering" (Brunton, 2013, p. xxii). Carpenter, Staudt Willet, Koehler, and Greenhalgh (2019) noted that educational researchers may take varied approaches to spam depending on their research goals. For example, researchers may eliminate spam from their analysis if their goal is to accurately represent learning activity or compare learning activity across contexts. In contrast, researchers focused on investigating participants' experiences of social media are more likely to retain spam content (i.e., to represent that experience fully).

Machine vs. Human Analysis: Social media data is often structured in a standardized way that facilitates *machine analysis*; indeed, machine techniques such as sentiment analysis, topic modeling, and classification algorithms allow researchers to approximate the kinds of analyses traditionally requiring human effort. For example, Greenhalgh, Staudt Willet, Rosenberg, and Koehler (2018) found that machine analysis could reasonably infer Twitter users' geographic location based on unobtrusively-collected data, presenting several practical advantages over human analysis of obtrusively-collected data. However, *human analysis* of data can account for nuance and meaning in ways that even the most advanced machine analyses cannot match. Researchers' choice between machine- or human-driven methods should, therefore, be driven by considerations of the *amount* of data researchers went to analyze (with more substantial amounts making human analysis impractical) and the *degree of nuance* desired (with higher levels of nuance being beyond the capability of machine analysis). Innovative combinations of machine and human analyses—such as Nelson's (2017) *computational grounded theory*—allow researchers to leverage the advantages of both approaches.

Networks: The networked aspect of social media platforms and education researchers' interest in interpersonal interaction invites LDT scholars to employ network analyses to understand social media activity (e.g., Rosenberg et al., 2019). Within network analysis, researchers may decide to focus on processes of *influence* (e.g., how ideas spread and affect individuals) or *selection* (e.g., how someone decides who to interact with). When representing networks visually, researchers must make choices about what features to highlight (e.g., how often users contribute, how frequently two users interact, or demographic information about users).

Disseminating Research

This step includes considerations related to *writing* the results of social media research studies as well as *sharing* and *publicizing* the final products. Although these issues are not unique to social media researchers—for a broader discussion of open science practices in education research, see van der Zee and Reich (2018)—some considerations stand out in this specific context:

Sharing Data: As previously described, some of the advantages of doing research based on social media data include their relative availability and authenticity. Given these advantages, it is logical for researchers to wish to share these data with others—whether as examples in written manuscripts (to support a study's assertions) or as publicly available datasets (to allow others to complete similar studies). Although the former goal is logical—and the latter laudable—we urge caution and recommend that researchers do so with the considerations in the Conducting Ethical Research section in mind.

Sharing Code: Collecting and analyzing social media data requires sustained attention to technical detail, and many education researchers do not have access to the technical training that is necessary to be aware of all the nuances involved in these processes (Kimmons & Veletsianos, 2018). Thus, while we urge caution before sharing social media *data* with other education researchers, we see few—if any—disadvantages to researchers' open sharing of the *code* they develop (indeed, our own research has long depended on software that other scholars have developed and shared). Sharing code helps communicate detailed methodological decisions, allows communities of scholars to validate findings (even without sharing data), helps others build upon the work, and can help establish standards for the rigor and quality of social media research in the field of LDT. Platforms such as GitHub (https://github.com) and the Open Science Framework (https://osf.io) facilitate the sharing of code online; however, the full benefits of sharing code are dependent on the documentation strategies that we have previously described.

Publishing and Publicizing Research: One interesting aspect to doing social media research is that social media themselves can help researchers publish and publicize their findings—a phenomenon called "social scholarship" (see Greenhow, Gleason, & Staudt Willet, 2019). In addition to blogs and other general-audience social media platforms, there are a number of academic-focused platforms that may be of use to LDT researchers seeking to promote and share their work; however, we encourage careful evaluation of the business models, privacy policies, and other practices of these platforms before using them. More importantly, researchers should consider whether and how to share their findings with social media users or communities whose data made that research possible. This is important not only in the context of ongoing considerations of what constitutes ethical relationships with those who contribute to research (Maxwell, 2013; Selwyn, 2019) but also given that social media users have expressed particular interest in learning about research based on their data (Fiesler & Proferes, 2018).

Conclusion

The ready availability of social media data provides many opportunities—but also considerable challenges—for LDT researchers. LDT research, social media platforms, and research tools are all diverse, not to mention continually evolving. Thus, scholars interested in using social media data in their research should familiarize themselves with the broad considerations that guide effective research in this domain—not just detailed walkthroughs of how to complete a specific task. In this chapter, we have described some of the key considerations associated with six critical steps in social media-focused LDT research: Conducting Ethical Research; Framing the Research; Organizing the Research Process; Collecting Data; Analyzing Data; and Writing, Sharing, and Publicizing Research. These considerations allow for numerous valid approaches to social media research and help researchers avoid the arguably more-numerous pitfalls they could encounter.

References

boyd, d., & Crawford, K. (2012). Critical questions for big data: Provocations for a cultural, technological, and scholarly phenomenon. *Information, Communication & Society*, *15*(5), 662-679. doi:10.1080/1369118X.2012.678878

boyd, d., Golder, S., & Lotan, G. (2010). Tweet, tweet, retweet: Conversational aspects of retweeting on Twitter. In *Proceedings of the 43rdh Annual Hawai'i International Conference on System Sciences*. Los Alamitos, CA: IEEE Computer Society.

Bruns, A., & Burgess, J. (2016). Methodological innovation in precarious spaces: The case of Twitter. In H. Snee, C. Hine, Y. Morey, S. Roberts, & H. Watson (Eds.), *Digital methods for social science: An interdisciplinary guide to research innovation* (pp. 17–33). New York, NY: Palgrave Macmillan.

Brunton, F. (2013). *Spam: A shadow history of the Internet*. Cambridge, MA: MIT Press.

Carpenter, J. P., Kimmons, R., Short, C. R., Clements, K., & Staples, M. E. (2019). Teacher identity and crossing the professional-personal divide on Twitter. *Teaching and Teacher Education*, *81*. doi:10.1016/j.tate.2019.01.011

Carpenter, J. P., Staudt Willet, K. B., Koehler, M. J., & Greenhalgh, S. P. (2019). Spam and educators' Twitter use: Methodological considerations and challenges. *TechTrends*. doi:10.1007/s11528-019-00466-3

Carpenter, J., Tani, T., Morrison, S., & Keane, J. (2018, March). Exploring the education Twitter hashtag landscape. In E. Langran & J. Borup (Eds.), *Proceedings of Society for Information Technology & Teacher Education International Conference 2018* (pp. 2230-2235). Association for the Advancement of Computing in Education (AACE).

Creswell, J. W. (2014). *Research design: Qualitative, quantitative, and mixed methods approaches* (4th ed.). Thousand Oaks, CA: SAGE.

Creswell, J. W., & Miller, D. L. (2000). Determining validity in qualitative inquiry. *Theory into Practice, 39*(3), 124-130. doi:10.1207/s15430421tip3903_2

Fiesler, C., & Proferes, N. (2018). "Participant" perceptions of Twitter research ethics. *Social Media and Society*, *4*(1). doi:10.1177/2056305118763366

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

Gao, F., & Li, L. (2017). Examining a one-hour synchronous chat in a microblogging-based professional development community. *British Journal of Educational Technology*, *48*, 332-347. doi:10.1111/bjet.12384

Gleason, B. (2018). Adolescents becoming feminist on Twitter: New literacies practices, commitments, and identity work. *Journal of Adolescent & Adult Literacy*, *62*, 281-289 doi:10.1002/jaal.889

Glesne, C. (2016). *Becoming qualitative researchers: An introduction* (5th ed.). Boston, MA: Pearson Education, Inc.

Golder, S., Ahmed, S., Norman, G., & Booth, A. (2017). Attitudes toward the ethics of research using social media: A systematic review. *Journal of Medical Internet Research*, *19*(6). doi:10.2196/jmir.7082

Greenhalgh, S. P., Rosenberg, J. M., & Wolf, L. G. (2016). For all intents and purposes: Twitter as a foundational technology for teachers. *E-Learning and Digital Media*, *13*, 81-98. https://doi.org/10.1177/2042753016672131

Greenhalgh, S. P., Staudt Willet, K. B., Rosenberg, J. M., & Koehler, M. J. (2018). Tweet, and we shall find: Using digital methods to locate participants in educational hashtags. *TechTrends, 62*, 501-508. doi:10.1007/s11528-018-0313-6

Greenhow, C. M., Gleason, B., Marich, H., & Staudt Willet, K. B. (2017). Educating social scholars: Examining novice researchers' practices with social media. *Qwerty – Open and Interdisciplinary Journal of Technology, Culture and Education*, *12*(2), 30-45. Retrieved from http://www.ckbg.org/qwerty/index.php/qwerty/article/view/269/

Greenhow, C. M., Gleason, B., & Staudt Willet, K. B. (2019). Social scholarship revisited: Changing scholarly practices in the age of social media. *British Journal of Educational Technology*, *50*, 987-1004. doi:10.1111/bjet.12772

Greenhow, C., Li, J., & Mai, M. (2019). From tweeting to meeting: Expansive professional learning and the academic conference background. *British Journal of Educational Technology*, *50*, 1656-1672. doi:10.1111/bjet.12817

Gruzd, A., Mai, P., & Kampen, A. (2017). A how-to for using Netlytic to collect and analyze social media data: A case study of the use of Twitter during the 2014 Euromaidan Revolution in Ukraine. In L. Sloan & A. Quan-Haase (Eds.), *The SAGE handbook of social media research methods* (pp. 513-529). Los Angeles, CA: SAGE.

Junco, R. (2013). Comparing actual and self-reported measures of Facebook use. *Computers in Human Behavior, 29*, 626-631. doi:10.1016/j.chb.2012.11.007

Kimmons, R., & Johnstun, K. (2019). Navigating paradigms in educational technology. *TechTrends*. doi:10.1007/s11528-019-00407-0

Kimmons, R., & Veletsianos, G. (2018). Public Internet data mining methods in instructional design, educational technology, and online learning research. *TechTrends*, *62*, 492-500. doi:10.1007/s11528-018-0307-4

Krutka, D. B., Asino, T. I., & Haselwood, S. (2018). Editorial: Eight lessons on networked teacher activism from #OklaEd and the #OklaEdWalkout. *Contemporary Issues in Technology and Teacher Education*, *18*(2). Retrieved from https://www.citejournal.org/volume-18/issue-2-18/social-studies/editorial-oklaed-and-the-oklaedwalkout-eight-lessons-on-networked-teacher-activism

Krutka, D. G., Heath, M. K., & Staudt Willet, K. B. (in press). Foregrounding technoethics: Toward critical perspectives in technology and teacher education. *Journal of Technology and Teacher Education*.

Latzko-Toth, G., Bonneau, C., & Millette, M. (2017). Small data, thick data: Thickening strategies for trace-based social media research. In L. Sloan & A. Quan-Haase (Eds.), *The SAGE handbook of social media research methods* (pp. 199-214). Los Angeles, CA: SAGE.

Lee, R. M. (2015). *Unobtrusive measures*. doi:10.1093/OBO/9780199846740-0048

Maxwell, J. A. (2013). *Qualitative research design: An interactive approach*. Thousand Oaks, CA: SAGE Publications, Inc.

Mishra, P., Koehler, M. J., & Greenhow, C. (2016). The work of educational psychologists in a digitally networked world. In L. Corno & E. M. Anderman, *Handbook of educational psychology* (3rd ed., pp. 29-40). New York, NY: Routledge.

Moreno, M. A., Goniu, N., Moreno, P. S., & Diekema, D. (2013). Ethics of social media research: Common concerns and practical considerations. *Cyberpsychology, Behavior, and Social Networking, 16,* 708-713. doi:10.1089/cyber.2012.0334

Nelson, L. K. (2017). Computational grounded theory: A methodological framework. *Sociological Methods & Research*. doi:10.1177/0049124117729703

Ngai, E. W. T., Tao, S. S. C., & Moon, K. K. L. (2015). Social media research: Theories, constructs, and conceptual frameworks. *International Journal of Information Management*, *35*(1), 33–44. doi:10.1016/j.ijinfomgt.2014.09.004

Peng, R., & Matsui, E. (2017). *The art of data science*. Retrieved from https://bookdown.org/rdpeng/artofdatascience/

Penuel, W. R., & Frank, K. A. (2016). Modes of inquiry in educational psychology and learning sciences research. In L. Corno & E. M. Anderman (Eds.), *Handbook of educational psychology* (3rd ed.; pp. 30-42). New York, NY: Routledge.

Proferes, N. (2017). Information flow solipsism in an exploratory study of beliefs about Twitter. *Social Media + Society*, *3*(1). doi:10.1177/2056305117698493

Proferes, N., & Walker, S. (2020). Researcher views and practices around informing, getting consent, and sharing research outputs with social media users when using their public data. In *Proceedings of the 53rd Annual Hawai'i International Conference on System Sciences*. Los Alamitos, CA: IEEE Computer Society.

Remler, D. K., & Van Ryzin, G. G. (2015). *Research methods in practice: Strategies for description and causation* (2nd ed.). Thousand Oaks, CA: SAGE Publications Inc.

Rosenberg, J. M., Greenhalgh, S. P., Wolf, L. G., & Koehler, M. J. (2017). Strategies, use, and impact of social media for supporting teacher community within professional development: The case of one urban STEM program. *Journal of Computers in Mathematics and Science Teaching*, *36*, 255-267.

Rosenberg, J. M., Reid, J. W., Dyer, E., Koehler, M. J., Fischer, C., & McKenna, T. J. (2019, August 21). Exploring the Next Generation Science Standards chat (#ngsschat) professional network on Twitter through social network analysis. Open Science Framework preprint. doi:10.31219/osf.io/uwza6

Selwyn, N. (2019). *What is digital sociology?* Medford, MA: Polity Press.

SIGCHI Ethics Committee. (2017, November 30). Do researchers need to follow TOS? [Medium post]. Retrieved from https://medium.com/p/f3bde1950d3c/

Sloan, L., & Quan-Haase, A. (2017). *The SAGE handbook of social media research methods.* Thousand Oaks, CA: SAGE Publications Inc.

Sloan, L., Quan-Haase, A., Kitchin, R., & Beninger, K. (2017). *Social media research & ethics* [Streaming video]. doi:10.4135/9781526413642

Staudt Willet, K. B., Koehler, M. J., & Greenhalgh, S. P. (2017). A tweet by any other frame: Comparing three theoretical frameworks for studying educator interactions on Twitter. In L. Liu & D. Gibson (Eds.), *Research highlights in technology and teacher education 2017* (pp. 63-70). Waynesville, NC: Association for the Advancement of Computing in Education (AACE). Retrieved from http://www.learntechlib.org/p/180960/

Suomela, T., Chee, F., Berendt, B., & Rockwell, G. (2019). Applying an ethics of care to Internet research: Gamergate and Digital Humanities. *Digital Studies/Le Champ Numérique*, *9*(1), 4. doi:10.16995/dscn.302

Townsend, L., & Wallace, C. (2016). *Social media research: A guide to ethics.* Aberdeen, U.K.: University of Aberdeen. Retrieved from http://www2.port.ac.uk/research/ethics/CurrentDownloads/filetodownload,198032,en.pdf

Tufekci, Z. (2014). Big questions for social media big data: Representativeness, validity, and other methodological pitfalls. In E. Adar, & P. Resnick (Eds.), *Proceedings of the Eighth International AAAI Conference on Weblogs and Social Media*. Palo Alto, CA: The AAAI Press.

U.S. Department of Health and Human Services (2018). *Revised common rule.* Retrieved from https://www.hhs.gov/ohrp/regulations-and-policy/regulations/finalized-revisions-common-rule/index.html

van der Zee, T., & Reich, J. (2018). Open education science. *AERA Open*, *4*(3), 1-15. doi:10.1177/2332858418787466

Van Osch, W., & Coursaris, C. (2015). A meta-analysis of theories and topics in social media research. In T. X. Bui & R. H. Sprague (Eds.), *Proceedings of the 48th Annual Hawai'i International Conference on System Sciences* (pp. 1668-1675). Los Alamitos, CA: IEEE Computer Society.

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*, 372-382. doi:10.1177/1556264617725200

Welser, H. T., Smith, M., Fisher, D., & Gleave, E. (2008). Distilling digital traces: Computational social science approaches to studying the Internet. In N. Fielding, R. M. Lee, & G. Blank, *The SAGE handbook of online research methods* (pp. 116-141). Thousand Oaks, CA: SAGE Publications, Ltd.

Wenger, E. (1998). *Communities of practice: Learning, meaning, and identity*. New York, NY: Cambridge University Press.

Xing, W., & Gao, F. (2018). Exploring the relationship between online discourse and commitment in Twitter professional learning communities. *Computers & Education, 126,* 388-398. doi:10.1016/j.compedu.2018.08.010