Digital Ethics

Book Section

How to cite:

For guidance on citations see FAQs.

© 2016 Rebecca Whiting and Katrina Pritchard

https://creativecommons.org/licenses/by-nc-nd/4.0/

Version: Accepted Manuscript

Copyright and Moral Rights for the articles on this site are retained by the individual authors and/or other copyright owners. For more information on Open Research Online’s data policy on reuse of materials please consult the policies page.

oro.open.ac.uk
Chapter 35

Digital Ethics

Rebecca Whiting and Katrina Pritchard

Corresponding author: katrina.pritchard@open.ac.uk

This chapter will be published in The SAGE Handbook of Qualitative Business and Management Research (Eds. Catherine Cassell, Ann Cunliffe and Gina Grandy) due out in 2017. Please note this is not the final edited version.

Abstract

We consider digital ethics, the moral principles or rules of behaviour that govern and guide qualitative internet research from its inception to publication and the curation of data. A number of overarching tensions are identified: flux and uncertainty regarding these rules, the type and status of ethical guidance, the lack of transparency around ethics in practice and the ‘problematic’ nature of qualitative research. Four key debates are then explored namely determining human participation, working with the private/public dilemma, seeking informed consent and from whom and deciding on anonymization or attribution. Looking ahead at the future directions, we consider the areas of researcher role and protection. We conclude with how we might channel the reflexivity that qualitative researchers already embrace when engaging with issues of validity, and use this for an ‘ethics as process’ case-based approach which features ongoing reflexive questioning of ethical considerations throughout the research cycle.
Key words: ethics, human participation, anonymization, informed consent, private/public, confidentiality, internet research, reflexivity
Introduction

In the context of qualitative research, ethics are generally defined as the moral principles or rules of behaviour that guide research from inception through to publication and the curation of data. The main concern is to minimize risk of actual or potential harm whilst ensuring the maximum benefit of the research (ESRC, 2015a). Our focus here is ‘digital ethics’ referring to ethical considerations in research that uses and/or studies web-based activities and settings. It should be noted that this is variously referred to as ‘e-research’, ‘internet-mediated’, ‘web’, ‘digital’ or ‘online’ research. For simplicity we use the term internet research as an umbrella categorization. Our primary interest is in the application of digital ethics to qualitative research in the field of business and management though we draw from useful sources across a range of disciplines.

The aim of the chapter is to provide an overview of the key ethical issues in internet research: assessing what is human participation; when informed consent is required and from whom; the dilemma of what is public or private on the internet; and decisions about anonymization or attribution. We situate these issues within a discussion of overarching tensions. These encompass flux and uncertainty surrounding both internet technology and the guidance developed to address it; the type and status of such guidelines; a lack of transparency in how digital ethics are actually addressed in practice and the particular challenges of what is seen as ‘problematic’ qualitative research.

Our starting point is to scope the ethics context for researchers in the field of business and management. We then examine the literature that has considered the ethics of internet research. This is used to highlight various overarching tensions before examining the key issues in greater detail. In doing so, we consider a selection of different approaches that a qualitative researcher
might like to consider when reviewing ethical guidance and practice to address these dilemmas in their own research. It is not our intention to summarize or indeed review all possible different ethics guidelines that might conceivably apply. Examples are included of available material, ranging from the formal (published ethics guidance) to the more informal (published research that may nevertheless become accepted practice). We conclude with suggestions as to how qualitative researchers might address the practicalities of dealing with ethical dilemmas when carrying out internet research in the field of business and management.

**Scoping ethics in business and management**

For business and management researchers, there is unlikely to be a single specific body or ethics code that applies; we must always draw on guidance from a developing range that may extend beyond this field to include our own institutions, professional and funding bodies and specialist digital ethics guidance. For example, a number of bodies support and represent the academic community and its engagement with professional activities in the field of business and management. These generally serve particular geographical constituencies, though the events they organize and their members’ work may be international. Member activities may include research, teaching, learning and practice. Researchers can check whether the relevant body for their location publishes a code of ethics or not. For example, the Academy of Management (AoM, United States) and the British Academy of Management (BAM, in combination with the Chartered Association of Business Schools and the Higher Education Academy) both issue such guidance. Their codes are widely drawn to address the full range of academic activities (not just research) though neither specifically addresses digital ethics. In the case of BAM, the guidance is a starting point to encourage members to take responsibility for their own ethical practice; this is supplemented by links to additional resources such as the ethical codes of other bodies in the fields of social anthropology, sociology and psychology (including the British Psychological Society) and funders (including the ESRC).
On the other hand, as at April 2016, neither the European Academy of Management nor the Australian and New Zealand Academy of Management issue such guidance. Bodies focusing on the delivery of teaching in Management and Business Schools, for example, the European Institute for Advanced Studies in Management, the European Foundation for Management Development and the Association to Advance Collegiate Schools of Business, also do not publish research ethic guidelines though they may offer codes of practice on the need to include ethics in programme curricula and doctoral training. As at April 2016, none of these refer to internet research.

It should be noted however that new codes are emerging as organizations update their guidelines to reflect the increasing significance of internet research. This covers a ‘wide range of activities’ (British Psychological Society, 2013: 1) across both qualitative and quantitative approaches. Digital research tools and techniques are nascent, developing alongside the internet as it becomes more socially interactive and where content is constantly changing. Not only are traditional research methods re-imagined, new forms of data are made available to study (for example, tweets namely 140 character microblogs, YouTube videos, below-the-line reader comments on news stories), new collaborations are formed and new tools are made available (including proprietary ones by social media platforms such as Twitter). Procter et al., (2013) provide an example of an agreement between Twitter and the Guardian Newspaper which involved the former providing a Twitter corpus to the latter and its academic collaborators. The Twitter search application program interface (API) however only provides 1% of actual traffic (Burnap et al., 2015). Procter et al., (2015) discuss the development of tools specifically to address the limitations of the availability and analysis of social media data by researchers.
This changing landscape has implications for ethics in business and management. One aspect relates to the scope of institutional ethics committees. Traditionally, committees have often sought to determine whether data to be collected are considered primary or secondary; the former are generally seen as requiring greater ethical scrutiny. But we need to consider if this classification is still meaningful when digital technology allows for a ‘redistribution of methods between researchers, devices, information and users, in online environments’ (Marres, 2012: 161) and researchers explore new ways of accessing textual and visual data. See Pritchard and Whiting, in this volume, for a brief discussion of the ethical and practical considerations of using stock photos, one genre of visual data in internet research.

Primary data is usually defined as that obtained directly and actively by the researcher from participants. The interaction between participant and researcher is critical rather than its mode. In internet research this could be via an online interview. In primary data analysis, the same researcher analyses this data. With secondary data, there is no direct interaction between researcher and research subject. Existing information is created or becomes available to the researcher. In internet research, data could be press reports of employee use of blogs and social media (Richards, 2012). In secondary data analysis, the researcher who analyses the data is not the one who created it. However, while such data is not generated by direct research interaction with participants, the researcher still needs to be active in collecting (or, we might say, constructing) it. Richards (2012) describes this process, specifically monitoring and recording press reports via his own blog. So in the context of internet research, typologies of primary and secondary data begin to blur. We suggest that it is more useful to look beyond the primary/secondary dualism (Bishop, 2007) and to consider the more explicit question of whether or not the research involves human participants (discussed in detail below).
The scope of regulated internet activities also extends beyond research. Ethical considerations were first mentioned in academic publications in the nineties (Sharf, 1999) and more formal guidelines for researchers appeared soon after, for example, from the Association of Internet Researchers (Ess, 2002). Potentially problematic research behaviours were identified such as ‘lurking’ (reading online posts without contributing) and ‘harvesting’ (collecting and using the words of others, even in an ostensibly public domain and with consent, but without contributing to the purpose of the chat room or forum). Lurking was, however, seen as a methodological advantage by others (Mann and Stewart, 2000). At the same time, the responsibilities of other organizations such as news media and internet service providers also came under scrutiny (Broadcasting Standards Commission, 1999). Legislation now covers issues such as data protection, freedom of information, the ‘right to be forgotten’, freedom of speech, and whistle-blowing. Indeed, for what was originally conceived as an open and unregulated space, the internet has become the focus of a great deal of policy, law and governance (Palfrey, 2010). So whilst our focus is on ethical rather than other types of regulatory codes, these can never be entirely separate and as researchers we need to take all these into account. This may pose particular challenges for multi-national research teams in business and management. For example, in the area of data protection the European Union recognizes an Internet Protocol Address (the numerical label assigned to a computer or other device which is connected to a network that uses the internet for communication) as personally identifiable data but the US does not (Buchanan and Hvizdak, 2009). As we discuss below, this is significant in decisions by researchers about what ‘personal’ information about their participants to redact in research publications or data sharing.

**Overarching Tensions**

Four overarching tensions provide a backdrop to a consideration of digital ethics. These concern the state of flux and uncertainty surrounding ethical codes in this field, the type and status of such
guidance, the particular challenges presented by qualitative research and the lack of transparency that flows from an absence of explicit examination of ethical issues in much published internet research. With an increase in such research, we found the emergence of specialist fields focusing on a particular feature such as Twitter use (Zimmer and Proferes, 2014) but also an emergent field specifically examining digital ethics. The topic has been the subject of several books (Ess, 2009; McKee and Porter, 2009; Whiteman, 2012), book chapters (Anderson and Jirotka, 2015; Buchanan, 2010; Buchanan and Zimmer, 2015; Eynon et al., 2008) and a special issue in the journal Qualitative Research in Psychology (Morison et al., 2015). We now explore the overarching tensions noted above and consider how they provide a backdrop to an examination of digital ethics.

Flux and uncertainty

Rapidly changing technology leaves researchers and institutional ethics committees wrestling with specific situations not directly anticipated by such codes. Guidelines relating specifically to digital ethics are themselves also subject to flux in response to new technologies or new regulations. For example, the UK’s Economic and Social Research Council (ESRC) has published two sets of ethics guidelines within 5 years (in 2010 and 2015) and the Association of Internet Researchers (AoIR) has published two within 10 years (2002 and 2012). What is regarded as ethical also evolves over time (Schultze and Mason, 2012). Consequently the field of digital ethics feels unfixed in comparison to non-digital ethics, for example, where ethical practice for ‘standard’ interviews is quite well known. Most research bodies address digital ethics to some degree but how they do so is less clear and more varied. Quite contrasting stances may be adopted by different institutional ethics committees to similar issues and may differ from published guidance from particular governing bodies. We note from our own experience that it can be difficult to get institutional ethics committees to engage with the issue of whether human participation is involved in internet research. What is clear however is that new technologies such as wearable cameras (Mok et al., 2015), 3D printing, digital video and
photography, use of avatars, and new forms and sources of online data such as leaked material from organizations (Michael, 2015) will continue to present new challenges. This brings us to what we can expect with regard to the type and status of ethical guidance available.

Type and status of guidance

The state of technological flux raises issues about the type of ethical guidance that is required or desirable. Where specific guidance is issued that directly addresses digital ethics, this can give rise to tensions as to its status vis à vis previous more general ethics guidance. Some institutions now specify which takes precedence. For example, the British Psychological Society (BPS) specifies that its Ethics Guidelines for Internet-Mediated Research (British Psychological Society, 2013) is supplemental and subordinate to the Society’s Code of Human Research Ethics (British Psychological Society, 2014) and to its overarching Code of Ethics and Conduct (British Psychological Society, 2009). Further, generic ethical guidelines tend to assume that researchers can anticipate forms of data and modes of participant engagement, describing in advance how to ensure participants’ rights are secured. This is increasingly difficult to achieve in an era of open access and data sharing (possible future use of data by other researchers) and given we cannot predict how participants themselves might in future engage with the internet (Saunders et al., 2015). Likewise, institutions cannot be expected to anticipate all future settings for the ethical guidance they issue. In similar vein, the ESRC identifies internet research as an area where ethics committees might need to consult independent experts for their guidance on research proposals (2015a).

Significantly, recent guidelines from the AoIR (Markham and Buchanan, 2012), the BPS (British Psychological Society, 2013) and the ESRC (2015a) all recommend a contextualized, continual approach to ethical appraisal for internet research. This involves a highly reflexive approach from the
researcher who is required to consider ethics not just at the outset of the project but throughout the research lifecycle. It also involves being sensitive to the specific internet context. We consider this continual, contextualized approach in more detail below. Likewise authors in this area reject the idea of a straightforward formula for conducting ethical internet research but rather emphasize the researcher’s responsibility to consider guidance within the context of their own research and ethical practice (Whiteman, 2012). The extent to which such consideration is readily available as a resource for other researchers is the issue we explore next.

Lack of transparency

Although we identified a literature in this area, there is a general absence of consideration of ethics issues in published work across a range of academic disciplines. For example, in the field of information systems (IS) Schutze & Mason note ‘currently insufficient discussion [...] about the ethics of Internet research’ (2012: 302), criticising how much published internet research does not explicitly engage with ethical issues. In their useful summary of academic research using Twitter data, Zimmer & Proferes (2014) report that only 16 of the 380 studies they analysed mention ethical issues relating to research design and data collection (and when they did, they adopted a variety of conflicting positions). These authors attribute this to early dominance of social media research by computer science and related disciplines which have historically been outside the scope of ethical review boards because their methods were not traditionally seen as involving human participants (Carpenter and Dittrich, 2011). This is not to say that subsequent researchers have failed to address ethical issues but rather that this is not reported in their publications. This absence and resulting lack of transparency of course simply exacerbates the other tensions we have outlined, and may present particular issues in qualitative research as we consider next.
‘Problematic’ qualitative research

We noted an early view that qualitative internet research poses particular challenges, from concerns over quality issues to wider ethical considerations. For example, its ‘highly personal’ probing (DeLorme et al., 2001: 271) is seen as requiring greater reliance on researcher integrity than quantitative research such as surveys. This is not the first context in which qualitative research is seen as more inherently risky and problematic than quantitative studies (Cassell and Symon, 2011). Qualitative methods, such as those involving close textual analysis, are seen to raise particular ethical challenges as we discuss in our section on ‘anonymity or attribution’ below. For example, a potential solution to inadvertent disclosure of the source of data by direct quoting is through ‘cloaking’. This involves changing quotations by paraphrasing the material to prevent readers from identifying the original site by placing text in a search engine. But this is inappropriate where the chosen methodology is conversation or discourse analysis because of the need to use text in its original form (British Psychological Society, 2007), as we discuss further below. The situation is further complicated by research practice and protocols developed within primarily quantitative studies (for example, COSMOS, 2015) that may affect what is accepted as ethical practice in internet research but which may be impractical or inappropriate for some qualitative research.

These four overarching tensions set the background for our consideration of the key debates in the area of digital ethics.

Key debates in digital ethics

We now turn to examine in more detail key debates taking place within digital ethics today. Throughout this chapter, we touch on the ethical guidelines of three institutions (AoIR, BPS and ESRC) selected from a range of bodies that business and management researchers might need to
negotiate in the course of their research. When starting our own internet research in 2011, we reviewed guidelines from a range of bodies and decided that the AoIR guidance was the most advanced in terms of digital ethics. This made it relevant to our project which we carried out in the UK and started in a department of organizational psychology. Whilst we do not suggest that the selected three are necessarily representative of similar bodies in other countries, they usefully illustrate a contrasting range of approaches to the issue of digital ethics, including how internet research is defined. As discussed above, all researchers will have to negotiate similar combinations of ethical guidance that might apply to the particular circumstances of themselves and their research. In addition to formal ethics guidelines, specific research projects may publish statements of their ethical practice. We have therefore additionally considered one such statement relating to social media research (COSMOS, 2015). This statement is partly driven by the need to abide by the terms of use of their data providers, terms geared more towards developers than researchers (specifically, Twitter’s Developer Policy and Developer Agreement), as required by the ESRC (2015a). Project ethics statements are a practical way for researchers to provide details of their own ethical engagement which may be adopted by others as good practice or critiqued through publication and citation. We consider aspects of these various guidelines and statements as we now turn to address the key debates in more detail.

*What is public or private on the internet?*

This question goes to the heart of a range of ethical issues facing internet researchers today and largely determines whether material can be used as data in research, particularly without explicit consent. The starting point is that although the internet is widely regarded as a public space, there are parts which users may conceive as private. The challenge for internet researchers is that much engagement with the internet takes place simultaneously in places regarded as private, such as the home, and in public, such as an open discussion forum (British Psychological Society, 2013). As
Barnes explains, ‘sitting at home alone typing into a computer may feel like a private exchange. However, once private information is posted on the internet, it becomes available for others to read. We have no control over who can read our seemingly private words’ (Barnes, 2006). She suggests this confusion may be partly a function of the design and sign-up processes of particular websites that may imply a degree of privacy to otherwise fully public sites. Usage of such sites is often bound by notoriously complex terms and conditions regarding the status of material posted (Hull, 2015).

Determining what is public or private therefore forms a critical part of the context for considering digital ethics. It is recognised that these are contested terms and that the distinction between private and public is more of a spectrum than a binary divide. For example, Sveningsson Elm (2009) argues that online spaces should be considered along a continuum of public to private, reflecting the grey area between the two.

Much focus is therefore on determining how to classify the internet context along this spectrum for the purpose of assessing ethical choices for the researcher. Some argue that posting almost any material on the internet can be considered public behaviour and therefore subject to the same ethical guidelines as other behaviours in a public space (Gosling and Mason, 2015). This accords with an established tradition that observing behaviour in public spaces (pre-internet, these would have been physical locations) does not require advance notice that research is being carried out or consent from those present (Roberts, 2015). Others reject this, invoking consideration of social norms and practices in determining whether such observation and use are acceptable; they argue that internet users may have reasonable expectations of privacy and of their behaviour not being made the subject of research (Gosling and Mason, 2015). An alternative view regards online postings as not involving human participants (Rodriquez, 2013), an angle we explore in the next section of this chapter.
So how might a researcher determine whether a particular internet location is a private or public space and if the material there can be used as data in research? In contrast to other aspects of ethical considerations, this is one area where researchers have begun to detail their decision-making about the private/public status of their research data (Roberts, 2015). Although a comprehensive review of the literature on this topic is beyond the scope of this chapter, we have identified a number of different approaches which we briefly examine.

One view is where researcher understanding of what is public is driven by what technology makes it possible to access (Schultze and Mason, 2012). This might involve assessing whether the site in question requires registration (Schotanus-Dijkstra et al., 2014). But on its own this would represent quite a narrow approach, one which focuses more on the space and less on behaviour. It risks ignoring norms and practices which apply to the internet as much as other social spaces. The majority of studies now additionally include more nuanced tests that seek to contextualize the internet setting such as the relative vulnerability of the population being studied (Holtz et al., 2012), the degree to which the material has already been publicly viewed (Marcus et al., 2012), the sensitivity of the topic (Trevisan and Reilly, 2014) and the intended audience (Fleischmann and Miller, 2013).

Another aspect of this contextualization is exploring whether users of a particular internet location perceive it to be a public or private space and crucially whether they might expect a researcher to access user-generated material and use it for research. Here, research on the privacy and usage expectations of different groups of internet users is helpful. Without it, ‘we risk constructing an understanding of Internet research ethics that is not supported by reality’ (Dawson, 2014: 430). A survey specifically looking at user views reported a mix of opinions with some considering that no...
online space is truly private (Beninger et al., 2014). Research on user privacy understanding and expectation (Paechter, 2013; Roberts, 2015; Schultze and Mason, 2012) suggests some users may adapt behaviours to fit with their perceptions of what is private or public in different internet settings. Schultze and Mason (2012) however conclude that users of internet chatrooms and online forum regard their communications as private (even on a website that its creators intended as being in the public domain). Most users in these cases considered that research without informed consent was unacceptable (see, for example, Chen et al., 2004; Hudson and Bruckman, 2004). The challenge for qualitative researchers in business and management is that most empirical examples are drawn from health sciences where ethical concerns, for example, in relation to perceived vulnerability to harm through research, may be higher. This is consistent with ethical guidance which imposes a greater obligation on the researcher to protect as the vulnerability of participants increases (Markham and Buchanan, 2012).

Schultze and Mason (2012) helpfully list characteristics for use in assessing expectations of internet privacy: these are group size, communicative purpose, social status of the community, intended audience for contribution, community membership policy, norms and expectations, and content storage and accessibility. These bring together many issues debated here and may provide a workable framework to guide researchers in this area. Recognising the constructed nature of a concept like privacy is also useful. Hull’s (2015) examination of privacy self-management online positions privacy as an individual commodity that internet users are willing to trade for other market goods. This usefully reminds us that privacy in online settings may be given up (‘traded’, in Hull’s terms) for the convenience of accessing a website that allows sharing information with family and friends but not necessarily for the convenience of researchers seeking online data.
The private vs. public dilemma is also specifically addressed in some ethical guidelines. The ESRC suggests that full ethics review is potentially required for internet research ‘when the understanding of privacy in these settings is contentious where sensitive issues are discussed’ (2015a: 10). This highlights the difficulty for institutions surrounding the public/private distinction. Some provide a guiding framework (such as the BPS’s advice to consider the ‘potentially damaging effects for participants’ of using the material) rather than attempt precise categorizations (British Psychological Society, 2013; Markham and Buchanan, 2012). This is not straightforward either as Markham and Buchanan neatly summarize: ‘the uniqueness and almost endless range of specific situations defy attempts to universalize experience or define in advance what might constitute harmful research practice’ (Markham and Buchanan, 2012: 7). Reviewing internet-mediated research, the BPS (2013) suggests researchers consider variable interpretations of public and private, advising significant time be invested in detailed consideration of the particular empirical circumstances. Thus it is incumbent on researchers to consider to what extent someone sharing information via the internet may expect to be subject to academic investigation and publication. This involves considering when, how and where the material was posted and its subsequent use and reuse across the web. So for example an individual who actively encourages distribution of their material might be regarded differently from someone attempting to restrict access even if both are publically visible. Further consideration would need to be given in cases where an individual has subsequently retracted information.

Are we dealing with human participants?

A research decision to identify and label participants, respondents or research subjects is significant not least since it implies particular ethical commitments and power relationships (Amis and Silk, 2008). It also identifies those who are not participants, thus setting up a number of other categories including ‘researcher’. Online data produced by individuals (for example, material on websites, postings to online communities) forms the basis for what has been described as ‘passive’ online
The key debate is whether the use of such data should be conceptualized as human subject research (requiring ethical consideration) or, as we debate in our Introduction, merely as secondary analysis (Bradley and Carter, 2012).

More generally, the internet, particularly the interactive web, is seen to have increased the entanglement between users and the technologies they employ (van Doorn, 2011). People’s identities become hybrid ‘simultaneously performed in physical and digital spaces [...] as Web 2.0 applications inherently entangle them with the digital material that results from their online activities’ (Schultze and Mason, 2012: 302). These authors problematize whether we can distinguish between the results of what people do in these digital spaces (send emails, upload photos, create profiles, comment on online media stories, write blogs) and their physical human identity. It is suggested that the two are inextricably intertwined (Boyd, 2006) and their virtual presence online are ‘psychological extensions and projections of each user’s identity’ (Schultze and Mason, 2012: 302).

Schultze and Mason (2012) suggest an interesting parallel with archaeology. Initially this discipline engaged with the material remains of human cultures (including long-deceased human bodies and their related artefacts). Over time, however, it developed from a field that prioritized the development of a taxonomy of found objects to one that encompassed the ‘rights, interests, sensitivities, cultures and attributes of the past inhabitants of research sites and, especially, with the rights and concerns of these ancient parties’ contemporary surrogates’ (2012: 309). These authors set out a framework to assess whether you are dealing with a human participant in an internet study. This involves rating from low to high on three dimensions (degree of entanglement, interaction/intervention and expectation of privacy). Their framework has been described as
‘purposefully conservative’ and likely to result in more research being classified as involving human participation than would be the case within disciplines such as IS (Zimmer, 2012: 313), which illustrates how ‘human participation’ is a rather contested term. In summary, this debate highlights the extent to which a ‘human participants’ approach is the most suitable model to apply to internet research when deciding issues of digital ethics, with some authors questioning the appropriateness of its application (Bassett and O’Riordan, 2002).

Do we need to obtain informed consent and, if so, from whom?

One starting point is whether the research involves collecting data in active or passive ways, the former where researchers engage directly with participants in, or about, online spaces and the latter where there is no direct interaction with participants (Morison et al., 2015). Both approaches require consideration of ethical issues but that of obtaining informed consent is generally more likely with active methods. With passive research approaches, where there is no direct intervention or interaction between researcher and participant, the research may be judged as not involving ‘human participation’ and fall outside the requirement for formal application to an institutional review board or ethics committee. For example, Langer and Beckman (2005) contend that message board contributions are analogous to readers’ letters to newspapers and do not require notice or consent to be used as data.

But determining whether we need to seek consent in internet research is closely related to considerations of what is private or public, as debated above, and by no means straightforward. For example, the COSMOS project adopts a ‘Big Data’ approach to the study of social media data across various social and policy contexts. Their ethics statement (COSMOS, 2015) states that in research outputs they do not directly quote individual Twitter users without their informed consent.
They do however directly quote from Twitter accounts maintained by public organizations (such as government departments, law enforcement and local authorities) without seeking prior informed consent (paragraph 11). The statement relates to a particular research project and is not guidance as such, though if cited and adopted it could take on the status of accepted practice. There are a number of points to note. First, the statement rejects the notion that by having a public Twitter stream, an individual has given implied consent to its use in research. Second, the practice relates to what happens in research publications and not during data analysis. Third, the approach involves making a judgment call on the status of the Twitter account in question (is it maintained by an individual or by a public organization? what is a public organization?). Fourth, where consent is not possible, the project team are able to represent the content of tweets in alternative ways (such as topic clustering, word clouds or themes). In our view, this practice may make it difficult for those qualitative researchers unable to rely on alternative ways of presenting Twitter content where close textual analysis is required, as we discuss further below. Significantly perhaps, COSMOS have thus far published work applying social network analysis and sentiment analysis which rely on meta data and textual analysis approaches which do not require the identification of individual tweets in publications (for example, Burnap et al., 2015).

An additional challenge is from whom we seek consent. How might a researcher acquire informed consent from someone who wants to remain anonymous (Markham, 2005)? This should be possible through contacting the online persona of the individual though, as Paechter (2013) points out, the researcher should check the person they are dealing with is an adult and not a child. This may be less of a consideration in online sites relevant in the field of business and management. What if the internet site involves a number of different contributors? In the context of a divorce support forum, Paechter (2013) took the view that the site owner was able to give consent to research using the site’s postings on behalf of the community but that consent to interviews should be given separately.
and individually. This importantly reminds us that internet research is often combined with offline data and methods. This combined information can threaten to increase the risk to a participant’s privacy and potential harm, making informed consent more desirable.

Lastly, thinking about consent as an ongoing process becomes particularly complex in terms of foreseeing possible future uses of internet material if transferred to data archives. In the context of Twitter, one particular challenge is when tweets (or whole Twitter accounts) are later deleted by their users: this poses difficult issues for researchers depending on when this occurs (during collection, analysis or after publication). COSMOS define these as ‘high risk’ data which are precluded from publication under the Twitter Developer Agreement. Data may also include retweets which have been retweeted from a private restricted Twitter account into the public domain which is also seen as problematic (Meeder et al., 2010). This has prompted debate as to whether it is appropriate to archive Twitter data for future research given accounts include personally identifiable information including geo-locational data about the user; likewise whether researchers should respect a user’s desire to opt-out or have unwanted tweets removed from any archiving function (Zimmer and Proferes, 2014). As before, the focus of debate centres on the privacy expectations of the Twitter users such as the extent to which they were aware of the fully public nature of their tweeting activity.

Should we anonymize or attribute?

Deciding whether internet material should be anonymized or attributed in research is closely linked to views about the status of online material, the stage in the research process, whether passive online qualitative research involves human participants or is just secondary data analysis (Bradley and Carter, 2012). It also depends on whether those who post material are seen as research
participants or authors (Beaulieu and Estalella, 2012). As noted above, some regard online messages as equivalent to readers’ letters to a newspaper (Langer and Beckman, 2005) or as literary works (Kozinets, 2010). This would mean it would be inappropriate to anonymize them or change any pseudonyms. Indeed, in this view it would be important to attribute the material to however the authors have self-identified themselves. A researcher might however need to consider who has ownership of the words and associated copyright. This might be the author, the online community, or the owner of the site on which it is posted (Roberts et al., 2008). So the answers to these questions would determine whether the issue was one of obtaining informed consent or complying with copyright (or both). But as we have already seen, this may be overly simplistic without a fuller consideration of other contextual matters. So once again there are grey areas which add complexity when we consider the handling of publicly accessible material.

We have discussed elsewhere in relation to our own research (Pritchard and Whiting, 2012) the process of ethical review which resulted in a focus on what we judged to be ‘public’ areas of the internet (for example, those not requiring membership or registration). Our research involves consideration of a range of web-based textual and visual data which represents a critical communicative context for understanding constructions of age and ageing at work (see Pritchard and Whiting, in this volume). Our study involved the systematic collection of Web 2.0 material about age at work using internet tools in a daily automated search process over 150 days during 2011/2. Data included online news with reader comments, reports, images such as stock photos, blog posts, and tweets. A week’s data was sampled to assess the range of data types, issues and voices. From this, more detailed research questions were developed including how generations are enrolled in the construction of age-related employment issues (Pritchard and Whiting, 2014) and how understandings of gendered ageing are constructed and interpreted via stock photos (Pritchard and Whiting, 2015).
We retained whatever names were adopted by the internet user (not attempting any anonymization) in the data collection, management and analysis stages of our research. This seemed critical to our understanding of our data and the role of different voices within it. However, in our publications we have maintained a broad categorization of sites, sources and individuals to ensure anonymity (Pritchard and Whiting, 2014, 2015). This follows recommendations that online identities should be treated with the same care (re anonymity and confidentiality) as ‘real’ names (British Psychological Society, 2013).

A further specific concern is the traceability of data via search engines, even if anonymized. This is problematic for qualitative researchers who rely heavily on quotes to illustrate their analysis. Both the AOIR (Markham and Buchanan, 2012) and the BPS (British Psychological Society, 2013) suggest ‘cloaking’ quotes; the practice of subtly altering the original, providing this fits within the research design principles. This has significant implications and may create more problems than it solves. Firstly, search engines often list alternatives for any search, so even an extensive rewording could result in the source being identified. Secondly, a rephrased quote might inadvertently direct the search to an incorrect source or a source which has also used this text. Third, cloaking will not be appropriate where the chosen methodology is conversation or discourse analysis because of the need to use text in its original form (British Psychological Society, 2007). For these reasons we do not deploy cloaking but this practice is currently receiving a good deal of attention and, along with many of the issues raised so far, will require further debate.

**Going forward: Future Directions**
We now turn to consider likely future developments within digital ethics, including two key areas of further research. First is the matter of enhancing understanding of the researcher role and issues of researcher protection in internet research. These are comparatively undeveloped but we see them as an essential part of the second area, namely the adoption of a wider reflexive process of digital ethics to address the currently limited transparency in publications.

**Researcher role and protection**

With academics increasingly encouraged to participate in social media to enhance research impact (Brent Zook, 2015), this raises the wider issue of roles where internet researchers also engage with the web in other capacities. In our own research, blogging and tweeting have resulted in our online content appearing in the alerts we set up to collect data for our research; potentially we as researchers then become the ‘researched’. Whiteman (2012) notes how researchers move between the roles of ‘observer and ‘participant’ in the online field and the difficulty in maintaining separation between these in qualitative research. This theme is developed by Paechter (2013) in analysis of her hybrid insider/outsider status in respect of her research site (a divorce support website), in which she also participated in a personal capacity. Her account describes the challenge of keeping professional and personal identities separate in this ethnographic research. Such dual roles within an online community positions the researcher as an insider (a regular member of the forum) and simultaneously as an outsider (a researcher observing and analysing the community). Paechter (2013) and Barratt (2012) usefully discuss how they handled the disclosure required in their online research settings. Paechter used a pseudonym in her personal capacity, her actual first name in her researcher role and her hybrid identity was only made explicit to the site’s management group. This involved an ethical responsibility not to mix these two roles so that ‘people do not become afraid that things said to [her] in [her] pseudonymous role will be incorporated into the research’ (Paechter, 2013: 82), part of a wider duty to act with integrity and good faith. Barratt
(recruiting participants from various online forums of which she was, sometimes, a ‘lurking’
member) set up a dedicated project website and provided links to this and her university website as
part of a legitimacy building process. She reiterates the importance of displaying technical and
cultural competence (such as language and humour) to participants as well as a shared history of
personal use of online settings (Barratt, 2012).

There are also issues of the safety and protection of the researcher. In one sense, our greatest
protection is through enacting ethical behaviour. But within current ethical debates we find little
attention given to researcher protection, notwithstanding ethical procedures should protect
researchers as well as participants (ESRC, 2015a). Internet research may increase a researcher’s
digital footprint, highlighting their identity to those who might seek to influence or engage with the
research topic. This will depend on the mode of digital engagement, for example, using an
institutional or commercial email address, maintaining a public research blog or deploying other
social media. Our initial engagement was tentative but over time we expanded both our personal
and project-specific use of the web (Pritchard and Whiting, 2012). This raised our research profile,
providing opportunities for engagement both online (such as sharing outputs) and offline (such as
invited presentations). Further, maintaining a research presence provides a contact for those who
may want to find out more or enquire as to whether they feature in our data. Overall our research
presence has not posed any additional risk than would be encountered during our day-to-day use of
the internet. Moreover with increased debate about research impact, such online engagement
seems likely to increase. However, dual or hybrid roles for researchers can problematize
classifications of active or passive research. The unexpected challenges they present within internet
research also underline the need for ongoing reflexivity on the part of the researcher in respect of
both ethical considerations but also with regard to methodological adjustments.
Our second area concerns the call for greater ethical reflexivity (Roberts, 2015) and greater transparency in ethical considerations in published work involving qualitative internet research (Dawson, 2014; Saunders et al., 2015). This might include, for example, presenting ‘practical illustrations of the challenges ... encountered in order to open up a discussion about how to manage anonymity issues in the face of the increasing accessibility of online information’ (Saunders et al., 2015: 135). Likewise this approach needs to encompass ‘forward reflexivity’ (Pritchard and Whiting, 2012: 350) as researchers need to anticipate issues such as how anonymity may be compromised by a participant’s own behaviour (Saunders et al., 2015). Papers that debate ethical issues with which researchers have grappled are very useful, particularly when they provide solutions which meet with local ethical approval (Sharkey et al., 2011). The field of health research is particularly helpful here, providing a useful pointer for qualitative researchers in business and management. As Dawson (2014) points out, this obligation extends to editors and reviewers to hold authors to account to provide this detail. As ever with qualitative research, though, we fear that word or page limits imposed by journals may militate against a move to accommodate yet further methodological detail.

All of this suggests that digital ethics cannot be addressed as a one-off tick-box exercise. One useful concept here is ‘ethics as process’ which ‘views ethics as relationally dynamic and revisited frequently over the course of a study’ (Sharkey et al., 2011: 753). This approach prioritizes the relationship between participant and researcher though this can present a challenge where there is no direct contact between the two (as in our own study). Nevertheless we consider this to be a promising framework for re-thinking digital ethics. We also recognize this as familiar territory for many qualitative researchers given the strong parallels with both reflexivity and methods of analysis in qualitative research which tend towards the circular and iterative (Morison et al., 2015).
Given how the more considered specialist guidelines and statements (AoIR and BPS) suggest a case-based approach with ongoing reflexive questioning, we suggest that ‘ethics as process’ would involve considering ethical issues at each stage in the research process: design, data collection / construction, analysis, publication, dissemination, and data sharing. Whilst the latter has been less prominent in business and management than say health sciences, it is likely to increase particularly in publicly funded research. For example, the ESRC expects data sharing and requires data management plans for research council funded work in the UK (ESRC, 2015b). In proposing a processual approach, we suggest these stages are not necessarily wholly discrete. Qualitative research is often an iterative process and there will almost certainly be temporal overlap between stages, particularly where data collection is an ongoing and partially automated procedure as in our own research. In other words, decisions made at one stage may need to be re-visited later and may affect subsequent iterations of the same process. Nor do we imply that once certain ethical considerations (such as anonymity) are addressed at one stage they are finished with, since they may re-appear as the internet context itself changes. We see this as part of a wider trend towards issues such as informed consent being seen as ongoing ethical obligations (Anderson and Jirotka, 2015) requiring negotiations over time with participants rather than a one-off tick box event (Mok et al., 2015), as reflected in guidance from the UK Data Service. This provides that ‘Consent for participation in research, for data use and for data sharing can be considered at different stages of the research, giving participants a clearer view of what participating in the research involves and what the data to be shared consist of’ (UK Data Service, 2015).

Conclusions
Rapid technological developments continue to present new challenges in the area of digital ethics, a situation made more difficult by a relative lack of transparency and reflexivity in much of the published internet research. We echo calls for an approach that moves away from reliance on fixed rules to one that highlights the need for continuing and reflexive practice. As qualitative researchers in the field of business and management, we can then work within an iterative framework that goes beyond ‘procedural ethics’ to examine ‘ethics as process’ and which continues throughout the research cycle.

We suggest that, far from being ‘problematic’, qualitative research is well suited to this approach given the parallel with demonstrating validity. Traditionally this might involve constructing a narrative ‘to explain...what was done and why’ in order to demonstrate ‘why and how the [research] findings are legitimate’ (Phillips and Hardy, 2002: 79). As Mason (2002: 190) concludes, the best way to demonstrate validity is to ‘explain how you came to the conclusion that your methods were valid’. Applied to digital ethics, and in the absence of more definitive rules, we suggest that working through the type of questions identified in say the AoIR guidelines in order to explain how you came to the conclusion that your methods were ethical may be the best way to demonstrate ethical compliance. We acknowledge that this may be an imperfect tool given that we as researchers have to accept our lack of perfect self-knowledge and that we may also be using such a narrative as a ‘device to invite trust’ (Hardy et al., 2001: 534). Perhaps all one can do, following Leitch & Palmer (2010), is to be explicit, as best one can, about one’s decision-making as a researcher in terms of the ethical decisions we have made.

References


Author Biographies

You can read more about Rebecca and Katrina’s research at: http://ageatwork.wordpress.com/ and follow them on twitter @ageatwork

Rebecca Whiting is a lecturer in the Department of Organizational Psychology at Birkbeck, University of London. Her research interest is in taken-for-granted aspects of the contemporary workplace, including the discursive construction of work identities (for example, the older worker), concepts (such as age, gender and work-life balance), and the ways in which work is organized. She is also
interested in the particular challenges of qualitative e-research and visual methodologies. Rebecca can be contacted at r.whiting@bbk.ac.uk

Katrina Pritchard is a Senior Lecturer in the Department of People and Organisations, OUBS, The Open University. Her research interests include the construction of identity and professional knowledge, digital media and devices at work and diversity, with a specific focus on age. Katrina is interested in a broad range of methodological issues in organizational studies including digital and visual approaches. Katrina can be contacted at katrina.pritchard@open.ac.uk