30

Ethics in Digital Research

Katrin Tiidenberg

INTRODUCTION

Research ethics as a topic of both public and scholarly debate tends to (re)surface when things go wrong. The history of research ethics could be told in our mistakes, and our collective attempts to learn from them. Ostensibly, we can start that history from the dehumanizing experiments of World War II, the Tuskegee syphilis study, and Stanley Milgram's groundbreaking yet disturbing research into human behavior. It can be said that (the reveal of) these mistakes led to the UN Declaration of Human Rights (1948), the Nuremberg Code (1949), the Declaration of Helsinki (1964), and the Belmont Report (1979); meant for protection of human subjects in biomedical and behavioral research; and continuously relevant in ethical management of most research happening with people today (see Mertens, Chapter 3, this volume).

Following the breadcrumb trail of research ethics failures through decades, we could tentatively add the uproar following the 2014 publication of the Facebook 'emotional contagion' study (Kramer et al., 2014) to the list. It meant researchers altering 689,000 Facebook users' news feeds to explore how exposure to emotional content influences what they posted. In May 2016 a student researcher leaked the data of 70,000 users of OKCupid (a dating platform), claiming he did so for the benefit of the scholarly community (Resnick, 2016). This indicates that while we may have learned from our past mistakes, and our newer ones may cause comparatively less harm, research ethics needs constant reflection. Neither the phenomena we study, the contexts we study them in, nor public perceptions of what is permissible are static.

The fact that an increasing amount of (social) research happens on, about, or with the help of the internet only complicates matters. While there is no consensus on the topic, compelling arguments have been made about the ethical specificity of digital context (Markham and Buchanan, 2015). Beaulieu and Estelalla (2012) even claim that internet research means a certain remediation of research practices, and thus a transformation of research objects, tools, and relations. Concurrently, digital research brings together a plethora of scholars with what can be diametrically opposing methodological, paradigmatic, epistemological, and ontological training and worldviews. Studies by scholars who consider themselves internet researchers coexist with experimental, correlational, and observational studies conducted online simply because it is convenient (Merriman, 2015; Tolich, 2014). This makes it quite difficult to agree on the need for, and content of, reasonable practices and sufficient standards for research.

In what follows, I outline some of the more persistent ethical issues that scholars involved in digital research face. Classic ethical concepts like informed consent, confidentiality, anonymity, privacy, publicity, and harm can be difficult to operationalize in a socio-technical context that is persistent, replicable, scalable, and searchable (boyd, 2010). In daily lives we often interact with software, interfaces and devices in ways that turn what we are used to considering an ‘interactional context’ into an ‘active participant’ (Markham, 2013). Scholars partaking in digital research, therefore, often find themselves faced with a lot of gray areas. Their individual sense of what is right and wrong; their discipline’s conventions; the legal and institutional conditions of approval; and the competition for professional relevance in a world where a lot of research is undertaken by private companies like Facebook, may at times clash or collapse.

CHANGING DISCOURSES ABOUT THE INTERNET, SOCIABILITY, AND RESPONSIBILITY

Before moving on to ethics in digital qualitative data collection, I want to briefly address some of the current thinking on online interactions and sociability, particularly from the perspective of responsibility. Such discourses feed into and filter trending attitudes in research ethics.

Jose van Dijck (2013) offers a compelling theory of social media-driven changes in various social norms in her book *The Culture of Connectivity: A Critical History of Social Media*. For instance, the meaning and the norm of ‘sharing’ has, according to her, markedly shifted during the past decade. The coded structures of social media platforms like Facebook (see Ditchfield and Meredith, Chapter 32, this volume) impose buttons like ‘Share’ as social values. These ‘have effects in cultural practices and legal disputes, far beyond platforms proper’ (van Dijck, 2013, p. 21), but are alarmingly ambiguous in their meaning. In the example provided by van Dijck, sharing connotes both users distributing their own information to each other, as well as the sharing of that personal information by service providers with third parties (2013, p. 46). Similarly, Markham (2016, p. 192) points out that sharing has become the ‘default relationship between the self and technological infrastructures’, which discursively naturalizes the massive harvesting and storage of personal data by platforms. Markham (2016, p. 194) goes on to point out the dangers of the ‘this is just how the internet works’ discourse, which frames privacy as an individual burden, and removes ‘agency from corporate interests, platform designs, and algorithmic activities that, in fact, quite powerfully and actively mediate how one’s personal activities online become public and publicly available’.

These discursive and attitudinal shifts operate in a context where, despite Facebook CEO Mark Zuckerberg’s attempts to convince us that our need for privacy indicates an unhealthy desire to hide something (Kirkpatrick, 2010a) or is perhaps a relic of the past (Kirkpatrick, 2010b), more than half of American social networking site users (58 percent) have changed their main site’s privacy settings to only be accessible to friends (Madden, 2012). Practically, it does little to limit non-friends' access to much of our interactions on the site. Our comments on our friends' posts are governed by their privacy settings, not our own; and our friends get notifications of our comments on our other friends' posts, even if they are not themselves connected. It is not surprising then, that nearly half of social media users feel that managing their privacy controls is difficult (Madden, 2012), and they 'have limited control over how their data is used online' (Microsoft Trustworthy Computing, 2013). Yet, paradoxically, users also feel they are solely responsible for their privacy online (40 percent of all Europeans and 46 percent of all Americans, Kügler, 2014). Understandably this tension leads to the (American) public trusting Facebook even less than they trust the IRS or the post-Snowden NSA (boyd, 2016).

Additionally, the legally and morally dubious model of 'effective consent' (disclosure via terms of service agreement) has become a de facto standard in the industry (Flick, 2016, p. 17) and again operates by assigning responsibility to individual users instead of corporate players (2016, p. 20). Facebook's data policy (last revised January 30, 2015), for example, reveals that we have all agreed to them collecting information about what we do on the platform, what others do (sending us messages, uploading images of us), the constellations of people and groups around us, our device use (including geolocation), payments, third party websites and apps that we use that use Facebook services (e.g. when you log on to Slideshare using Facebook), information from companies we use that are owned by Facebook (i.e. Instagram, Whatsapp) and finally, and perhaps most eerily, about us and our activities 'on and off Facebook from third party partners', which might as well encompass one's entire web usage.

These troubling shifts in discourse of individual responsibility can be detected in the social research community itself (cf. Flick, 2016, or boyd, 2016, on lack of scholarly consensus regarding the ethical aspects of the aforementioned Facebook emotional contagion study, and Weller and Kinder-Kurlanda, 2015, for social media researchers' attitudes toward ethics in social media research). Researchers are having a hard time agreeing on what data is public and what data is private, and how publicly accessible data should be treated. Debates over boundaries and best practices of internet research ethics are ongoing (Flick, 2016; Markham and Buchanan, 2015; Mauthner et al., 2012). Various professional organizations – Association of Internet Researchers (Markham and Buchanan, 2012); the Norwegian National Committee for Research Ethics in the Social Sciences and the Humanities (NESH, 2014), The SATORI Ethics assessment in Internet Research Ethics (Shelley-Egan, 2015) – urge researchers to ask themselves those difficult questions, while increasingly realizing that the 'ethical guides of traditional disciplines are of limited usefulness' (Beaulieu and Estalella, 2012, p. 10).

DIGITAL QUALITATIVE DATA COLLECTION

Qualitative data collection on/in/through the internet is wide and varied, and as mentioned above, used by scholars from different disciplinary backgrounds. A brief look at the contents of this very *Handbook* reveals interviews (see Roulston and Choi, Chapter 15, this volume), focus group discussions (see Morgan and Hoffman, Chapter 16, this volume), observations (see Wästerfors, Chapter 20, this volume); and collection of textual, visual, audio and media data for narrative (see Murray, Chapter 17, this volume), conversation (see Jackson, Chapter 18, this volume), and discourse analyses (see Rau et al., Chapter 19, this volume) or performative ethnographies (see Denzin, Chapter 13, this volume). All of these can and are successfully conducted online and/or about internet-related phenomena. Our everyday lives weave through mediated and non-mediated contexts, thus delineating data collection by its digitality is problematic at best (for a persuasive complication of the online–offline divide in qualitative inquiry see Orgad (2009), and following responses by Bakardjieva (2009), and Gajjala (2009); see also Fielding, Chapter 37, this volume). Hence, it may be more sensible to focus on which internet-specific tensions arise in various methodological steps of qualitative data collection. After all, as Markham (2006) astutely points out, all of our methods decisions – from asking questions and defining field boundaries to interpreting data – are, in fact, ethics decisions (see also Mauthner et al., 2012, for a distilled discussion on what our methods 'do' ethically; see also Mertens, Chapter 3, this volume).

This approach shifts our focus from the collectables – from what we gather and create as data – to the process. We start thinking less about whether something was 'publicly' accessible and hence fair game to be grabbed and analyzed, and more about whether the fact that we can technically access it automatically means we should. Reviewing contributions from the early 2000s Eynon et al. (2008, p. 27) point out that, while digital research is not 'intrinsically more likely to be harmful than face-to-face methods', it can make it more difficult to evaluate risks of harm, and complicate judging participants' and wider publics' reactions to research. Analytically, this can be linked to the internet's affordances for human sociability – the fact that much of what used to be ephemeral in our everyday lives has become visible and traceable, often in 'forms divorced from both the source and the intended or actual audience' (Markham, 2011, p. 122; see also Markham, Chapter 33, this volume). Thus, while social media affordances of persistence, replicability, scalabilty and searchability (boyd, 2010) allow researchers unprecedented access to aspects of meaning-making or identity construction; they are also tinged with ambiguities of whether, what for, when, and for how long these processes should be observed, collected, and preserved for the sake of research.

CONTESTED CONCEPTS

Typically the focus of research ethics, as outlined in various declarations, acts, and guidelines, is on maintaining beneficence (minimization of harm and maximization of benefits), respect, and justice for people involved (Markham and Buchanan, 2012; see Mertens, Chapter 3, this volume). How these are translated into actual research practices (e.g. seeking informed consent or manipulating data for confidentiality, anonymity or privacy) in digitally saturated contexts continues to be an issue of significant debate. In the following I will describe some of the resurfacing complications surrounding these concepts.

Human Subjects Research

The 'human subjects model' can be considered a reaction to the harmful medical and experimental research conducted in the first half of the twentieth century, and is built on the concepts of confidentiality, anonymity, and informed consent; all derived from the basic human right to privacy (Eynon et al., 2008). While there are different approaches to what exactly counts as human subjects research, a typical definition focuses on interaction between the researcher and the participants, and the traceability of collected data to individuals (Walther, 2002). While in the earlier years of digital research there were some who advocated for considering online data text (White, 2002), it has become more common to be cautious, when estimating the 'humanness' of any data. This is more complicated in research where the unit of analysis is not a person, a group of people, or human behavior, but perhaps a malicious software attack, or density of a social network. Scholars may, in these cases, claim exemption from the human subjects model and the related ethics board review (Dittrich, 2015). However, as pointed out in the AoIR ethics guidelines:

> because all digital information at some point involves individual persons, consideration of principles related to research on human subjects may be necessary even if it is not immediately apparent how and where persons are involved in the research data. (Markham and Buchanan, 2012, p. 4)

Informed Consent

The idea of informed consent is grounded in principles of individual autonomy and beneficence. Broadly, it means that researchers commit to giving detailed information on the purpose, duration, methods, risks, and benefits of the study to participants, while participants have an absolute right to withdraw at any time (Marzano, 2012, p. 443).

The concept has a long history in medical and bioethics, where it is seen as an oversight mechanism to guarantee that research prioritizes participant welfare. Most ethics boards require that all research projects they deem human subjects research incorporate informed consent, or explicitly apply for an exception. Decisions over the need for, and type of informed consent procedures are based on the assumed steepness of risks. The more risk, the more formal (i.e. a signed form instead of an oral agreement) the informed consent process needs be. Risks are considered higher with research involving sensitive data or vulnerable participants, neither of which are as unproblematic as they may seem (cf. Egan et al., 2006, for a study where research participants with brain injuries found an ethics committee's ideas about their vulnerability patronizing and unhelpful). Exceptions tend to be given when risks of participating in the study are seen as minimal, for example, because of aggregation of data that is claimed to make it impossible to identify individual participants (cf. Zimmer, 2010, for how this as an assumption has backfired in the case of aggregated data collection from Facebook; see Ditchfield and Meredith, Chapter 32, this volume) and/or when research is not considered human subjects.

There are multiple tensions that arise when addressing the suitability of the informed consent model for digital research. On the one hand, there are worries that the mediated context makes it more difficult for the consent-seekers to 'determine autonomy, competence and understanding, and for consenters to understand the ramifications of the disclosure' (Flick, 2016, p. 17). On the other hand, it is quite common to claim that some spaces online can be considered public domain, and thus everything posted there can be considered 'naturally occurring data' (see Potter and Shaw, Chapter 12, this volume) and used without seeking any kind of explicit consent (cf. Rodham and Gavin, 2006, on informed consent and using data from message boards).

Additionally, the informed consent model is predicated on the expectation of research participants' autonomy, competence, and ability to understand risk; and assumptions of it being possible for researchers to imagine and predict future harm, including, for example, from storing data in a cloud, or sharing data in a data bank. Both of these assumptions are increasingly challenged as well (Mauthner, 2012; Markham, 2015; Markham and Buchanan, 2015).

Finally, voices from the ethnographic and feminist research traditions (Lomborg, 2012; Beaulieu and Estalella, 2012) point out the insufficiency and inappropriateness of rigid consent forms, and instead advocate for informed consent as a continuous negotiation (Lawson, 2004); a series of waivers of expected and behavioral social norms (Manson and O'Neill, 2007); or a situated decision that the researcher makes by focusing primarily on avoiding harm rather than consent per se (Markham and Buchanan, 2015). These approaches seem to be backed by studies about research participants' expectations toward the research process. Lewis and Graham (2007) found that participants reacted unfavorably to the idea of written consent, and were more interested in naturalistic, authentic approaches to information-giving.

Public or Private?

One of the more heated debates pertaining to digital research ethics is about what kinds of spaces, interactions, and data should be considered private, and which can be considered public. As Baym and boyd (2012, p. 322) point out, social media, thanks to its architecture and affordances, exponentially increases the potential for visibility and public engagement, thus requiring new skills and new mechanisms of control.

It is enticing to focus on the technical accessibility of information and define the internet as a vast public sphere. Categorizing it as such would seemingly release researchers from the difficult choices of making their presence known or seeking consent. This line of thinking is well illustrated in the following quote:

> it is important to recognize that any person who uses publicly available communication systems on the internet must be aware that these systems are, at their foundation and by definition, mechanisms for the storage, transmission, and retrieval of comments. While some participants have an expectation of privacy, it is extremely misplaced. (Walther, 2002, p. 207)

I would draw a parallel between the logic above and me saying that anyone traveling in the city must be aware that cars stop for pedestrians at lights and zebras. While it may be ‘misplaced’ for a person to cross randomly, I would not run them over based on my assumed right of way. Fortunately, an increasing number of social media researchers are less preoccupied with what people ‘must’ be aware of, and instead recognize that people and groups have particular expectations toward the privacy and publicity of their interactions no matter what their settings are (Bakardjieva and Feenberg, 2000; Ess and Jones, 2004; McKee and Porter, 2009; Sveningsson-Elm, 2009; Nissenbaum, 2010; Markham and Buchanan, 2012; Robards, 2013; Ess, 2014; Fileborn, 2015).

However, already a decade ago some authors (Barnes, 2006; Acquisti and Gross, 2006) noted a ‘privacy paradox’, where people claim they value privacy, yet their online practices seem to be counterproductive to maintaining it. A recent study from Hargittai and Marwick (2016) shows that while young adults may somewhat misunderstand risk or how effective particular privacy-protective behaviors are, these are not the sole reason for the privacy paradox. Rather, Hargittai and Marwick (2016, p. 3752) suggest that ‘users have a sense of apathy or cynicism about online privacy, and specifically believe that privacy violations are inevitable and opting out is not an option’.

Defining something as private or public has implications for how we assume it should be treated in a research context. Can we look at it? Can we analyze it? Can we reproduce it? Should we alter it for the sake of confidentiality? How should we ask about using it? In a context, where terms of user agreements are entirely dictated by service providers, and produce accessible, user-generated content as a side effect, some claim that we should even assume that most publicity is unintended (Merriman, 2015). I find Donald Treadwell’s (2014) thinking that differentiates intent of publication from publicity quite helpful here. According to him (2014, p. 51), most internet content – while public in the way billboards are – is much closer to informal discussions, or thinking aloud; than to stable opinions that have been published with intent.

The scope of this chapter does not allow us to fully delve into the complexity and the philosophical underpinnings of the concept of privacy, but Helen Nissenbaum’s widely cited work (2004, 2010) is an excellent source. She suggests interpreting privacy through the lens of ‘contextual integrity’ (cf. also McKee and Porter, 2009, on 'perceived privacy', and Warrell and Jacobsen, 2014, on 'intended audiences'). In our everyday lives, we all move through a plurality of different realms, each of which involve a distinct set of norms, roles, and expectations (Nissenbaum, 2004, p. 137). These include norms of information flow. As long as the information is flowing appropriately (Nissenbaum, 2010, p. 2) we feel our privacy to be maintained. The difficulty for researchers lies in operationalizing this concept. Do we commit to always asking what people's expectations are? This is undoubtedly not possible in many research situations. Similarly it is naïve to assume that people's expectations are stable or informed.

Additionally, recent years' key texts (Markham and Buchanan, 2012; Lomborg, 2012; Markham and Buchanan, 2015) recommend 'the distance principle' as a mechanism of thinking about privacy. The distance principle examines the distance between the researcher and the participants, but more importantly between data collected and the persons who created whatever content the data consists of. Through that, the potential for causing harm, and the appropriate course of action in terms of informed consent is assessed (Lomborg, 2012, p. 22). The smaller the distance, the more careful we need to be. Distance is considered to be smaller between a small sample of identifiable status updates and the people who posted them, than it is, for example, between the people who have tweeted and a large sample of automatically scraped, aggregated tweets.

Finally, in some good advice from Markham and Buchanan (2015, p. 6) it might not be all that helpful to ask if something is private or public in contexts, where 'information flow is constant and public, where people are always connected, or where cutting, pasting, forwarding, reposting, and other mashup practices remix our personal information in globally distributed, complex networks'. Instead, they suggest (2015) that we focus on people's expectations; on the sensitivity and vulnerability of both people and data; and primarily, on the impetus to do no harm.

Anonymity and Confidentiality

Anonymity and confidentiality are classic promises made to research participants in social research, and concepts often contemplated together, although their focus is slightly different. Ethics review boards systematically require both for approval. While confidentiality means accessing and sharing personal information only as authorized by the person concerned (and typically includes assuring participants their data will not be accessed by anyone but the researchers), anonymity is about ensuring that the person cannot be identified from the research data (Felzmann, 2013, p. 20). Typically anonymity is deemed sufficiently established when 'personally identifiable information' like names and ID numbers are stripped (for a discussion on the differences between the US and European definitions of personally identifiable information and its implications for anonymity, see Zimmer, 2010, p. 319).

The plausibility of either of those promises is questionable in a context where data-mining technologies can link participants to the 'information they produce and consume via a range of mobile devices, game consoles and other internet based technologies' (Markham, 2012, p. 336), and potential risks to security and integrity of data are manifold (Buchanan et al., 2011). Incidentally, it has occasionally been implied that internet pseudonyms, which participants choose for themselves, are far enough removed from their legal identities, and are thus enough to ensure confidentiality. This assumption creates ample difficulties (see Sveningsson, 2004, for a discussion on the necessity of protecting participants' internet pseudonyms as well as their legal identities). Similarly some researchers (Kendall, 2002) have had experiences of their participants rejecting the anonymity researchers attempt to provide by changing names and details. This puts the scholar in a difficult position between respecting and empowering the participants, predicting possible harm, and institutional demands of their IRB approval.

Beaulieu and Estalella (2012, p. 11) point out that for mediated ethnographies, which use direct quotes from the web, removing identifying details and assigning new pseudonyms is not enough. They talk about 'traceability' instead of anonymity, and suggest it shifts our focus toward 'exposure, ownership and authorship' (2012, p. 5) of content published online. This may mean that ethnographers are simply no longer in a position to offer subject protection, as anonymization has become effectively impossible (2012, p. 12).

Alongside these discussions there are also questions regarding the security of data storage – the format it is stored in, its location, the duration of storage. Researchers are taking steps to increase security by using encryption, passwords, onscreen working methods, and tracking software (Aldridge et al., 2010), all of which, while helpful, are not guarantees of security.

Sharing and Storing Qualitative Data

It is more and more common for funding agencies and research governance institutions to require that researchers share their data in digital archives and depositories. This requirement can even be linked to withholding of final grant payments (Mauthner, 2012). Philosophically, it relies on an admirable expectation that information and research results are 'public goods', access to which is a basic right (Willinsky, 2006). However, it also implies normalization of standardized, automatized and regulated data collection and storage (Mauthner, 2012). This is particularly problematic for qualitative researchers, because it undermines the ontological, epistemological, and ethical implications of trust, rapport, and the dialogic co-construction of data – all long-standing traditions in qualitative inquiry. The emotional relationship that develops between a researcher and a participant during some qualitative research is seen as creating an additional layer of ethical responsibility, which is, arguably, not available when qualitative data is accessed from a data bank (Crossen-White, 2015, citing Richardson and Godfrey, 2003). Perhaps even more importantly, seeking informed consent to share qualitative interview data in an archive constitutes different 'moral and ontological conditions of possibility' for storytelling, which may alter the very stories we are told (Mauthner, 2012, p. 164). Yet, anonymizing qualitative data to the extent where it is shareable in good conscience, may lead to it losing so much of its contextual integrity that the scientific value of its future use becomes questionable (see Corti, Chapter 11, this volume).

Stolen and Hacked Data

Finally, a short note on using stolen or hacked data in research. Unfortunately, due to malicious privacy hacks and failures of technology, sets of data not intended to be publicly shared, viewed or researched, are regularly made available online. These data may offer interesting insights into various aspects of human co-existence. They may also be unproblematically taken advantage of by corporate or individual developers, researchers, or journalists, thus presenting temptation for scholars to 'make something good out of a bad thing'. Consequentialist claims that no further harm is coming to those whose data is reused, are sometimes employed to justify these desires. As a researcher interested in visual self-presentation, sexuality, and shame, I would have found analyzing the leaked images of the Snappening (thousands of Snapchat accounts were hacked and photos leaked in 2014) or the Fappening (a collection of almost 500 private pictures of celebrities were leaked in 2014) quite gratifying. Similarly, the leaked Ashley Madison data would have probably been of interest to scholars researching online sexual behavior, dating, interpersonal relations, or gender. What can be said about this?

While using hacked or stolen data is so far mostly absent from ethics guidelines, it is sometimes discussed among members of professional organizations (e.g. the AoIR mailing list) or at conferences. In line with the dilemmas described above, concepts of privacy and publicity are employed (the data are, after all, now 'public'), as well as conditions of aggregation and anonymization, for non-qualitative research. Because of the lack of published deliberations on the topic, I would here rely on Ben Zevenbergen's email (AoIR mailing list, ethics discussion, October 2015, cited with permission), which complements this chapter's contextual-ethics approach, and summarizes many of the opinions voiced in that discussion. Zevenbergen pointed out, and I agree, that using stolen, leaked, and hacked data for research adds more unintended audiences to it, and implicitly condones (perhaps even incentivizes) the act of hacking and publishing ill-sourced datasets, and should thus be avoided.

MAKING CHOICES

Considering the above-described complexity, it is unsurprising that experts are reluctant to recommend clear-cut one-size-fits-all guidelines. Instead a case-based, inductive approach is often recommended. Turning ethical decision-making into a deliberative process during all steps of inquiry enables 'a more proactive role in determining how best – on a case-by-case basis – to enact beneficence, justice, and respect for persons' (Markham and Buchanan, 2015, p. 8). To illustrate, I offer some examples of ethics-related decision-making in digital qualitative data collection from some of my own recent projects.

I will be drawing on examples from two research projects – first, an ethnography with a community of sexy-selfie enthusiasts on Tumblr.com, and then, a study of how pregnant women present themselves on Instagram. In both cases people post scantily clad (or unclad) pictures of their bodies on the internet, and the data is public in terms of the posts being accessible to everyone (one needs to have downloaded the Instagram app in the case of Instagram, but there is not even a need to have an account in the case of Tumblr).

The research questions of the Tumblr study (Figure 30.1) meant I was collecting data ethnographically (see Buscatto, Chapter 21, this volume), which included talking to people; and my data collection spanned years. In addition, the topic involved nudity and sexuality, and I was aware from my discussions with the participants that they perceived the space as somewhat private, despite it

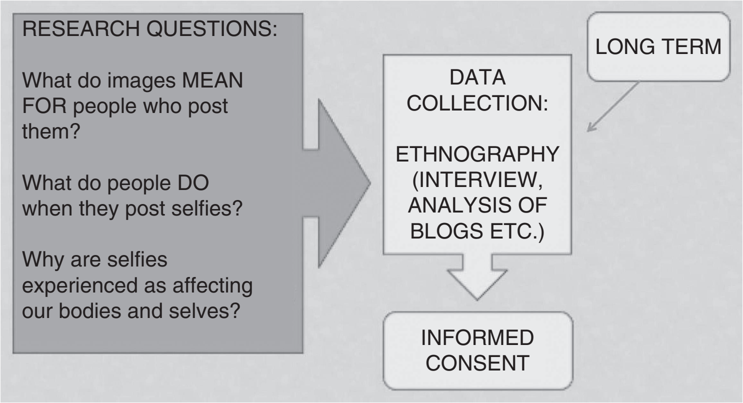

Figure 30.1 Tumblr study, research context

being technically publicly accessible. Thus I approached it as sensitive data. Taking all this into account, my choice to ask for informed consent is unsurprising.

During the Tumblr study, I found myself particularly drawn to the ideas of the ethics of care. Held (2006, p. 9) has defined care as both a value and a practice. The ethics of care ideally prescribes 'relations of trust and mutual respect' (Boellstorff et al., 2012, p. 129), and is seen as something that goes beyond avoiding harm. Based on recommendations in literature, I attempted to practice an ethics of care through dialogic consent, accurate portrayal, ethical fabrication, and doing good. These manifested as the following:

1. Despite having solicited 'blanket consent' at the beginning of my study, I double-checked with participants whenever entering a new stage of research ('I will now start looking at your images, is it still okay for me to do so?'), and when I wanted to include particular images in presentations or publications.
2. I kept interested participants in the loop of what I was doing to the data they helped me create via a research blog. It allowed me to do occasional member checks regarding some of my interpretations.
3. Markham (2012) has articulated the idea of ethical fabrication for protecting participants' privacy in contexts where public and private are shifting or difficult to interpret. She offers composite accounts, fictional narratives, and remix techniques as examples. I incorporated this idea, and devised some techniques particularly suitable for visual data. I edited all of the images I reproduced with an IOS application that made them look like pencil sketches, which retained visual and compositional detail, while reducing recognizability. I also somewhat altered the wording in the direct quotes from the web, doing reverse Google searches to make sure the altered text no longer (at least based on Google's data crawlers' current capabilities) led back to the blogs I studied.
4. While it is difficult to measure one's beneficial impact on the people studied without sounding hopelessly pretentious, it has been my understanding from five years' worth of conversations, that being a part of my research project has created enjoyable networks and carved out a space of self-reflection for my participants, which has had a therapeutic effect and assisted in them developing a certain sense of empowerment.

The second study I want to touch on had a markedly different context (Figure 30.2), both in terms of the questions and the practicalities. I was interested in people's self-presentations through the content they had chosen to publish on Instagram.

The practicalities of the project only allowed a month for data collection, but I had high-level technical assistance, which meant I could streamline it by experimenting with Instagram's API, which I had been curious about beforehand. Instagram doesn't have an internal messaging system[1] or reveal account holders' email addresses, thus my only option of reaching out to the approximately

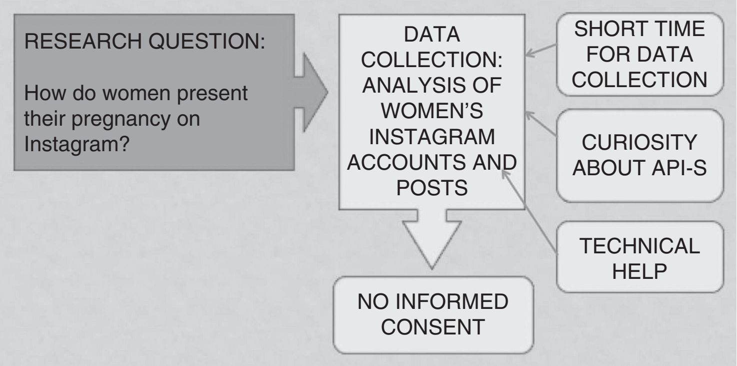

Figure 30.2 Instagram study, research context

250 accounts I included in the sample was by publicly commenting on their photos. I thought this was likely to be interpreted as 'creepy', and decided against it. This meant either forgoing informed consent or giving up on the project based on an assumption that the users would find my analysis of their images 'creepy' as well.

Table 30.1 shows my risk analysis to decide whether to continue without informed consent. Compared to my Tumblr study, people's practices indicated a markedly different perception of privacy. Where on Tumblr[2] people went to considerable length to protect their anonymity or, what has been called 'plausible deniability' by one of my participants, on Instagram real names and locations were regularly posted, and people systematically hashtagged their content to increase its searchability and visibility (e.g. concise informative hashtags and Instagram-specific attention-driven hashtags like #follow4follow or #like4like).

Based on the normative stances regulating pregnancy versus those policing sexually explicit conduct in Western capitalist societies, I decided the potential harm was much higher in the case of my research accidentally outing someone as a sex-blogger on Tumblr, than it was if I accidentally exposed someone as posting pregnancy- and family-related content from under their full name on Instagram. With significant unease, I thus decided to continue the study without informed consent, but tried to incorporate some of my practices of care developed during the Tumblr study.

a. I set up an Instagram account for purposes of accountability, described my study in the profile space, and offered an email address, where I could be reached (no one has emailed me, 18 people followed me back). I followed all of the accounts that had made it into my sample from this account. To turn accountability into a process, I used the researcher account to now and then go and 'like' some posts on the accounts in my sample.
b. I kept up my visual ethical fabrication techniques, and anonymized names and locations.
c. I 'outsourced' my member checks by engaging in regular dialog with trusted colleagues to make sure I portrayed these women accurately and fairly.

CONCLUSION

The purpose of this chapter has been to unsettle the approach to research ethics that equates it with a formalized list of rules, and can be seen as made dominant by the standardizing and streamlining attempts of ethics review boards, funding agencies, and research institutions today. Looking at widely used dictionary definitions of ethics we see that it may be interpreted as a consciousness of moral importance (Merriam Webster definition 2d) or a system of values (Merriam Webster definition 2a). In that case it becomes

Table 30.1 Comparing risk and privacy for the Tumblr study and the Instagram study

	NSFW TUMBLR	*PREGNANCY ON INSTAGRAM*
PERCEPTION OF PRIVACY	MORE PRIVATE • names, faces, locations, tattoos systematically removed • no hashtags, or personalized hashtags not intended for platform-wide searchability	MORE PUBLIC • names, faces, locations regularly included • hashtags suitable for searchability
POTENTIAL HARM FOR INDIVIDUAL	SEX = MORAL PANIC Accidental outing of participants' sexual preferences and lifestyles could cause harm to career, reputation, and personal relationships. People kept their blogs hidden from most of their other social networks.	PREGNANCY = SOCIALLY SUCCESSFUL STATE. Increase in social and moral capital for women, but are pregnant women vulnerable by default? What about the possible harm to unborn children?

Box 30.1

For other recent examples where qualitative researchers describe their ethics-related decision-making in great detail see Bianca Fileborn's and Stine Lomborg's work. Fileborn (2015) used Facebook to recruit study participants, and experienced a loss of control over where and with whom her recruitment advertisement was shared. She writes of the interesting conundrum of accountability, intended audiences, and her possible roles as a researcher, when friends of her friends comment on her study under these shared posts.

Lomborg (2012, pp. 24–9) describes her decision-making regarding the necessity of informed consent in a Twitter- and blog-based research project. While all of her data were, supposedly, both public and non-sensitive, the perceived privacy of her informants led her to opting for informed consent.

impossible, if not absurd, to rely on an external checklist. After all, how does one practice consciousness through a list of mandatory steps? A checklist-driven mentality presumes that institutional boards and individual scientists are able to predict ethical issues. Yet, we know, even from the relatively short history of the internet that there may be issues '"downstream" and only rise to the surface due to a change in Internet architecture, Internet norms, or even legal changes' (Markham and Buchanan, 2015, p. 10).

Thus, to bring the chapter to a close, I would offer an anti-checklist checklist; a set of reminders for those planning digital qualitative data collection and open to the approach of research ethics as situated, responsible decision-making. These may serve as reminders at critical junctures in specific projects (Markham and Buchanan, 2012), and hopefully shift our orientation from the past to the future (Markham, 2015).

- Our discourses about both (research) ethics and the internet are a result of 'tangles of human and non-human elements, embedded in deep – often invisible – structures of software, politics and habits' (Markham, 2015, p. 247). It's important to interrogate our assumptions, talk to colleagues, read texts by scholars from different disciplines.
- Despite the dominant discourse of personal responsibility, the technological affordances of networked sociality seem to leave our privacy at other people's discretion much more than before. Just because something is technically accessible and collectable, doesn't mean it should be accessed and collected.
- Having previous experience with internet research, or being an avid internet user, does not guarantee our understanding of other people's internet use. Behavioral expectations and perceptions do not seamlessly translate from space to space and group to group.
- All methods questions are ethics questions – 'most basically, a method is nothing more or less than a means of getting something done. And every choice one makes about how to get something done is grounded in a set of moral principles' (Markham, 2006, p. 16). Thus, we need to consider the ethical implications in our methods of defining field boundaries; accessing participants; raising a sample; collecting, organizing, analyzing, and archiving information; representing ourselves and others in writing; framing knowledge; and maintaining professional autonomy (see Markham, 2006; and Mauthner et al., 2012).
- We should avoid being lulled into complacency by the seemingly increasing regulation of research ethics. We are still responsible for our own research, even after our ethical review forms have been approved (Mauthner, 2012). Neither the possibility nor sufficiency of informed consent, confidentiality, or anonymity; the definition and implications of vulnerability or beneficence; the delineation of something as private or public; or what publicity indicates for research are obvious or uniformly observable in digital settings. Instead, they almost always depend on the context. Having an ethics review board approval and following the steps outlined in it may be a good start, but it does not guarantee a problem-free research process, nor does it absolve the researcher from being constantly engaged.

Granted, approaching research ethics as a personal pledge to be critically situated in all of one's research-related decisions is not an overly comfortable stance. It is future oriented, carries an expectation of the unexpected, and demands a certain willingness to stomach uncertainty. Concurrently, we may claim that recurring ethics breaches indicate individual researchers' lack of ability to use self-reflection as outlined above. But lack of clarity and the need for ongoing dialogue and adjustments in how research practices are taught and honed is, not exactly new or unfamiliar for us as scholars. We know how to do this. It means extending the qualitative inquiry's epistemological and ontological sensitivity to context to study design and data collection. It means paying attention to what and how we teach.

Notes

1 Instagram Direct did not allow starting conversations with just text at that time, so I would have had to send an image to reach out, and since these accounts were not following me, they would have shown up as requests not as messages.

2 Both Tumblr and Instagram allow public and private accounts, the content posted to public accounts can be accessed without having a Tumblr account on Tumblr, but one has to have the Instagram app to be able to search the public content on Instagram. A viewer does not have to become a follower to view the public content of the posters neither on Instagram nor on Tumblr.

FURTHER READING

Markham, Annette N. and Buchanan, Elizabeth (2015) 'Ethical considerations in digital research contexts', in James D. Wright (ed.), *Encyclopedia for Social & Behavioral Sciences.* Oxford: Elsevier Science, pp. 606–13.

Nissenbaum, Helen (2010) *Privacy in Context: Technology, Policy, and the Integrity of Social Life*. Stanford, CA: Stanford University Press.

Zimmer, M. (2010) '"But the data is already public": On the ethics of research in Facebook', *Ethics and Information Technology*, 12(4): 313–25.

REFERENCES

Acquisti, A. and Gross, R. (2006) 'Imagined communities: Awareness, information sharing, and privacy on Facebook', *Lecture Notes in Computer Science*, 4258: 36–58.

Aldridge, J., Medina, J., and Ralphs, R. (2010) 'The problem of proliferation: Guidelines for improving the security of qualitative data in a digital age', *Research Ethics*, 6(1): 3–9.

Bakardjieva, Maria (2009) 'A response to Shani Orgad', in Annette N. Markham and Nancy K. Baym (eds.), *Internet Inquiry: Conversations about Method*, Los Angeles: Sage, pp. 54–60.

Bakardjieva, M. and Feenberg, A. (2000) 'Involving the virtual subject', *Ethics and Information Technology*, 2(4): 233–40.

Barnes, S. B. (2006) 'A privacy paradox: Social networking in the United States', *First Monday*, 11(9). Retrieved from http://firstmonday.org/article/view/1394/1312.

Baym, N. K. and boyd, d. (2012) 'Socially mediated publicness: An introduction', *Journal of Broadcasting & Electronic Media*, 56(3): 320–9.

Beaulieu, A. and Estalella, A. (2012) 'Rethinking research ethics for mediated settings', *Information, Communication & Society*, 15(1): 23–42.

Boellstorff, Tom, Nardi, Bonnie, Pearce, Celia, and Taylor, T. L. (2012) *Ethnography and Virtual Worlds: A Handbook of Method*. Princeton University Press, Kindle Edition.

boyd, danah (2010) 'Social network sites as networked publics: Affordances, dynamics, and implications', in Zizi Papacharissi (ed.), *Networked Self: Identity, Community, and Culture on Social Network Sites*. New York: Routledge, pp. 39–58.

boyd, d. (2016) 'Untangling research and practice: What Facebook's "emotional contagion" study teaches us', *Research Ethics*, 12(1): 4–13.

Buchanan, E., Aycock, J., Dexter, S., Dittrich, D., and Hvizdak, E. (2011) 'Computer science security research and human subjects: Emerging considerations for research ethics boards', *Journal of Empirical Research on Human Research Ethics*, 6(2): 71–83.

Crossen-White, Holly L. (2015) 'Using digital archives in historical research: What are the ethical concerns for a "forgotten" individual?', *Research Ethics*, 11(2): 108–19.

Dittrich, D. (2015) 'The ethics of social honeypots', *Research Ethics*, 11(4): 192–210.

Egan, J., Chenoweth, L. I., and McAuliffe, D. (2006) 'Email-facilitated qualitative interviews with traumatic brain injury survivors: A new accessible method', *Brain Injury*, 20(12): 1283–94.

Ess, Charles (2014) *Digital Media Ethics*. Cambridge, Malden: Polity Press.

Ess, Charles and Jones, Steve (2004) 'Ethical decision-making and internet research: Recommendations from the AoIR ethics working committee', in Elizabeth Buchanan (ed.), *Readings in Virtual Research Ethics: Issues and Controversies*. London: Information Science Publishing, pp. 27–44.

Eynon, Rebecca, Fry, Jenny, and Schroeder, Ralph (2008) 'The ethics of internet research', in Nigel Fielding, Raymond M. Lee and G. Blank (eds.), *The Handbook of Online Research Methods*, London: Sage, pp. 23–41.

Felzmann, Heike (2013) 'Ethical issues in internet research: International good practice and Irish research ethics documents', in Cathy Fowley, Claire English and Sylvie Thouseny (eds.), *Internet Research, Theory and Practice: Perspectives from Ireland*. Dublin: Research-publishing net, pp. 11–32.

Fileborn, B. (2015) 'Participant recruitment in an online era: A reflection on ethics and identity', *Research Ethics*, 12(2): 97–115.

Flick, C. (2016) 'Informed consent and the Facebook emotional manipulation study', *Research Ethics*, 12(1): 14–28.

Gajjala, Radhika (2009) 'Response to Shani Orgad', in Annette M. Markham and Nancy K. Baym (eds.), *Internet Inquiry: Conversations about Method*. Los Angeles: Sage, pp. 61–8.

Hargittai, E. and Marwick, A. (2016) '"What can I really do?" Explaining the privacy paradox with online apathy', *International Journal of Communication*, 10(2016): 3737–57.

Held, Virginia (2006) *The Ethics of Care: Personal, Political, Global*. Oxford: Oxford University Press.

Kendall, Lori (2002) *Hanging Out in the Virtual Pub: Masculinities and Relationships Online*. Berkeley, CA: University of California Press.

Kirkpatrick, Marshall (2010a) *The Facebook Effect: The Inside Story of the Company That Is Connecting the World*. New York: Simon & Schuster.

Kirkpatrick, Marshall (2010b) 'Facebook's Zuckerberg Says The Age of Privacy is Over, ReadWrite', Read Write. Retrieved from http://readwrite.com/2010/01/09/facebooks_zuckerberg_says_the_age_of_privacy_is_ov

Kramer, A. D. I., Guillory, J. E., and Hancock, J. T. (2014) 'Experimental evidence of massive-scale emotional contagion through social networks', *Proceedings of the National Academy of Sciences*, 111(24): 8788–90.

Kügler, Dennis (2014) 'Individuals should be responsible for their online privacy, not governments', says survey. Retrieved from https://www.ivpn.net/blog/individuals-responsible-online-privacy-governments-says-survey.

Lawson, Danielle (2004) 'Blurring the boundaries: Ethical considerations for online research using synchronous CMC forums', in Elizabeth A. Buchanan (ed.), *Readings in Virtual Research Ethics: Issues and Controversies*. Hershey, PA and London: Information Science Publishing, pp. 80–100.

Lewis, J. and Graham, J. (2007) 'Research participants' views on ethics in social research: Issues for research ethics committees', *Research Ethics*, 3(3): 73–9.

Lomborg, S. (2012) 'Personal internet archives and ethics', *Research Ethics*, 9(1): 20–31.

Madden, Mary (2012) 'Privacy management on social media sites', *Pew Research Center's Internet & American Life Project.* Retrieved from http://pewinternet.org/Reports/2012/Privacy-management-on-social-media.aspx.

Manson, Neil C. and O'Neill, Onora (2007) *Rethinking Informed Consent in Bioethics*. New York: Cambridge University Press.

Markham, A. N. (2006) 'Method as ethic, ethic as method', *Journal of Information Ethics*, 15(2): 37–55.

Markham, Annette N. (2011) 'Internet research', in David Silverman (ed.), *Qualitative Research: Theory, Method, and Practices* (3rd edn). London: Sage, pp. 111–27.

Markham, A. N. (2012) 'Fabrication as ethical practice: Qualitative inquiry in ambiguous internet contexts', *Information, Communication and Society*, 15(3): 334–53.

Markham, Annette N. (2013) 'Dramaturgy of digital experience', in Charles Edgley (ed.), *The Drama of Social Life: A Dramaturgical Handbook*. Burlington: Ashgate, pp. 279–94.

Markham, Annette N. (2015) 'Producing ethics [for the digital near future]', in R. A. Lind (ed.), *Producing Theory in a Digital World 2.0: The Intersection of Audiences and Production in Contemporary Theory*, *Volume 2*. New York: Peter Lang, pp. 247–56.

Markham, Annette N. (2016) 'From using to sharing: A story of shifting fault lines in privacy and data protection narratives', in Bastiaan Vanacker and Don Heider (eds.), *Digital Ethics*. London: Peter Lang, pp. 189–205.

Markham, Annette N., and Buchanan, Elizabeth (2012) 'Ethical Decision-Making and Internet Research, Recommendations from the AoIR Ethics Working Committee (Version 2.0)'. Retrieved from http://aoir.org/reports/ethics2.pdf.

Markham, Annette N., and Buchanan, Elizabeth (2015) 'Ethical considerations in digital research contexts', in James, D. Wright (ed.), *Encyclopedia for Social & Behavioral Sciences.* Waltham, MA: Elsevier, pp. 606–13.

Marzano, Marco (2012) 'Informed consent', in Jaber F. Gubrium, James A. Holstein, Amir B. Marvasti and Karyn D. McKinney (eds.), *The SAGE Handbook of Interview Research: The Complexity of the Craft*. London: Sage, pp. 443–56.

Mauthner, Melanie S. (2012) '"Accounting for our part of the entangled webs we weave": Ethical and moral issues in digital data sharing', in Tina Miller, Maxine Birch, Melanie Mauthner and Julie Jessop (eds.), *Ethics in Qualitative Research.* London: Sage, pp. 157–76.

Mauthner, Melanie, Birch, Maxine, Miller, Tina, and Jessop, Julie (2012) 'Conclusion: Navigating ethical dilemmas and new digital horizons', in Tina Miller, Maxine Birch, Melanie Mauthner and Julie Jessop (eds.), *Ethics in Qualitative Research* (2nd edn). London: Sage, pp. 176–87.

McKee, Heidi A. and Porter, James E. (2009) *The Ethics of Internet Research. A Rhetorical, Case-based Process*. New York: Peter Lang.

Merriman, B. (2015) 'Ethical issues in the employment of user-generated content as experimental stimulus: Defining the interests of creators', *Research Ethics*, 10(4): 196–207.

Microsoft Trustworthy Computing (2013) 2013 privacy survey results. Retrieved from http://download.microsoft.com/download/A/A/9/AA96E580-E0F6-4015-B5BB-ECF9A85368A3/Microsoft-Trustworthy-Computing-2013-Privacy-Survey-Results.pdf.

Nissenbaum, H. (2004) 'Privacy as contextual integrity', *Washington Law Review*, 79(119): 119–59.

Nissenbaum, Helen (2010) *Privacy in Context: Technology, Policy, and the Integrity of Social Life*. Stanford, CA: Stanford University Press.

Norwegian National Committee for Research Ethics in the Social Sciences and the Humanities (NESH) (2014) *Ethical Guidelines for Internet Research*.

Orgad, Shani (2009) 'Question two: How can researchers make sense of the issues involved in collecting and interpreting online and offline data?', in Annette N. Markham and Nancy K. Baym (eds.), *Internet Inquiry: Conversations about Method*, Los Angeles, CA: Sage, pp. 33–53.

Resnick, B. (2016) 'Researchers just released profile data on 70,000 OkCupid users without permission', Vox. Retreived from: https://www.vox.com/2016/5/12/11666116/70000-okcupid-users-data-release

Richardson, J. C. and Godfrey, B. S. (2003) 'Towards ethical practice in the use of archived transcript interviews', *International Journal of Social Research Methodology*, 6(4): 347–55.

Robards, B. (2013) 'Friending participants: Managing the researcher–participant relationship on social network sites', *Young*, 21(3): 217–35.

Rodham, K. and Gavin, J. (2006) 'The ethics of using the internet to collect qualitative research data', *Research Ethics*, 2(3): 92–7.

Shelley-Egan, Clare (2015) *Ethics assessment in different fields: Internet Research Ethics*. Retrieved from http://satoriproject.eu/media/2.d.2-Internet-research-ethics.pdf.

Sveningsson, Malin (2004) 'Ethics in internet ethnography', in Elizabeth Buchanan (ed.), *Readings in Virtual Research Ethics: Issues and Controversies*. London: Information Science Publishing, pp. 45–61.

Sveningsson-Elm, Malin (2009) 'How do various notions of privacy influence decisions in qualitative internet research?', in Annette, N. Markham and Nancy K. Baym (eds.), *Internet Inquiry: Conversations about Method*. Los Angeles: Sage, pp. 69–87.

Tolich, M. (2014) 'What can Milgram and Zimbardo teach ethics committees and qualitative researchers about minimizing harm?', *Research Ethics*, 10(2): 86–96.

Treadwell, Donald (2014) *Introducing Communication Research: Paths of Inquiry* (2nd edn). Los Angeles: Sage.

Van Dijck, Jose (2013) *The Culture of Connectivity: A Critical History of Social Media*. Oxford University Press, Kindle Edition.

Walther, J.B. (2002) 'Research ethics in Internet enabled research: Human subjects issues and methodological myopia', *Ethics and Information Technology*, 4(3): 205–16.

Warrell, J. G. and Jacobsen, M. (2014) 'Internet research ethics and the policy gap for ethical practice in online research settings', *Canadian Journal of Higher Education Revue canadienne d'enseignement supérieur*, 44(1): 22–37.

Weller, Katrin and Kinder-Kurlanda, Katharina (2015) '"I love thinking about ethics!" Perspectives on ethics in social media research', Selected Papers of Internet Research, *Proceedings of ir15 – Boundaries and Intersections*. Retrieved from http://spir.aoir.org/index.php/spir/article/view/997.

White, M. (2002) 'Representations or people?', *Ethics and Information Technology*, 4(3): 249–66.

Willinsky, J. (2006) *The Access Principle: The Case for Open Access to Research and Scholarship*. Cambridge, MA: Massachusetts Institute of Technology.

Zevenbergen, Ben (2015) E-mail to the Association of Internet Researcher's mailing list, on the ethics of using hacked data, cited with permission.

Zimmer, M. (2010) '"But the data is already public": On the ethics of research in Facebook', *Ethics and Information Technology*, 12(4): pp. 313–25.