SAGE Research Methods

The SAGE Handbook of Online Research Methods

The Ethics of Online Research

The Ethics of Online Research

Rebecca Eynon Jenny Fry Ralph Schroeder

This chapter considers some of the main ethical issues that researchers are likely to encounter in Internet-related research. These issues have been discussed for some time and some guidelines for researchers are established (Markham and Buchanan, 2012); however, there is still considerable debate about the ethics of Internet research – not least because the Internet is still in a formative phase and new phenomena continue to emerge. In this chapter, we will discuss some of the major issues that have been debated and give some indication of how to go about addressing them.

One of the challenges to developing a coherent ethical approach to Internet research is that as the Internet evolves as a space for social interaction and information dissemination, the methods necessary to capture and document such activities are also emergent. Consequently, consideration of ethical issues in a context-independent manner, divorced from matters of methodology and conceptual frameworks, would be limited in scope and usefulness. In this chapter, we therefore discuss novel ethical dilemmas for Internet researchers in the context of three predominant approaches to gathering Internet-based data: use of online methods to gather data directly from individuals, analysing online interaction within virtual environments, and social media as a research laboratory. Prior to this discussion, we reflect on how ethical issues relating to Internet research might differ from research in traditional settings.

New Technology, Old and New Ethics

Ethical Governance in Traditional Research Settings

Thus far, the governance of Internet research has been heavily influenced by the well established guidelines in (offline) social research (Basset and O'Riordan, 2002). These guidelines are typically at the national level (e.g. research councils such as the Australian Research Council (ARC) and the Economic and Social Research Council (ESRC) in the UK) or committees set up at an institutional level (e.g. ethical review committees/ethics committees (UK), Institutional Review Boards (US) or Human Subjects Review Boards (AU)).

Both these mechanisms for external research governance (i.e. beyond that of the individual researcher or research group) have historical roots in the 'human subjects research model'. Three ethical concepts are at the core of institutional and professional research governance based on the 'human subjects model': confidentiality, anonymity and informed consent. These are derived from the basic human right to privacy, although these rights are interpreted differently in different jurisdictions (e.g. the EU and the US; see Reidenberg, 2000; Rule, 2007). However, these regulations originate from the medical sciences and are not always applicable to the social sciences. The human subjects research model is widely used in discussions of online research ethics, although the suitability of applying this model in some online contexts has been questioned (Basset and O'Riordan, 2002).

Institutional governance of research (the interrelationship between legal and ethical interventions) not only varies between institutions, but also from country to country (Buchanan and Ess, 2009). These range from close intervention, which in extreme cases can hinder the progress of research, to minimal guidance that relies on the self-policing of researchers. Differentiation in institutional/professional ethical rules and guidelines illustrates a tension between external (structural) governance and the freedom of self-regulation among individual researchers. Such institutional provisions do not necessarily exempt researchers from further ethical obligations and responsibilities.

Researchers have different relations with research participants and data provided by them, depending on the method and approaches used. For example, it is not uncommon for ethnographers to develop a trust relationship with the people from the communities they observe, and they often come to perceive themselves as custodians of the data they gather. Ethical practices are also shaped by personal ethical frameworks,

as well as researchers' cultural and professional ones. As Ess (2006) argues, any emphases on the rights of research participants must be considered further alongside other important rights and values – including (deontological) emphases on the importance of knowledge developed through research and (more utilitarian) emphases on research knowledge as contributing to public policy and debate, along with researchers' rights and interests in pursuing knowledge. This is a recurring issue in relation to the ethics of Internet research and closely related to considerations of 'harm' to research participants (see Ess, 2006, 2013).

There is a blurring of the boundary between ethical and legal considerations and provisions. Ess makes a useful distinction between institutional or legal requirements as against the ethical requirements that can go beyond these (Ess, 2002). In addition to the requirements set by Research Ethics Committees and professional bodies, there are also laws regarding privacy and data protection that govern research in different countries. In Internet research, however, the institutional and legal context may be uncertain because research participants may be online in any geographical context. The global reach of the Internet may thus, as Ess (2006) suggests, entail that researchers take heed of contexts which go beyond their own jurisdictions. This also applies to considerations over and above these institutional and legal requirements, such as what we might do as individual researchers out of a sense of fairness. Here, as well, it is necessary to think 'globally' because values such as privacy may be culturally specific and what is considered an appropriate balance between privacy and freedom of expression will vary between cultures (Fry, 2006; Nakada and Tamura, 2005).

New Ethics for New Settings?

Why should online research require separate or additional treatment? This 'meta-issue' has itself been a major debate that runs through the various individual topics in Internet research (Buchanan and Ess, 2009). Walther (2002), for example, has argued that many of the features of Internet research are similar to those found in other media or in existing offline research. Walther's (2002) arguments are directed against those (particularly Frankel and Siang, 1999) who argue the opposite; namely, that new rules are required for this novel setting because, to give just one example, people may misrepresent their identity online (to which Walther replies that they can also do this offline).

Despite continuing disagreements in this debate, Ess (2002) argues that there has been a convergence on the view that research ethics for online settings are not special and can be derived from the ethics for offline settings. We shall encounter a number of instances later. At the same time, we shall also argue that in some cases there are special considerations that are needed for online research, such as the changed nature of disclosure and informed consent. This arises from the increased domestication (Silverstone et al., 1992) of the Internet in everyday life and the possibilities for technical and methodological innovation.

As noted earlier, Internet research is regulated in a similar way to other areas of academic research, with institutional review boards and ethical committees, alongside professional associations providing guidance. Regardless of the specifics of online ethics, it is important to note that such review boards are not without their critics, and a number of researchers have highlighted concerns about the extent to which ethical review boards can apply a set of largely context-free guidelines in unproblematic ways, the extent to which such processes account for the messiness of real-life research, and to which such protocols pay sufficient attention to all aspects of the research process, including exiting the field site (Miller, 2013). Ultimately, researchers should not be put off engaging with the review board or assuming certain kinds of research cannot be done; however, given the innovation in this area it is likely that researchers should not rely solely on the judgements of ethical review boards or take on significant responsibility themselves (Lunnay et al., 2015). Indeed, the Association of Internet Researchers (AoIR) provided an updated set of guidelines for researchers in 2012, and part of this provides a useful framework of questions researchers should ask themselves as they consider the ethics of their research project (Markham and Buchanan, 2012: 8–10).

Sensitivity to Context

Sensitivity to context is important. The AoIR guidelines place an emphasis on this context-dependence, which entails respecting people's values or expectations in different settings. A few examples (in addition to those

provided in the AoIR guidelines) will suffice:

- Bloggers: the aim in this case is to disseminate the blogger's views, but should everything, including sensitive personal information contained in a blog, be disseminated via research?
- Search: from a legal point of view, the release of information by a search provider in anonymized form may not pose a problem, but clearly those who search don't expect to be potentially identified in relation to their search behaviour.
- Online games: the context here may be play, but even though these environments are public, is it appropriate to reveal players' names in research publications?
- Chatrooms: though a chatroom space may be public, the participants may feel they are part of a trusted community and use the space to communicate intimate details of their lives. Should consideration be given to reproducing the content verbatim in research communications and to what extent should social structures be protected from being disclosed or 'invaded' by researchers?

Internet research ethics thus need to be tailored to different contexts (Sveningsson-Elm, 2009). It may not be sufficient, for example, to stay within the strictures of copyright law (the institutional and legal requirements mentioned in the previous section) or to simply adopt the rule of 'fair use' as with offline publications (Walther, 2002; but see Ess, 2002: 3) in order to be ethically fair to research participants.

Approaches to Internet Research

The following sections deal with the particular ethical issues that arise in using online methods to gather data directly from individuals, analysing online interaction within virtual environments and social media as a research laboratory. The online methods used to gather data directly from individuals that are discussed include surveys, interviews and focus groups. In these cases, researchers use online tools to ask participants for responses to particular questions or issues. The study of online interaction in virtual environments includes participant observation and logging and visualizing the interaction between participants. The analysis of social media as a research laboratory involves capture and analysis of digital traces that people leave online, such as representation of self-identity and social interaction, alongside the use of large-scale experiments to nudge behaviour.

Use of Online Methods to Gather Data Directly from Individuals

For a detailed discussion of how to conduct interviews and surveys see Chapter 24 on interviews by O'Connor and Madge and the Internet survey section in Part IV of this Handbook. Here, we address the key ethical considerations of these online methods, which raise slightly different ethical challenges to the face-to-face context (Mann, 2003).

Benefits and Risks of Online Research

It is the investigators' responsibility to ensure, as far as they are able, that participants will not come to harm by taking part in any study. In the social sciences, psychological and physical harm to participants may be caused, for example, by research that evokes bad memories or reduces a person's sense of pride or dignity, or by cases where the anonymity of the participant is not maintained as originally agreed (Bier et al., 1996). Trying to ensure harm is not caused by the study is particularly challenging as there may well be unintended consequences of research unforeseen by the researcher (Rees, 1991: 147).

Online research is not intrinsically more likely to be harmful than face-to-face methods, yet it does pose different challenges (Kraut et al., 2004). In online research it is more difficult to assess the risk of participants coming to harm because fewer studies have been conducted from which researchers can learn, and it is harder to judge individuals' reactions to the research (e.g. if a person is getting distressed by an interview question or if a participant feels insulted or harassed by other group members in an online discussion) (Bier et al., 1996; Mann and Stewart, 2000). Strategies to try to address these issues include building a good rapport with participants, establishing 'netiquette' in group discussions (Mann and Stewart, 2000) and

providing participants with an easy way to leave the study (Hewson et al., 2003; Nosek et al., 2002).

A second issue is the potential of harm to the researchers. Given the anonymity of the Internet, researchers can come across or receive distressing information of numerous kinds. Examples include people who are contemplating suicide (Lehavot et al., 2012), people who are considering/have committed a crime, people who are bullying others or people who are grieving (Carmack and Degroot, 2014). What a researcher does with such information has ethical, and in some cases legal, implications. It is important for researchers to anticipate and assess these risks prior to beginning the study as far as is possible in order to reduce the potential of harm to themselves and their participants (for detailed advice see Stern, 2003).

Ensuring Confidentiality

Harm can also occur due to breaches of confidentiality and anonymity caused by the misuse of storing or using the data (Fox et al., 2003). Researchers have a responsibility to ensure the confidentiality of data and the privacy of participants at all stages of the process, during all interactions with the participants and when the data is transmitted and stored (Nosek et al., 2002). Given that the perceived anonymity of the Internet may encourage people to discuss topics or disclose more details than they would be willing to in face-to-face situations (Meho, 2006), researchers need to ensure that participants' perceptions of anonymity are met, or if not, made explicit to the participant (see section on informed consent later).

In terms of securely transmitting data, potential solutions include the use of encryption, use of data labels that are meaningless to anyone but the researcher, and the separate transmission of personal data and experimental data (Nosek et al., 2002). In terms of data storage, the data needs to be protected from other people accessing it or tampering with it; this can be an issue in the networked systems commonly in place in universities (Fox et al., 2003). Password-protecting computer directories, saving personal data and experimental data separately (Kraut et al., 2004), encrypting the files so no one else can read them, or coding the data in a way that reduces the likelihood of people being able to trace the data to a specific individual (Pittenger, 2003) are all possible strategies.

The issue of ensuring confidentiality whilst interacting with the participants may arise at various points throughout the research. Participants may wish to contact the researchers up to and including the debriefing stage at the end of the study, yet directly emailing the researchers may compromise anonymity in a number of ways. First, email addresses are often identifiable as they can contain names, geographical location and organizational affiliation. Although people can make use of anonymous email services to cover their identity, these are not 100 per cent effective and tend to promise 'best efforts' as opposed to true anonymity. Second, a copy of all emails is retained on the server of the sending account, any transmitting server and on the destination server and these copies are frequently retained on back up recordings for a number of years (Fox et al., 2003). These issues can be particularly problematic when certain activities are carried out online, for example if verifiable names and addresses or signed agreements are required to fulfil informed consent procedures and/or institutions require personal details when participants are rewarded for the research in the form of prizes or payment. Email should be reduced to a minimum with offline methods or alternative web-based methods utilized where appropriate, for example setting up a forum on the research website for participants to ask questions (Fox et al., 2003) and when offering prizes for participating in the research (a technique that in itself raises ethical questions), maintaining anonymity by purchasing online gift certificates and then providing the certificate number to the participant (Kraut et al., 2004).

Informed Consent

Individuals who choose to participate in any research project must do so on the basis of informed consent, where the individual understands what the goal of the research is and what they are agreeing to do, the potential risks and benefits of taking part, and have details of alternative options that may benefit them. Participants must have the option to ask anything they wish and understand that participation is voluntary and that they can withdraw at any time (Anderson, 1998). In practice, gaining truly informed consent is not straightforward in any context. The nature of informed consent changes throughout the research process and thus needs to be constantly renegotiated (Bier et al., 1996; Sin, 2005). Further, it is difficult to ascertain if

informed consent is truly given by the participant, for example problems and misunderstandings arise when potential participants do not read documents carefully or fail to ask for clarification from the researcher (see Varnhagen et al., 2005). In face-to-face contexts it is potentially easier to evaluate if the participant is fully informed about the study compared to online environments. Owing to the distance between the researcher and the participant in online settings, this is sometimes more difficult. It is harder to determine whether the participant truly understands what they are consenting to and it may take more time to gain consent because it may require more online discussions to ensure the participants fully understand the implications of participating. This additional online interaction may put participants off clarifying or asking all the questions they wish about the research (Mann and Stewart, 2000). To try to ensure participants are truly informed in online settings, techniques of increasing the readability of the document can be used (e.g. reducing the amount of text, use of subheadings and use of colour). Using quizzes to check understanding can be another means, although this extra burden on the participants increases the risk of dropout (Varnhagen et al., 2005). Despite these challenges the advantage of online consent when compared to face-to-face consent is that participants are likely to feel less pressure to enter into and remain in the study and are therefore more likely to enter and participate in the research freely.

A second important issue is verifying the participant's ability to give informed consent (Kraut et al., 2004). Verifying the ability of an individual to give informed consent is harder in online environments because it is more difficult to know whether or not the online sample includes 'vulnerable groups' (e.g. young people,[1] the elderly or people with mental health issues) and because the extent to which individuals are able or competent to give informed consent varies widely and this is more difficult to judge online. Reducing the chances of a vulnerable group (e.g. young people) being part of a research project can partly be addressed by the recruitment strategy utilized. For example, sending specific invitations to known adult participants to access a password-controlled site (Pittenger, 2003) or designing advertising materials that are unlikely to attract or interest young people when employing a more 'broad brush' strategy (Nosek et al., 2002) may help. Other options include asking for information that only adults would have, such as credit card information, although such activities can increase dropout (Kraut et al., 2004). In practice, verifying identity is really an issue only in research involving controversial topics and/or where the study presents higher risks to potential participants (Pittenger, 2003). Indeed, whether one should try and obtain online consent for high-risk studies at all is open to question (Kraut et al., 2004).

The issues considered in this section have included protecting participants from harm, ensuring confidentiality and informed consent. Indeed, these questions are perhaps becoming increasingly challenging to address where the distinction between researcher and researched is becoming more blurred in some settings, such as crowd-sourcing and sites that offer a complex mix of support in return for data (e.g. patientslikeme.com) (Janssens and Kraft, 2012; O'Connor, 2013).

It is a balancing act for researchers to ensure that participants are protected, but at the same time not placing unnecessary and excessive burdens on participants in terms of completing informed consent procedures, ensuring security, etc. (Kraut et al., 2004). Although it is impossible to predict all eventualities in online research, it is useful to always pilot test instruments and consent forms because what works in one context with one group may not work and/or may well produce different ethical questions in another situation (Meho, 2006).

Analysing Interactions in Virtual Environments

To address the questions raised by online environments, we take online virtual worlds as an example. These include social spaces where people, in the form of avatar representations of themselves, interact with each other in the virtual setting for various purposes including gaming, socialising and collaborating (see Bailenson and Schroeder, 2008).

Online Social Settings

Virtual spaces in which people interact online as avatars can be treated as social worlds. These social settings are perhaps most often akin to 'third places' (public parks, coffee shops, street corners) – places that are

neither public nor private but in between (Oldenburg, 1989).

Online social spaces exemplify the imperative discussed earlier: to be sensitive to the values and aims of people in different online settings. This sensitivity to context will involve treating different virtual worlds in different ways. There may be events or whole worlds in which people interacting online are behaving in a public way, as in a public meeting or in a virtual world that is open to all for commercial or educational purposes. It may also be, however, that certain spaces within a virtual world, such as an online church (Schroeder et al., 1998), although formally public, include interactions that should be treated as private – such as when personal details are revealed or if a whole online world is expressly designed to provide a private forum for interaction among a group that would be difficult in an offline setting (or in another virtual setting).

For offline participant observation or fieldwork there have been extensive debates in anthropology about the role of the observer and the degree to which researchers should engage in covert or overt observation and these will provide some guidance (see, for example, Angrosino and Rosenberg, 2011; Horst and Miller, 2012; Levy and Hollan, 1998), but virtual spaces present some unique challenges, which are discussed later.

The Role of The Observer

A real difference in online versus offline settings is the researcher's ability to hide completely – or lurk – in the online world. There may be a trade-off in this case between the advantages of covert observation which does not disturb the environment, and revealing one's identity as a researcher, which ensures transparency and participation, but may also lead to changed behaviour on the part of the subjects (for a particularly striking example, where the researcher became 'stalked', see Hudson-Smith, 2002). Anecdotally there have been a number of cases when many researchers descended on an online environment and there was resentment against their presence. Although ultimately the decision to disclose the presence of a researcher is down to the individual project, care needs to be taken to treat each research site with respect. The well-established rule in anthropology – to leave the field so that future researchers are not disadvantaged – must be an important consideration. There are a number of strategies that can be used to approach online communities, gain informed consent and make it clear to all participants what the researcher's role is: approaching key stakeholders for permission to research the community, using a name in the community that highlights your research status and providing a link in your online interactions in the group that communicates more information about your research for anyone to access (see also Roberts, 2015).

Studies of Online Populations

In the physical world people can be covertly recorded (as with closed-circuit television cameras), but in online worlds the possibilities of recording, reproducing and analysing interactions, especially covertly, are more powerful. This raises ethical issues because people using these environments do not necessarily expect to have all their behaviour recorded – even when theoretically giving their consent to this through the end-user license agreement when downloading the software (Chee et al., 2012).

For example, Penumarthy and Börner (2006) analysed where people moved and when they focused their attention in an online virtual world for education. This kind of recording of behaviour is unlikely to be objectionable. If, however, they had counted the number of times that avatars had engaged in particularly unsavoury behaviour, even in a public place, users might reasonably object to this kind of surveillance. There is a fine line between when data about a large online game is aggregated to reveal patterns about behaviour without violating participants' sense that they are under surveillance – and the opposite. The same applies to the analysis of small-scale groups, which can be analysed down to the granularity of the finest details of interaction (Schroeder and Axelsson, 2006).

Thus researchers will need to adapt ethical considerations to the novel technological possibilities and constraints of online virtual worlds. They will continue to face the choice between Kantian duty-based or 'deontological' ethics, with their absolute respect for the individual's aims, as against the calculation of consequentialist or utilitarian ethics, which weighs the balance of harms and benefits. Final choices are likely to be based on the ethical dispositions of the researcher, the nature of the group being researched, what

research questions are being asked and how the data will be used (see, for example, Boellstorff et al., 2012; McKee and Porter, 2009; Horst and Miller, 2012).

In terms of the uses and limits of virtual environments for experimental research, see Eynon et al., (2008) for a detailed discussion of the benefits versus the harms and, in particular, discussion of Slater et al.'s (2006) virtual reconstruction of the Milgram experiment.

Social Media as a Research Laboratory

The particular characteristics of social media and the way in which they are used provide new challenges for research ethics. Social media has the potential to be ubiquitous and large populations of users are constantly connected from multiple devices. The very 'publicness' of the sites, such as Facebook, YouTube, Twitter and Instagram, imbue interactions with a performative quality, with some individuals using multiple sites to promote themselves in the same way that a company might promote a brand. Individuals are defined by various characteristics from their friendship circles to their consumption patterns. In these spaces we observe the blurring of the distinction between private and public/the public sphere and commerce. The inhabitants of these spaces are akin to a melting pot with teenagers, university students, professionals, celebrities, grandparents, the healthy, the vulnerable, the benevolent and malevolent intermingling in the spaces that comprise the social web.

Determining potential harm in this melting pot is complex, added to which imperatives to share data, advanced data processing capacities and interest in big data by commerce and governments raise a number of new challenges for Internet researchers.

Privacy in Public Online

What unites a whole range of research in this area is the question of what constitutes a private act and how researchers might deal with the issue of 'privacy in public' (Nissenbaum, 1998). A particularly perplexing question is the extent to which researchers should take measures to protect an individual's privacy when the sources of data are publicly accessible. In this regard institutional review boards and ethics committees have proven limited (Zimmer, 2010) and in any case, as Henderson et al. (2013) argue, the moral responsibility of researchers to conduct research in an ethically robust way goes beyond the sphere of formal institutional guidelines.

In physical public environments individuals can adapt their behaviours in certain ways in order to create 'partial privacy' when an 'audience' is present – in a restaurant a couple can request a table in an out-of-the-way corner, on a train people having a conversation about a private matter can lower their voices, at a public event an individual may choose to avoid being photographed. Furthermore, members of the 'audience' can avert their 'gaze' in response to subtle social cues that indicate that 'partial privacy' is desired. In other words, privacy is managed in physical spaces through an awareness of mutual attention. To some extent the privacy settings offered by social media platforms afford the management of 'privacy in public', but compared to the subtle and complex ways in which individuals negotiate their privacy in physical public spaces they are crude and limited in their functionality. Individuals are not necessarily aware of privacy settings or think they have been evoked when in actuality they have not (Henderson et al., 2013). The implication of treating publicly accessible social media profiles and other objects as 'public' is that potential ethical dilemmas are likely to be overlooked. This is one of the reasons why there have been a number of initiatives encouraging Internet researchers to share the ethical dilemmas that they have experienced (Markham and Buchanan, 2012) and why 'rules' can be limited given the contextual nature of the ethical issues that emerge.

In law, privacy is a qualified right protected in major legislation such as the 4th Amendment to the US Constitution 1791 and the European Convention on Human Rights 1950. Data privacy laws are gaining ground and have been adopted in more than a hundred countries around the world, even if implementation in practice is lagging behind (Greenleaf, 2013). Protecting privacy is important from the perspective of a number of ideals: human dignity, individual autonomy, freedom to behave and to associate with others without the continual threat of being observed, freedom to innovate and freedom to think. These ideals reflect the

strong relationship between privacy and identity. In constructing their social media profiles individuals are also constructing, or 'performing', multiple aspects of their identity. In contrast to anonymous environments such as chat rooms, virtual worlds and Massively Multi-player Online Games, many social media sites afford targeted identity performance and as a result have been termed by some researchers as 'Nonymous' environments (Grasmuck et al., 2009). Different social media platforms afford different types of performance depending on technical and social affordances of their design. According to Grasmuck et al. (2009), for example, Facebook affords three types of targeted performance: 'Self as social actor' (implicit visual claims through photographs and wall posts); 'self as consumer' (listing cultural preferences that define a user, lists of consumption preferences and tastes, such as books, movies, music and appreciated quotes); and 'first-person self' ('about-me' entries, explicit self-description).

In conceptualizing identity performance in such environments several authors have drawn on Goffman's (1959) dramaturgical concept of 'front stage/back stage' (Grasmuck et al., 2009; Hookway, 2008; Rosenberg, 2010; Schultze, 2014). 'Front stage' is where an individual can project their 'possible hoped for selves' (Grasmuck et al., 2009: 165) and 'back stage' serves as a private space where individuals can be free from the 'scrutinizing gaze of others' (Rosenberg, 2010: 27). Photographs of 'self' and 'others', for example, are often selected 'back stage' in private somewhat unguarded moments, without a specific audience in mind and are presented front stage and viewed widely. Grasmuck et al. (2009) argue that the casual selection of photographs with close friends in mind may reveal more to a broad audience than a Facebook user would divulge in a face-to-face interaction where a sense of audience is more acute.

Privacy also plays a role in the management and maintenance of relationships, for example what an individual is willing to share with close friends or a 'neutral' professional (such as a family doctor or a counsellor) might be different to family members. In a study of YouTube, Lange (2008) observed varied levels of 'privately public' behaviour in video making and sharing. An example of this is the 'coming out' video, a recognized genre on YouTube (Thelwall et al., 2012) whereby an individual may choose to come out online, using various mechanisms to target the performance to a specific social group, for example by using a pseudonym that only close friends will recognize and partially hiding the location of a video by using limited/cryptic tags (Lange, 2008). In today's society with pervasive uses of technology, it is not feasible for an individual to expect total privacy and this is reflected in data protection legislation, which rather than being about protecting privacy per se is about giving individuals some control about the information that flows outwards from them.

The traceability of both text and non-text based data2 via Internet search engines and the use of data mining tools raises the question of how to represent the data when it comes to dissemination and publication, for example should verbatim quotes be used and how should images be represented? Anonymization typically involves the removal of personally identifiable information3 such as full name, residential address and date of birth. Whereas this level of anonymization might be sufficient to protect the privacy of an individual in a standalone dataset, Internet search and data mining tools enable the re-identification of an individual via triangulation. Users of social media platforms typically have profiles across multiple platforms and it is this unique overlapping feature that makes the ease of re-identification particularly problematic (Narayanan and Shmatikov, 2009).

Researchers are exploring innovative ways to anonymize data. Markham (2012) has been developing techniques related to 'fabrication' in representing qualitative data, which involves developing typical examples or scenarios that are comprised of composite objects that collectively mask individuals. This approach is controversial, however, and Markham (2012) describes the difficulty experienced by colleagues in getting a manuscript accepted for publication where they had created composite blogs, rather than using verbatim quotes from actual blogs. Some researchers report experiencing the opposite issue of how to acknowledge those individuals who wish to be disclosed and have their data attributed to them (Tilley and Woodthorpe, 2011).

Social media environments are also unique in that they afford the collective construction of identity through sharing, tagging, commenting and automatic feeds from 'friends'. Individuals may work tirelessly on 'front stage' management, but family and friends might reveal glimpses of 'back stage'. This collective element of identity construction further complicates distinguishing between private/public and is just one example of why private/public should be treated as a continuum rather than a straightforward dichotomy (Rosenberg,

2010). As Schultze and Mason (2012: 303) argue, individuals may be performing their identities according to 'situated assumptions of privacy'. Consequently, researchers cannot assume that an entire website is 'public' as a consequence of the intentions of platform developers, Terms of Service agreements or the technical capabilities of privacy settings.

The extent to which an individual may consider lack of mutual attention a breach of privacy will depend on a number of factors, such as cultural attitudes to privacy, individual privacy attitudes (Westin, 1967) and gender. It is possible that there may be gender differences in the extent to which individuals consider a lack of awareness of 'mutual attention' problematic because it is known that there is a gender dimension to privacy in the context of social media (Thelwall, 2011). Heightened social media privacy concerns amongst women are related to the increased likelihood that they will be victims of malevolent online behaviours such as cyberbullying, cyberstalking (Thelwall, 2011) and revenge porn. It is arguable that 'mutual attention' is at the heart of the ethical dilemma faced by researchers when using publicly available social media data. Of course, informed consent signals 'mutual attention' to research participants, but whether or not a social media profile actually represents a human participant, and thus evokes the human subjects model, has been the topic of much debate in the related literature.

Schultze and Mason (2012) propose the introduction of three new principles to the human subjects model: 'degree of entanglement', 'extent of interaction/intervention' and 'expectation of privacy', each of which could be measured on a sliding scale. The more private the activity and the space within which it occurs, the more the source of the data is seen as a human subject, and the more public, the more the source of the data is seen as the author of a text. Where there is a combination of a high-degree of 'entanglement' and a high expectation of privacy, then this would be indicative that seeking informed consent would be good practice. Consideration of the use of 'big data' for academic versus commercial research purposes can further highlight the complexity of some of the issues raised earlier regarding the use of data gathered from social media platforms, particularly with regard to anonymization, re-identification and recontextualization of data.

Analysing Big Data

Analysis of big data using data collected from social networking sites (and other digital traces) has been among the fastest growing areas of research in recent years. Here, we review two important studies, both using Facebook. The first was a study of the social networks on Facebook (Lewis et al., 2008). The study identified a number of patterns among the 'tastes' from the 'ties among Facebook friends from among students at a "private college in the Northeast U.S."' (Lewis et al., 2008: 331), thus ensuring the anonymity and privacy of those concerned. However, it took Zimmer (2010) little effort to figure out that the study had in fact been done at Harvard University, thus potentially being able to re-identify the subjects and creating a number of ethical issues that would be deemed unacceptable. Among the lessons from this episode is that there needs to be strict ways of ensuring anonymity and thus privacy. The study also raised a number of other issues, including whether consent is needed with a study using data from a private company and also whether it would be possible to make the data available to other researchers for re-use (which was intended in this case, but did not happen because of the concerns raised).

This issue of access to data was also raised in the second study, the 'social contagion' study (Kramer et al., 2014). This research took the form of a 'naturalistic' experiment, dividing almost 700,000 randomly selected Facebook users into two groups and filtering the content of their 'timelines' (their personalized news feeds) such that one group had more positive words and others were unchanged. This type of research, which analysed 3 million posts and 122 million words, certainly fits the definition of big data (Schroeder, 2014a). The finding was that users with more positive words in their feeds subsequently produced more positive words of their own, an important and large-scale confirmation of the 'social contagion' effect whereby what others do affects our own behaviour.

As already mentioned, one set of issues raised by this study is whether the privileged access to research data is afforded to some researchers by this type of study – one of the authors worked at Facebook – which creates unequal access to research materials. The added question here concerns the replicability of the study, which is an essential feature of scientific research and which is made impossible with this kind of proprietary data. The second set of issues revolves around whether carrying out this type of research violates Facebook

users' privacy. Here the reply from Facebook was that the study did not breach the 'terms and conditions' that users had signed and so the study did not break laws, even though Facebook has also said that it should have handled 'communicating the study' better (Guardian, 2014a). A related question was whether the study should have been subject to the scrutiny of a university 'institutional review board' and thus a different kind of 'consent' apart from the legal terms and conditions required by using the site. In this respect the response was that the academic researchers only carried out the analysis whereas the data collection issues were handled by Facebook. Another response came from research ethicists in an article in Nature (Meyer, 2014) who took a stance against the idea that this kind of experiment carried out with a commercial company was ethically unacceptable; instead, these ethicists argued that imposing strictures on this type of commercial research would only drive it underground to the detriment of advancing publicly available knowledge.

There is a third set of issues, which relates to big data and the very idea of undertaking large-scale research which essentially manipulates people (where 'manipulate' simply means doing something to them). These have been discussed by one of the authors of this chapter elsewhere (Schroeder, 2014b), but they are also broader than questions of research ethics and of law and, in this single study, which concern big data research generally. Although the ethical and legal questions around individual cases will likely be dealt with, this larger question will require a wider debate in society. The larger question concerns the fact that big data methods, often based on social media or other online behaviours, are becoming more widespread. If scientific knowledge about human behaviour based on these methods becomes more powerful, then it will also be able to manipulate people more powerfully. Academic researchers are typically not interested in this type of manipulation (but see the Lewis et al., 2008 study), but perhaps with this knowledge they are becoming the handmaidens of those who are (such as digital media companies, in this case Facebook).

These are difficult questions that relate to the role of science and technology in society as a whole and what the value of large-scale experiments and other 'manipulations' should be. Moreover, such studies should not be ruled out altogether. In some cases, if a greater good can be achieved – for example, if we could find out people's attitudes towards climate change and conserving energy by means of this type of big data study – surely the benefits could outweigh the costs of analysing how behaviour might be manipulated as long as there is minimal or no direct risk involved. It should also be remembered that this type of big data research does not always require a private sector platform – Wikipedia has also been used for big data research (Schroeder and Taylor, 2015) and the data are open to all for replication. Governments are also engaging in big data methods to influence people.

In any event, the spectre of using large-scale online platforms to potentially sway peoples' beliefs hovers uneasily above this type of study. And while academics may largely be disinterestedly concerned with greater knowledge about life online, their research may support non-academic uses which can now, more powerfully, alter peoples' hearts and minds. In the years to come, more people will be online more often and produce vastly greater amounts of digital traces. Academic researchers will need to think hard, beyond ticking boxes on research ethics forms and beyond collaborating with non-academic sources of these data, about the extent to which this type of research improves the world, while at the same time avoiding societal concerns not just about 'big brother' but also of a 'brave new world'. It can be remembered that this novel warned not so much of surveillance, about which there has recently been much discussion, but of a future in which manipulating minds was embraced by the public and seen as benefitting society, much as Facebook argued that the social contagion experiment served to improve the users' experience (Guardian, 2014b).

Triangulation of Datasets and Third-Party Reuse

Advancements in the development of resources and tools available on the Internet make the triangulation and third-party reuse of data much more likely. While a standalone dataset may preserve anonymity and privacy, new capabilities for aggregating and combining data could jeopardize such ethical integrity by enabling profiles of individuals to be constructed through triangulation (Oboler et al., 2012). As Kitchin (2013: 264) notes, big data is 'highly resolute, providing fine-grained detail on people's everyday lives', which is why Narayanan and Shmatikov (2009) were able to de-anonymize an anonymous Twitter graph by using a generic re-identification algorithm and triangulating the Twitter data with data available from Flickr. In doing so, Narayanan and Shmatikov (2009) illustrated that in the context of big data anonymization alone is not sufficient to protect privacy. Indeed, social media users have very little control over their data despite the

different levels of privacy settings offered by social media sites (Puschmann and Burgess, 2014).

The tracking capabilities built into the very infrastructure of the Internet itself and the tools being developed to exploit the gathering and aggregation of fine-grained data on a large scale mean that the role of researcher as custodian and gatekeeper of personal data becomes radically altered. Tools that enable data to be easily reused by third parties and recontextualized in novel ways undermine the notion of 'context', for example the norms, values and beliefs of groups within online social settings (see earlier section), as a heuristic for developing ethical practices that are socially and culturally appropriate. Gleibs (2014: 359) argues that in the context of big data researchers move away from a legal contract that demarcates private/public space and instead the right to use data becomes a complex psychological contract that needs to take into account perceptions and expectations about individuals' control over the flow of information that relates to them. Reuse and the emergent practice of data profiling by third parties reduces choice for both researcher and research participant in terms of how data is represented and how it travels through media and across actors. The researcher, therefore, may no longer be able to foresee all of the consequences and potential harm of their research, which has implications for 'informed consent' where it is deemed appropriate in large-scale studies. Gleibs (2014) discusses various mechanisms for technically implementing informed consent in the context of social media data and argues that through such technical mechanisms social media users should be 'reminded of the use of data for research and that data created on SNS can be mixed with other sources for new discoveries'(p. 366).

Furthermore, as noted earlier, the vast quantities of social science data being generated by the Internet are of significant commercial value (Schroeder, 2014b), with social media data being an important area of economic growth based on privileged access to data that provides insight into consumer behaviour (Puschmann and Burgess, 2014). Consequently, social science data generated and used by academic researchers may travel beyond the professional boundaries of the social science disciplines and into the private sector, whose practices in relation to ethical considerations may be governed by legal jurisdiction, rather than ethical codes of practice. Indeed, the business models and development goals of social media providers and commercial third parties differ, with commercial third parties wanting access to vast quantities of real-time data that enable them to model and predict user behaviour on an unprecedented scale (Puschmann and Burgess, 2014).

There is a school of thought in the application of novel technologies to social science that is on the side of pushing the boundaries until there is a legal intervention. For example, placing responsibility for privacy onto end-users to be aware of and understand the Terms of Service for the social media site. This can be problematic given that in many cases the technology and its capability for triangulating and reprocessing data is so novel that often legal intervention is lagging behind. Therefore, practice is often pushing the boundaries of ethical frameworks and legal interventions, and the analysis of social media data on a real-time or 'near real-time' basis, in particular, is likely to push these boundaries. As Thelwall and Stuart (2006) point out, some techniques are inherently illegal in their mechanisms. For example, web crawling is illegal because crawlers make permanent copies of copyrighted material without the owner's permission.

The rise in the use of big data for academic research thus raises the question of the extent to which Internet researchers should be concerned with the collection and use of potentially harmful data, given that we cannot anticipate all the ways in which it might be reused and by whom. In terms of research excellence, social scientists have always been encouraged to consider only collecting sufficient data to satisfy the immediate objectives of their research, but with the Internet the capabilities for collecting and storing data are so vast that the practicality or desirability of maintaining such practices in the context of new technologies, methods and techniques are brought into question. As illustrated in the earlier case study of Facebook, the opposite can also happen whereby data generated in the private sector can become available in the public domain and be used as a resource for academic research. Earlier high-profile examples of this were the public release of the Enron and AOL email databases on the Internet. The Enron database was released in the interests of transparency and accountability as part of a legal investigation, but it was not sufficiently anonymized and was retracted after two weeks (see Eynon et al. 2008).

Conclusion

One of the key challenges in guiding ethical decision making in Internet-based research is in its global

reach and the necessity to respect and incorporate diverse cultural practices, ethical governance and legal frameworks. What is different about Internet-based research in contrast to research in the offline world is that the research object is no longer clearly delineated by national boundaries and protected by national research governance. The emergence of novel methods across disciplines also brings an interdisciplinary focus to bear on the Internet as an object of study and challenges existing instruments of research governance that have traditionally been focused along disciplinary dimensions.

At the same time, the online world affords new modes of human interaction and related ethical practices are shaped by the researchers' objectification of those being researched, for example whether individuals participating in an online chatroom are perceived as research subject, research participant, artist (Bruckman, 2002) or author (Bassett and O'Riordan, 2002). There is also a potential convergence between research and commercial data on the Internet and a blurring of boundaries in crowdsourcing and wider public engagement initiatives that lead to questions over who is the researcher and who is researched. Development of aggregator tools and services have led to the informatization of data, whereby data acquires additional value beyond the immediate research context. Consequently, the potential for third-party reuse is much greater than in the offline world. Data sharing and reuse are institutional imperatives with many funding bodies now mandating the submission of datasets to data archives and repositories upon the completion of funding. This entails the development of practices and techniques to anonymize highly sensitive data, with some data being easier to anonymize than others (Markham, 2012; Saunders et al., 2015).

The context of social interactions in online worlds is also important to bear in mind. If we take the position that traces of interaction on the Internet are public and should be treated as such, for example participants have no rights to privacy considerations, how do we address the issue that online bodies and forms of expression have offline instantiations? To what extent do we need to protect these from harm? As tools for tracing social structures become more sophisticated, so too do our capabilities for triangulating data and getting a more holistic view of participants lives. Whereas participants may choose to draw a boundary between their online and offline worlds, and may in fact be online in order to escape the strictures of the offline world, the technologies currently being developed do not necessarily respect such boundaries. The question for us as social scientists is to what lengths we should go to discover people's intentions. This, of course, means that we must disclose ourselves as researchers, which could alter the kind of results we were hoping to obtain. In the context of digital research, participants may be, but are not necessarily, already in the Internet domain. We cannot therefore simply assume that they have chosen to be online or what their intentions are in being there. Again this raises the question of whether 'public in everyday life' is equivalent to 'public on the Internet'. All the while, the 'human subjects' research model remains in place and, as Bassett and O'Riordan (2002) have argued, what is required now is the trying and testing of different models of research governance.

The issues that we have raised in this chapter go beyond responsibilities towards a particular set of research participants and have implications for social, political and ethical aspects of social science research. A significant proportion of the world will not be represented in online research and researchers need to ask whether this is ethical. Certain groups are likely to be under-represented and are therefore less likely to gain benefits from participating. Such an emphasis on the interests of the information-rich may reinforce existing societal divisions (Mann, 2003; Eynon et al., 2009). Researchers have an ethical responsibility to ensure that the research they carry out is of high quality and that conclusions drawn from it can be inferred from the data collected (Pittenger, 2003). Finally, one obvious strategy to adopt under conditions with yet-to-emerge norms that have been sketched here is to be explicit about the ethical decisions that are made in order that others can learn from and debate the issues that arise when reporting findings.

Notes

1. Acquiring informed consent for participation in research by children is subject to legal frameworks and regulations that differ from country to country. See, for example, Wiles et al. (2005) for a discussion of the UK context.

2. An interesting example of how images can be searched using the service Google Images is provided by Henderson et al. (2013).

3. Definition of personally identifiable information according to EU/US data protection legislation.

References

Anderson, G. (1998) Fundamentals of Educational Research.
2nd edn.
London: Falmer Press.

Angrosino, M. and Rosenberg, J. (2011) 'Observations on observation: continuities and challenges'. In N.K. Denzin and Y.S. Lincoln (eds.), The SAGE Handbook of Qualitative Research.
4th edn.
London: Sage Publications. pp. 467–78.

Bailenson, J. and Schroeder, R. (2008) 'Research uses of multi-user virtual environments'. In R. Lee, N. Fielding and G. Blank (eds.), The SAGE Handbook of Online Research. London: Sage Publications, pp. 327–42.

Basset, E. and O'Riordan, K. (2002) 'Ethics of Internet research: contesting the human subjects research model', Ethics and Information Technology, 4 (3): 233–47.

Bier, M., Sherblom, S. and Gallo, M. (1996) 'Ethical issues in a study of Internet use: uncertainty, responsibility, and the spirit of research relationships', Ethics and Behavior, 6 (2): 141–51.

Boellstorff, T., Nardi, B., Pearce, C. and Taylor, T.L. (2012) Ethnography and Virtual Worlds: A Handbook of Method. Princeton, NJ: Princeton University Press.

Bruckman, A. (2002) 'Studying the amateur artist: a perspective on disguising data collected in human subjects research on the Internet', Ethics and Information Technology, 4 (3): 217–31.

Buchanan, E. and Ess, C. (2009) 'Internet research ethics and the Institutional Review Board: current practices and issues', Computers and Society, 39 (3): 43–9.

Carmack, H. J. and Degroot, J. M. (2014) 'Exploiting loss? Ethical considerations, boundaries, and opportunities for the study of death and grief online', OMEGA – Journal of Death and Dying, 68 (4): 315–35.

Chee, F. M., Taylor, N. T. and de Castell, S. (2012) 'Re-mediating research ethics: end-user license agreements in online games', Bulletin of Science, Technology & Society, 32 (6): 497–506.

Ess, C. (2002) 'Introduction to Special Issue on Internet research ethics', Ethics and Information Technology, 4 (3): 177–88.

Ess, C. (2006) 'Ethics and the use of the Internet in social science research'. In A. Joinson, K. McKenna, T. Postmes and U.-D. Reips (eds.), Oxford Handbook of Internet Psychology. New York, NY: Oxford University Press. pp. 487–503.

Ess, C. (2013) Digital Media Ethics. Cambridge: Polity Press.

Eynon, R., Fry, J. and Schroeder, R. (2008) 'The ethics of internet research'. In R. Lee, N. Fielding and G. Blank (eds.), The SAGE Handbook of Online Research. London: Sage Publications, pp. 23–41.

Eynon, R., Schroeder, R. and Fry, J. (2009) 'New techniques in online research: challenges for research ethics', Twenty-First Century Society, 4 (2): 187–99.

Fox, J., Murray, C. and Warm, A. (2003) 'Conducting research using Web-based questionnaires: practical, methodological, and ethical considerations', International Journal of Social Research Methodology, 6 (2): 167–80.

Frankel, M. and Siang, S. (1999) 'Ethical and legal aspects of human subjects research on the Internet'. Available at: http://www.aaas.org/sites/default/files/migrate/uploads/report2.pdf (accessed 1 June 2016).

Fry, J. (2006) 'Google's privacy responsibilities at home and abroad', Journal of Librarianship and Information Science, 38 (3): 135–9.

Gleibs, I. H. (2014) 'Turning virtual public spaces into laboratories: thoughts on conducting online field studies using social network sites', Analyses of Social Issues and Public Policy, 14 (1): 352–70.

Goffman, E. (1959) The Presentation of Self in Everyday Life. New York, NY: Doubleday.

Grasmuck, S., Martin, J. and Zhao, S. (2009) 'Ethno-racial identity displays on Facebook', Journal of Computer-Mediated Communication, 15: 158–88.

Greenleaf, G. (2013) 'Data protection in a globalised network'. In I. Brown (ed.), Research Handbook on Governance of the Internet. Cheltenham, UK: Edward Elgar, pp. 221–59.

Guardian (2014a) 'Facebook apologises for psychological experiments on users'. 2 July. Available at: http://www.theguardian.com/technology/2014/jul/02/facebook-apologises-psychological-experiments-on-users (accessed 15 June 2016).

Guardian (2014b) 'Facebook reveals news feed experiment to control emotions'. 29 June. Available at:

http://www.theguardian.com/technology/2014/jun/29/facebook-users-emotions-news-feeds (accessed 15 June 2016).

Henderson, M., Johnson, N., and Auld, G. (2013) 'Silences of ethical practice: dilemmas for researchers using social media', Educational Research and Evaluation: An International Journal on Theory and Practice, 19 (6): 546–60.

Hewson, C., Yule, P., Laurent, D. and Vogel, C. (2003) Internet Research Methods: A Practical Guide for the Social and Behavioural Sciences. London: Sage Publications.

Hookway, N. (2008) 'Entering the blogosphere: some strategies for using blogs in social research', Qualitative Research, 8 (1): 91–113.

Horst, H. and Miller, D. (2012) Digital Anthropology. London: Berg.

Hudson-Smith, A. (2002) '30 days in active-worlds – community, design and terrorism in a virtual world'. In R. Schroeder (ed.), The Social Life of Avatars: Presence and Interaction in Shared Virtual Environments. London: Springer. pp. 77–89.

Janssens, A. C. J. and Kraft, P. (2012) 'Research conducted using data obtained through online communities: ethical implications of methodological limitations', PLoS Medicine, 9 (10): e1001328.

Kitchin, R. (2013) 'Big data and human geography', Dialogues in Human Geography, 3 (3): 262–7.

Kramer A., Guillory J. and Hancock J. (2014) 'Experimental evidence of massive-scale emotional contagion through social networks', Proceedings of the National Academy of Sciences, 111 (24): 8788–90.

Kraut, R., Olson, J., Banaji, M., Bruckman, A., Cohen, J. and Couper, M. (2004) 'Psychological research online: report of the Board of Scientific Affairs' Advisory Group on the Conduct of Research on the Internet', American Psychologist, (59) 2: 105–17.

Lange, P. (2008) 'Publicly private and privately public: social networking on YouTube', Journal of Computer-Mediate Communication, 13: 361–80.

Lehavot, K., Ben-Zeev, D. and Neville, R. E. (2012) 'Ethical considerations and social media: a case of suicidal postings on Facebook', Journal of Dual Diagnosis, 8 (4): 341–6.

Levy, R. I. and Hollan, D. W. (1998) 'Person-centered interviewing and observation'. In H. R. Bernard (ed.), Handbook of Methods in Cultural Anthropology. Walnut Creek, CA: AltaMira, pp. 333–64.

Lewis, K., Kaufman, J., Gonzalez, M., Wimmer, A. and Christakis, N. (2008) 'Tastes, ties, and time: A new social network dataset using Facebook.com', Social Networks, 30(4): 330–42.

Lunnay, B., Borlagdan, J., McNaughton, D. and Ward, P. (2015) 'Ethical use of social media to facilitate qualitative research', Qualitative Health Research, 25 (1): 99–109.

Mann, C. (2003) 'Generating data online: ethical concerns and challenges for the C21 researcher'. In M. Thorseth (ed.), Applied Ethics in Internet Research. Trondheim, Norway: NTNU Publications Series No. 1, pp. 31–49.

Mann, C. and Stewart, F. (2000) Internet Communication and Qualitative Research: A Handbook for Researching Online. London: Sage Publications.

Markham, A. (2012) 'Fabrication as ethical practice: qualitative inquiry in ambiguous internet contexts', Information, Communication and Society, 5 (3): 334–53.

Markham, A. and Buchanan, E. (2012) 'Ethical decision-making and internet research. Recommendations from the AoIR Ethics Working Committee (Version 2.0)'. Available at: http://aoir.org/reports/ethics2.pdf (accessed 15 June 2016).

McKee, H. A. and Porter, J. E. (2009) The Ethics of Internet Research: A Historical, Case-Based Process. New York, NY: Peter Lang.

Meho, L. (2006) 'E-mail interviewing in qualitative research: a methodological discussion', Journal of the American Society for Information Science and Technology, 57 (10): 1284–95.

Meyer, M. (2014) 'Misjudgements will drive social trials underground', Nature, 511 (7509): 265.

Miller, T. (2013) 'Messy ethics: negotiating the terrain between ethics approval and ethical practice'. In J. McClancy and A. Fuentes (eds.), Ethics in the Field: Contemporary Challenges. Oxford: Berghahn.

Nakada, M. and Tamura, T. (2005) 'Japanese conceptions of privacy: an intercultural perspective', Ethics and Information Technology, 7 (1): 27–36.

Narayanan, A. and Shmatikov, V. (2009) 'De-anonymizing social networks', IEEE Symposium on Security & Privacy. Oakland, CA. Available at: http://www.cs.utexas.edu/~shmat/shmat_oak09.pdf (accessed 15 June 2016).

Nissenbaum, H. (1998) 'Protecting privacy in an information age: the problem of privacy in public', Law and Philosophy, 17: 559–96.

Nosek, B., Banaji, M. and Greenwald, A. (2002) 'E-research: ethics, security, design and control in

psychological research on the Internet', Journal of Social Issues, 58 (1): 161–76.

Oboler, A., Welsh, K. and Cruz, L. (2012) 'The danger of big data: social media as computational social science', First Monday, 17 (7).

O'Connor, D. (2013) 'The apomediated world: regulating research when social media has changed research', The Journal of Law, Medicine & Ethics, 41 (2): 470–83.

Oldenburg, R. (1989) The Great Good Place. New York, NY: Marlowe and Co.

Penumarthy, S. and Börner, K. (2006) 'Analysis and visualization of social diffusion patterns in three-dimensional virtual worlds'. In R. Schroeder and A.-S. Axelsson (eds.), Avatars at Work and Play: Collaboration and Interaction in Shared Virtual Environments. Netherlands: Springer, pp. 39–61.

Pittenger, D. (2003) 'Internet research: an opportunity to revisit classic ethical problems in behavioural research', Ethics and Behavior, 13 (1): 45–60.

Puschmann, C. and Burgess, J. (2014) 'The politics of Twitter data'. In K. Weller, A. Bruns, J. Burgess, M. Mahrt and C. Puschmann (eds.), Twitter and Society. New York, NY: Peter Lang, pp. 43–54.

Rees, T. (1991) 'Ethical issues'. In G. Allan and C. Skinner (eds.), Handbook for Research Students in the Social Sciences. New York, NY: Falmer Press, pp. 140–52.

Reidenberg, J. R. (2000) 'Resolving conflicting international data privacy rules in cyberspace', Stanford Law Review, 52 (5): 1315–71.

Roberts, L. D. (2015) 'Ethical issues in conducting qualitative research in online communities', Qualitative Research in Psychology, 12 (3): 314–25.

Rosenberg, A. (2010) 'Virtual world research ethics and the private/public distinction', International Journal of Internet Research Ethics, 3 (12): 23–37.

Rule, J. (2007) Privacy in Peril: How We Are Sacrificing a Fundamental Right in Exchange for Security and Convenience. New York, NY: Oxford University Press.

Saunders, B., Kitzinger, J. and Kitzinger, C. (2015) 'Participant anonymity in the Internet age: from theory to practice', Qualitative Research in Psychology, 12 (2): 125–37.

Schroeder, R. (2014a) 'Big data: towards a more scientific social science and humanities?'. In M. Graham and W. H. Dutton (eds.), Society and the Internet. Oxford: Oxford University Press, pp. 164–76.

Schroeder, R. (2014b) 'Big data and the brave new world of social media research', Big Data and Society, July–December: 1–11.

Schroeder, R. and Axelsson, A.S. (eds.) (2006) Avatars at Work and Play: Collaboration and Interaction in Shared Virtual Environments. Netherland: Springer Science & Business Media.

Schroeder, R. and Taylor, L. (2015) 'Big data and Wikipedia research: social science knowledge across disciplinary divides', Information, Communication and Society, http://dx.doi.org/10.1080/1369118X.2015.1008538

Schroeder, R., Lee, R. M. and Heather, N. (1998) 'The sacred and the virtual: religion in multi-user virtual reality', Journal of Computer-Mediated Communication, 4 (2).

Schultze, U. (2014) 'Performing embodied identity in virtual worlds', European Journal of Information Systems, 23(1): 84–95.

Schultze, S. and Mason R. O. (2012) 'Studying cyborgs: re-examining Internet studies as human subjects research', Journal of Information Technology, 27: 301–12. doi:10.1057/jit.2012.30.

Silverstone, R., Hirsch, E. and Morley, D. (1992) 'Information and communication technologies and the moral economy of the household'. In R. Silverstone and E. Hirsch (eds.), Consuming Technologies: Media and Information in Domestic Spaces. London: Routledge, pp. 15–31.

Sin, C. (2005) 'Seeking informed consent: reflections on research practice', Sociology, 39 (2): 277–94.

Slater, M., Antley, A., Davison, A., Swapp, D., Guger, C., Barker, C., Pistrang, N. and Sanchez-Vives, M.V. (2006) 'A virtual reprise of the Stanley Milgram obedience experiments', PLoS ONE, 1 (1): e39. doi:10.1371/journal.pone.0000039.

Stern, S. (2003) 'Encountering distressing information in online research: a consideration of legal and ethical responsibilities', New Media and Society, 5 (2): 249–66.

Sveningsson-Elm, M. (2009) 'How do various notions of privacy influence decision making in qualitative internet research'. In A. Markham and N. Baym (eds.), Internet Inquiry: Conversation about Method. Thousand Oaks, CA: Sage Publications, pp. 69–87.

Thelwall, M. (2011) 'Privacy and gender in the social web'. In S. Trepte and L. Reinecke (eds.), Privacy Online. Berlin: Springer-Verlag, pp. 251–65.

Thelwall, M. and Stuart, D. (2006) 'Web crawling ethics revisited: cost, privacy and denial of service', Journal of the American Society for Information Science and Technology, 57 (13): 1771–9.

Thelwall, T., Sud, P. and Vis, F. (2012) 'Commenting on YouTube videos: from Guatemalan Rock to El Big Bang', Journal of the American Society for Information Science and Technology, 63 (3): 616–29.

Tilley, L. and Woodthorpe, K. (2011) 'Is it the end for anonymity as we know it? A critical examination of the ethical principle of anonymity in the context of 21st century demands on the qualitative researcher', Qualitative Research, 11(2): 197–212.

Varnhagen, C., Gushta, M., Daniels, J., Peters, T., Parmar, N., Law, D., Hirsch, R., Takach, B. and Johnson, T. (2005) 'How informed is online consent?', Ethics and Behavior, (15) 1: 37–48.

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

Westin, A. (1967) Privacy and Freedom. New York, NY: Atheneum.

Wiles, R., Heath, S., Crow, G. and Charles, V. (2005) 'Informed consent in social research: a literature review. NCRM Methods Review Papers NCRM/001'. Available at: http://eprints.ncrm.ac.uk/85/1/MethodsReviewPaperNCRM-001.pdf (accessed 15 June 2016).

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

http://dx.doi.org/10.4135/9781473957992.n2