

CHAPTER 2

USERS' VIEW OF ETHICS IN SOCIAL MEDIA RESEARCH: INFORMED CONSENT, ANONYMITY, AND HARM

Matthew L. Williams, Pete Burnap, Luke Sloan,
Curtis Jessop and Hayley Lepps

ABSTRACT

Some researchers consider most social media communications as public, and posts from networks such as Twitter are routinely harvested and published without anonymisation and without direct consent from users. In this chapter, we argue that researchers must move beyond the permissions granted by “legal” accounts of the use of these new forms of data (e.g., Terms and Conditions) to a more nuanced and reflexive ethical approach that puts user expectations, safety, and privacy rights center stage. Through two projects, we present qualitative and quantitative data that illustrate social media users’ views on the use of their data by researchers. Over four in five report expecting to be asked for their consent and 9 in 10 expect anonymity ahead of publication of their Twitter posts. Given the unique nature of this online public environment and what we know about users’

The Ethics of Online Research
Advances in Research Ethics and Integrity, Volume 2, 27–52
Copyright © 2017 by Emerald Publishing Limited
All rights of reproduction in any form reserved
ISSN: 2398-6018/doi:10.1108/S2398-601820180000002002

views pertaining to informed consent, anonymity, and harm, we conclude researchers seeking to embark on social media research should conduct a risk assessment to determine likely privacy infringement and potential user harm from publishing user content.

Keywords: social media; digital data; Twitter; safety; privacy law; privacy rights; privacy infringement

INTRODUCTION

Social media platforms offer social scientists previously unrealized access to real-time naturally occurring¹ data, and researchers are using sites such as Twitter, Flickr, Tumblr, and Facebook to collect open source online communications and publish content without anonymisation or informed consent. In this chapter, we argue research conducted using social media platforms must be subject to the same ethical scrutiny as offline research, and in some cases, an additional level of scrutiny may be needed. Indeed, “online research presents new ethical problems and recasts old ones in new guises” (Jones, 2011). Researchers have struggled to adapt existing frameworks to a mode of inquiry that takes place in a rapidly changing medium that is characterized by a blurring of the public and private, where data are created at scale outside of a “for research” context. This blurring has resulted in a schism in the perceptions held by researchers and the users of social media with respect to what is reasonable use of these open data for social science research. This chapter provides qualitative and quantitative insights into these user views and argues for a reflexive ethical approach to using social media data in social research.

CONTEXT TO THE RESEARCH

The New Social Media, New Social Science (NSMNSS) network² was set up in 2012 as a community of practice to help foster links between social media research practitioners and across disciplines; to catalyze discussions, address challenges, and to share best-practice, approaches, tools, and experiences. The ESRC-funded Social Data Science Lab³ was established in 2015 and continues the successful Cardiff Online Social Media Observatory (COSMOS) program of research that ran between 2011 and 2015. The Lab brings together social and computer scientists to study the methodological, theoretical, empirical,

and technical dimensions of “New and Emerging Forms of Data” in social and policy contexts. This empirical social data science program is complemented by a focus on ethics and the development of new methodological tools and technical/data solutions for the UK academic and public sectors. One of the key issues that came out of NSMNSS network discussions and Lab research was the ethics of social media research, and what was felt to be a lack of guidance in this area. A literature review from Salmons (2014) identified some initial ethics sources, and as social media research has developed as a field within the social science community, ethical guidelines too have been developed (e.g., British Psychological Society, British Educational Research Association, and British Sociological Association guidelines).

However, often what are missing from the conversation are the views of users. How do they curate their digital lives? What do they understand about how their information is used and shared on the Internet? What do users think about their information being used by researchers in online and social media research? Further, one of the particular characteristics of social media that presents challenges to social researchers is how the various platforms mediate and alter the relationship between the researcher and research participant. Within this context, in 2013, the National Centre for Social Research (NatCen) and, in 2015, the Social Data Science Lab conducted exploratory qualitative and quantitative research with social media users into how they feel about their online posts being used in research and their understanding of this type of research. The goal of these research projects was to reveal insights that researchers and practitioners could apply in their research design, recruitment, collecting or generating of data, and reporting of results.

Existing Literature on Ethics in Social Media Research

Legislation on Data Collection and Use in the UK

Data extracted from Social Media Application Programming Interfaces⁴ (APIs) contain personal data⁵ meaning they are subject to the UK Data Protection Act (DPA), and as such it must be processed fairly and lawfully. In cases where informed consent cannot be sought from users (likely to be the majority of cases if thousands of posts are being subject to analysis), a social researcher should establish the fair and lawful basis for collecting personal information. A researcher can accept that social media network terms of service provide adequate provision to cover this aspect of the DPA. However, if the data have been collected using a service that provides additional derived data on users, such as sensitive personal characteristics (e.g., ethnicity and sexual orientation)

based on algorithms that make estimations, the legal issue of privacy may be compounded as it is not information that the user has chosen to make “public.” Under the DPA, sensitive personal information can only be collected if the user has a legitimate need for processing. Deriving sensitive information and making conclusions about a person or persons’ views or philosophy and publicly releasing this information without anonymisation could lead to stigmatization or actual bodily harm (in this case of extremist views for example), should the location of the social media persona be established. Furthermore, it is possible that legal proceedings could follow. The DPA allows cases to be brought on a personal basis, so it is possible that the researcher and not the institution could be liable. Within the EU, the General Data Protection Regulation will replace the DPA in 2018. It includes provisions for the erasure of personal data and restrictions on data dissemination to third parties. However, it also imposes limitations on the right to be forgotten, including cases in which data are processed for historical, statistical, and scientific purposes. To what extent these proposals will impact upon social media research is unclear.

Learned Society Guidelines on Social Media and Internet-Based Research

Several learned societies have recognized the need for and introduced ethical principles for research in digital settings, including the British Psychological Society (BPS), the British Society of Criminology (BSC), the British Educational Research Association (BERA), the European Society for Opinion and Market Research (ESOMAR), and the Association and the Association of Internet Researchers (AoIR). Broadly, most guidelines adopt the “situational ethics” principle: that each research situation is unique and it is not possible simply to apply a standard template in order to guarantee ethical practice. Maybe the most thorough set of guidelines are those developed by AoIR. AoIR was one of the first learned societies to introduce a set of guidelines, which are now in their second iteration. These guidelines highlight three key areas of tension: the question of human subjects online; data/text and personhood; and the public/private divide (AoIR, 2012).

The guidelines advance the idea that the notion of the “human subject” is complicated when applied to research within online environments. For example, can we say “avatars” are human subjects? Does digital representation and automation of some online “behaviors” call into question the definition of human subjects in Internet-based research? If so, then it may be more appropriate and relevant to talk of harms, vulnerabilities, personal identifiable information, and so on. In addition, the Internet complicates the conventional construction of “personhood” and the “self,” questioning the

presence of the human subject in online interactions. Again, can we say an avatar is a person with a self? Is digital information an extension of a person? In some cases, this may be clear-cut: emails, instant message chat, newsgroup posts are easily attributable to the persons that produced them. However, when dealing with aggregate information in social media repositories, such as collective sentiment scores for subgroups of Twitter users, the connection between the object of research and the person who produced it is more indistinct. Attribute data on very large groups of anonymised Twitter users could be said to constitute non-personalized information, more removed from the human subjects that produced the interactions as compared to, say, an online interview. In these cases, the AoIR (2012, p. 7) guidelines state “it is possible to forget that there was ever a person somewhere in the process that could be directly or indirectly impacted by the research.”

In relation to informed consent, BERA specifically state that social networking and other on-line activities present challenges and the participants must be clearly informed that their participation, and their interactions are being monitored and analyzed for research. On anonymity the guidelines state one way to protect participants is through narrative and creative means, which might require the fictionalizing of aspects of the research or the creation of composite accounts, such as in vignettes, providing generalized features based on a number of specific accounts. In relation to consent, ESOMAR states that if it has not been obtained researchers must ensure that they report only depersonalized data from social media sources. If researchers are using automated data collection services, they are recommended to use filters and controls to remove personal identifiers such as user names, photos and links to the user's profile. In relation to anonymity the guidelines state where consent is not possible their analysis must only be conducted upon depersonalized data, and if researchers wish to quote publicly made comments, they must first check if the user's identity can be easily discoverable using online search services. If it can, they must make reasonable efforts to either seek permission from the user to quote them or mask the comment.

Academic Literature

There is some consensus about what ethical research involves, at least at the level of abstract principles (Webster, Lewis, & Brown, 2013), concerning obtaining informed consent, maintaining confidentiality and anonymity, and minimizing risk of harm to participants and researchers. Yet we do acknowledge some of the robust opposition to “principlism” in social science research ethics, a perspective “imported” from biomedical ethics models (see Chapter X in this volume).

Informed Consent

Gaining informed consent is a vital part of the early stages of research and must be negotiated and secured as early as is practically possible in the process. Ethical guidelines state that participants should understand the purpose of the research, what taking part will involve and how the data will be used. Participants require this information to make an informed decision about participation (ESRC, 2012; GSR, 2006; MRS, 2012). Questions have been raised about the process of consenting and how it needs to be done differently for different forms of research (Iphofen, 2011), or whether consent is required for all types of online research, or whether there are exceptions.

There are two schools of thought on informed consent. One position is that data posted in open spaces without password or membership restrictions would usually be considered to be in the public domain. This means they can be used for research purposes without the need for informed consent from individuals (see e.g., ESOMAR, 2009; Thelwall, 2010). The need to gain informed consent becomes relevant when data are obtained from closed websites requiring login details. This is challenged by the Market Research Society (MRS) discussion paper on online data collection and privacy (2012). Researchers should take into account the unique nature of online public environments. Internet interactions are shaped by ephemerality, anonymity, and a reduction in social cues, leading individuals to reveal more about themselves within online environments than would be done in offline settings, blurring the public and the private (Joinson, 1998; Lash, 2001; Williams, 2006). Research has highlighted the disinhibiting effect of computer-mediated communication, meaning Internet users, while acknowledging the environment as a (semi-) public space, often use it to engage in what would be considered private talk (Williams, 2006). Online information is often intended only for a specific networked public made up of peers, a support network, or specific community, not necessarily the Internet public at large, and certainly not for publics beyond the Internet (boyd, 2014). When it is viewed by unintended audiences, it has the potential to cause harm, as the information is flowing out of the context it was intended for (Barocas & Nissenbaum, 2014; Nissenbaum, 2008). Accepting this view, the AoIR (2012, p. 7) guidelines state that social, academic, and regulatory delineations of the public-private divide may not hold in online contexts and as such “privacy is a concept that must include a consideration of expectations and consensus” within context. In the final analysis, the subject of informed consent for social media research remains contentious among social scientists and views change depending on the topic, website, and sample population being worked with.

Regardless of the stance an individual takes on informed consent, obtaining it from individuals can in practice be very difficult. Social media research

does not typically offer the opportunity to verbally reiterate what participants are consenting to; researchers can be less confident that the key pieces of information have been relayed and understood. Further, researchers cannot verify a participant's identity to assess their capacity to consent. Where researchers cannot be certain the participant is of age, for example (Bull et al., 2010), it may be necessary to contact the guardians of children if possible (British Psychological Society, 2013).

Confidentiality and Anonymity

Conducting online and social media research presents three key challenges related to confidentiality and anonymity: safe data collection and storage may depend on platform security; that participants may want to be credited for their information and therefore not want to remain anonymous; and the possibility of breaking confidentiality when reporting findings.

Existing ethical guidelines state that researchers should ensure no one knows who has said what in a report (i.e., anonymity) and that participant information should be securely stored and shared (i.e., confidentiality; ESRC, 2012; GSR, 2006; MRS, 2012). However, in research using social media data, the risks of not upholding confidentiality are greater as a researcher has less control over data protection than in offline research (British Psychological Society, 2013). There is a permanent record of any information that is posted (Roberts, 2012), and direct quotations from participants can be traced back to the original source (BPS, 2007) through search engines like Google. In this case, anonymity cannot be protected. Anonymity is also related to the issue of copyright. For example, in the attempt of anonymising participant data researchers may exclude the participant's name; however, some users may feel that they should be given credit for their information being used (Barrett & Lenton, 2010; Lui, 2010; Roberts, 2012).

In cases where consent is not provided to directly quote without anonymisation, Markham (2012) suggests some innovative methods for protecting privacy in qualitative social media research. Acknowledging that traditional methods for protecting privacy by hiding or anonymising data no longer suffice in digital settings that are archived and searchable, Markham advocates bricolage-style reconfiguration of original data that represents the intended meaning of interactions. However, given Twitter rules on not changing the content of tweets in publication or broadcast, researchers are required to generate synthetic data that retains the meaning and sentiment of the original post. While this may be suitable for general thematic analysis, it may not satisfy the needs of more fine-grained approaches, such as conversation and discourse analysis.

Harm and Risk to Participants and Researchers

Researchers have an obligation to avoid causing physical, emotional, or psychological harm to participants. Research should also be conducted in a way to minimize undue harm to the research team. Appropriate support for participants needs to be in place following a research interaction, just as it would in face-to-face settings. Discussions of sensitive or emotional topics (e.g., posts about criminal activity, financial problems, mental health issues and feelings of suicide, extramarital sexual activity, controversial political opinions, and activism) have the potential to put participants at risk of emotional or psychological harm (Townsend & Wallace, 2016). Furthermore, the use of algorithms in social media research can expose users to harm. Taking the example of cyberhate on social media, Williams and Burnap (2016) employed machine-learning algorithms to classify hateful content and users (see also Burnap & Williams, 2015). They report that automated text classification algorithms perform well on social media datasets around specific events. However, their accuracy decreases beyond the events around which they were developed due to changes in language use. Therefore, an ethical challenge arises about how researchers should develop, use, and reuse algorithms that have the consequence of classifying content and users with sensitive labels often without their knowledge. Where text classification techniques are necessitated by the scale and speed of the data (e.g., classification can be performed as the data are collected in real-time), researchers should ensure the algorithm performs well (i.e., minimizing the number of false positives via continual testing) for the event under study in terms of established text classification standards. Furthermore, researchers have a responsibility to ensure the continuing effectiveness of the classification algorithm if there is an intention to use it beyond the event that led to its design.

High-profile failures of big data, such as the inability to predict the US housing bubble in 2008 and the spread of influenza across the United States using Google search terms, have resulted in many questioning the power and longevity of algorithms (Lazer, King, & Vespignani, 2014). Algorithms therefore need to be openly published and transparent for reproducibility (including classifier configuration and threshold settings), such that they can be routinely tested for effectiveness and may need to be “refreshed” with new human input and training data if false positives are to be minimized, avoiding the mislabeling of content and users. Where such information is published, every effort must be made to maintain anonymity, including efforts to reduce the likelihood of deductive disclosure (e.g., the linking of different social media for a single user can reveal their identity; Stewart & Williams, 2005).

Public Attitudes

The Eurobarometer Survey 359 Attitudes on Data Protection and Electronic Identity in the EU (2011; $N = 26,081$) found that 58 percent of European Internet users read online privacy policies.⁶ Over 70 percent were aware of the purposes for which social media networks can and may collect, use, and share personal data of users. Around 70 percent of European citizens were concerned about how companies use their data. More recently, Evans, Ginnis, & Bartlett (2015) conducted a survey of users' attitudes toward social media research in government and commercial settings. While 60 percent of respondents reported knowing that their social media data could be shared with third parties under the terms of service they sign up to, the same proportion felt that social media data should not be shared with third parties for research purposes. These views softened when users were offered anonymity and where only public data were to be used in the research. In the report, the majority of users rejected the position that accepting the terms of service was enough to establish consent, preferring instead opt-in consent for each individual research project. To date no research has attempted to model the predictors of the views of the public toward the use of their social media data in various settings. An overview of the methodologies of the two projects outlined in this chapter follows.

METHODOLOGY

Both the NatCen and Social Data Science Lab projects sought to address the issues raised in the preceding section by putting a series of open and closed questions to users on the use of their social media data for research.

NatCen Qualitative Research

The research used two qualitative methods for collecting data on users' perceptions of their use of social media data in research:

- Four focus groups
- Two paired and two depth interviews

The sample for the fieldwork was partly recruited from the British Social Attitudes 29 survey and partly from an external recruitment agency. It was purposively selected to ensure the views of low, medium, and high users of

social media were included,⁷ and to ensure diversity in relation to a number of characteristics including age, gender, ethnicity, and use of a variety of social media platforms for different purposes. Individuals who did not use the Internet were excluded from the study. In total 34 participants took part in an interview, paired interview, or focus group.

Focus groups and interviews were conducted by NatCen researchers using a topic guide covering themes such as general use of social media; views on research using social media; and key messages to researchers using social media in their research. The topics explored were acknowledged to be difficult to explain to participants who may not be familiar with social media or the terminology used. Many of the topics covered also required participants to think hypothetically so vignettes were used to illustrate key points and stimulate discussion.

Social Data Science Lab Survey

The Bristol Online Survey tool⁸ was used to design and distribute the ethics survey via social networks. The use of online media in social research is now well established and can yield reliable and valid results in a short period of time (Williams et al., 2017). Nonprobability sampling was employed to derive the sample of respondents. While sample bias is a fundamental shortcoming of nonprobability sampling, Meyer & Wilson (2009) note that this is often the only option available to researchers embarking on exploratory research. Furthermore, as the hypotheses tested in this analysis are concerned more with the existence of inter-variable relations and strengths of association than estimating population prevalence, the use of nonprobability sampling does not fundamentally weaken the design of the study (Dorofeev & Grant, 2006). Moreover, our study is principally concerned with “soft” measures (attitudes, perceptions, and opinions), which have no absolute validity (they cannot be compared with any authoritative external measure). However, Meyer and Wilson (2009) caution that sampling bias can still affect analysis if a sample is significantly uncharacteristic of the target population. The sample does not deviate significantly from what we know about the population of Twitter users. As our sample reflects, Twitter users are more likely to be younger, low-to middle-income earners, and are less likely to have children as compared to the general population (Sloan, 2017; Sloan & Morgan, 2015; Sloan, Morgan, Burnap, & Williams, 2015; Sloan et al., 2013;). However, given the size of the sample and the violation of the normality assumption for ordered linear regression analysis, the bias was corrected and accelerated bootstrapping

technique was utilized⁹ (Efron & Tibshirani, 1993). Given the nature of the research topic, the authors made efforts to establish informed consent via the introduction page to the online survey. The research aims and objectives were clearly expressed, and all the respondents were informed that the data produced would be anonymised and would remain confidential. Those under 18 were not permitted to complete the survey.

Qualitative Findings

Behaviors on, and Awareness and Understanding of Social Media

To better understand the context within which social media users' views on social media research were formed, the interviews and focus groups explored participants' behaviors on, and awareness and understanding of, social media.

Online behaviors varied widely between participants depending on the platform type and the intention of its use (e.g., social, leisure or business purposes, or a combination of them). Within these contexts, participants demonstrated three distinct (but overlapping) behaviors:

- “Creators” post original content on platforms including text, videos, and images.
- “Sharers” re-tweet, share, or forward content posted by others.
- “Observers” read and view content on social media and other sites but tend not to pass on this information.

The extent to which participants were aware of and understood social media varied, and depended on their sources of information. These included terms and conditions of the platforms, friends, and family, and online sources such as search engines. Participants reported not reading terms and conditions, accepting them only to progress to using the websites, due to the density of the content. They also conveyed difficulty in staying up-to-date with the terms and conditions of platforms as they are “constantly changing.” As a result, participants were not always aware of their privacy settings, and whether they were still sufficient for their needs after terms and conditions had changed.

Social media websites were described as “boundless” as they cross international boundaries and, therefore, may be regulated by a country's laws that differ from the laws of the user's country of residence. This linked in to a lack of confidence about the regulation of the use of content shared on social media and what social media sites can store on users from different countries.

The sources of information accessed by participants helped them to understand and be aware of issues inherent in social media, including its public nature (which raised issues of data ownership) and the difficulty of permanently deleting information. These two characteristics then related to three key concerns participants had about using social media:

- Maintaining privacy
- Protecting the reputation and identity of themselves, friends, and family
- Ensuring safety

Participants employed a number of strategies to mitigate the potential risks they identified. Participants discussed restricting the type of content shared (e.g., personal details or content that may “shed you in a bad light”) to address risks of undue intrusion, reputational concerns, and safety. Participants also mentioned adjusting privacy settings so that only their family and friends could access their content. Parents discussed monitoring what their children access and post on social media websites, for safety purposes and to protect them from possible reputational damage. However, participants’ views on what was “sensitive,” and the extent to which it should be shared varied by user. Also, not all participants were aware of how to adjust privacy settings, and there was a view that platforms change quickly and one cannot stay up to date with what their settings now mean.

Awareness and Views of Research Using Social Media

Participants’ awareness of research using social media varied. Some participants struggled to understand how social media could be used for research, or found it difficult to distinguish between social and market research, while others (typically more frequent users) had a better understanding of the term. This awareness of using social media for research, as well as their knowledge of social media more generally was closely related to their views on the subject. Participants’ feelings about research using social media fell into three categories: acceptance, skepticism, and ambiguity.

Participants who accepted research using social media discussed the value of this methodological approach and the benefits it may have to society. These participants recognized that social media data collection methods could be beneficial when trying to understand broad social trends as the volume of data could mitigate extreme views, and that people may be more open or honest online than in, for example, a face-to-face survey. Accepting views were also expressed by those users who “self-regulated” online. These participants only

posted online what they were happy for others to access and, therefore, accepted that researchers may use their information and were comfortable with this.

Skepticism about social media research was expressed by participants and found to be related to uncertainty about the validity or value of data from social media, compared to traditional research methods. Participants also had concerns about the lack of transparency online, particularly in relation to who was conducting the research and its purpose. Participants were concerned that their data may be taken out of context to support something they did not agree with and that they were not able to confirm the legitimacy of “who they’re dealing with.”

Ambivalent participants, who expressed neither concern nor acceptance of social media research, felt that whatever their view it would not be listened to and that there was little they could do to stop it from happening. Participants worried about “Big Brother” culture and saw the use of social media data as inevitable; it was accepted that having your information used was part what happens when you put it online.

Informed Consent

Participants identified four key factors that influenced the importance of researchers gaining consent:

- Mode and content of the posts;
- Social media website being used;
- The expectations the user had when posting, and;
- The nature of the research

Participants who “self-regulated” did not think researchers needed to gain consent and this held true whatever the type and content of the post. However, others felt it was important. Some users suggested that researchers should ask to use any written content, in particular if it were to be published alongside the username. A different view was that the researcher did not need to ask for consent as long as they were sure the Tweet was an accurate representation of the users’ views. Images were identified as particularly problematic, and it was questioned what rights individuals had when they were included in photos posted by other social media users. Finally, participants suggested that researchers should go to greater lengths to gain consent and/or protect anonymity when using posts with more “sensitive” content.

The type of social media website was another factor in the qualitative study that influenced whether or not participants thought consent definitely needed

to be gained by a researcher. Social media websites with a “social” purpose were viewed differently from websites with a professional aim. Websites with a “social” purpose were thought to contain more “personal” content, whereas content posted to “professional” sites was less so. In light of this, participants thought that it would be acceptable for researchers to access the latter without gaining consent because the risks associated with being identifiable through personal information are lower.

Participants’ views were also influenced by user expectations. If a user intended for their post to be widely accessible, then it was felt that a researcher would not necessarily need to gain consent to use it, though this assumed users understood the openness and accessibility of the platform, which may not be the case. In contrast, if a user had not meant it to be public in the first place, or it was posted for a different purpose, then it should not be used. This was felt to be more important than the site from which a researcher took it. For researchers, this means that no matter how open or public a site is considered to be, the user’s expectation about how the post should be used is what should be considered.

Finally, the nature of the research and the nature of the organization also affected participants’ views on research ethics. Use of social media data was affected by the affiliation of the researcher and the purpose of the research. The type of organization that the research was affiliated with (e.g., charitable or commercial) influenced whether or not participants viewed research to be of “good quality.” Research being conducted by a not-for-profit organization (typically a university), rather than for “commercial” reasons, was preferred because the former were felt to be more “productive,” more “ethical” and “not exploitative.” Further, participants did not like to think of their social media posts being used to generate a profit for others although it was acknowledged that this was already happening. As elsewhere, participants of the opinion that once one posts to a social media platform you waive your right to ownership were not concerned about the affiliation of the researcher. Other users were unaware of the differences between not-for-profit and commercial researchers or did not care about the distinction, and therefore had little to say about how researcher affiliation might influence their desire to agree to informed consent. Although concern about the affiliation of the researcher was not widespread, concern about the “purpose” of the research was. Participants expressed worry about their posts being used to “drive an agenda” they would not have agreed to if the researcher had asked them. Using social media content to “drive an agenda” was seen differently from research offering a social benefit such as research aimed at providing more knowledge about a particular social issue or improving public services. In these instances, the potential benefit to society was seen to outweigh the risks to the user.

Quantitative Findings

The Social Data Science Lab ethics survey ($N = 564$) revealed that 94 percent of respondents were aware that Twitter had terms of service, and just below two-thirds had read them in whole or in part. Seventy-six percent knew that when accepting terms of service, they were providing consent for some of their information to be accessed by third parties (see Table 1). Least concern was expressed in relation to Twitter posts being used for research in university settings (84 percent of respondents were not at all or only slightly concerned, compared to 16 percent who were quite or very concerned). Concern in relation to Twitter being used for research rose significantly in government (49 percent were quite or very concerned) and commercial settings (51 percent were quite or very concerned). Respondents expressed high levels of agreement in relation to the statements on consent and anonymity in Twitter research. Just under 80 percent of respondents agreed that they would expect to be asked for their consent before their Twitter posts were published in academic outputs. Over 90 percent of respondents agreed that they would want to remain anonymous in publications stemming from Twitter research based in university settings.

Four factors emerged as statistically significant in relation to concern in university settings (see Table 2). Unsurprisingly, those with no knowledge of Twitter's terms of service consent clause were more likely (odds increase of 1.59) to express concern in this setting. Those who use the Internet for more hours in the day were also more likely (odds increase of 1.10) to express concern, but the effect was marginal. Of the demographic variables, parents of children under 16 were more likely (odds increase of 2.33) to be concerned compared to nonparents. Female respondents were more likely (odds increase of 1.92) to be concerned compared to male respondents. Several predictors emerged as significant for concern in government and commercial settings that were not significant for concern in university settings. Lesbian, gay, and bisexual (LGB) respondents were more likely to express concern over their Twitter posts being used in government (odds increase of 2.12) and commercial settings (odds increase of 1.92), compared to heterosexual respondents. Older respondents were also more likely to report higher degrees of concern in both these settings, as were those who had a higher level of Internet expertise. Those who posted information most often on Twitter were less likely to be concerned with their information being used in commercial settings.

Those respondents who reported familiarity with Twitter's terms of service consent clause were significantly less likely to expect to be asked for their informed consent by university researchers to publish content (odds decrease

Table 1. Sample Descriptives.

	Coding	%/M ^a	N ^b /SD
Dependent Variables			
Concern – University Research	Not at all concerned	37.2	136
	Slightly concerned	46.4	170
	Quite concerned	11.2	41
	Very concerned	5.2	19
Concern – Government Research	Not at all concerned	23.3	85
	Slightly concerned	27.7	101
	Quite concerned	25.5	93
	Very concerned	23.6	86
Concern – Commercial Research	Not at all concerned	16.8	61
	Slightly concerned	32.1	117
	Quite concerned	29.4	107
	Very concerned	21.7	79
Expect to be asked for Consent	Disagree	7.2	26
	Tend to disagree	13.1	47
	Tend to Agree	24.7	89
	Agree	55.0	198
Expect to be anonymised	Disagree	5.1	18
	Tend to disagree	4.8	17
	Tend to Agree	13.7	48
	Agree	76.4	268
Independent variables			
Frequency of posts daily	Scale (range: 1 “Less than once” to 7 ‘over 10’)	1.75	1.23
Postpersonal activity	Yes=1	37.7	161
Postpersonal photos	Yes=1	19.0	81
Knowledge of ToS consent	Yes=1	75.5	317
Net use (years)	Scale (range: 1 “Less than year” to 9 ‘15+ years’)	6.59	1.76
Net use (hours per day)	Scale (range: 1 “Less than hour” to 10 ‘10+ hours’)	6.03	2.52
Net skill	Scale (range: 1 “Novice” to 10 ‘Expert’)	7.69	1.60
Sex	Male = 1	48.93	276
Age	Scale (range: 18–83)	25.38	10.17
Sexual orientation	Heterosexual = 1	83.6	357
Ethnicity	White = 1	91.1	389
Relationship status	Partnered = 1	45.4	194
Income	Scale (range: 1 “below 10K” to 11 ‘100K+’)	3.72	3.07
Has child under 16	Yes = 1	7.3	31

Notes:^aMean and standard deviation given for scale variables.

^bReduction in sample size due to missing data; bootstrapped (BCa) results reported.

Table 2. Ordered Regression Predicting Concern about Using Social Media Data in Three.

	University			Government			Commercial					
	B	SE ^a	Wald Exp(B)	B	SE ^a	Wald Exp(B)	B	SE ^a	Wald Exp(B)			
Frequency of posts	0.048	0.093	0.266	1.05	-0.067	0.089	0.564	0.089	-0.227*	0.09	6.304	0.80
Postpersonal activity	0.216	0.246	0.771	1.24	0.101	0.235	0.184	0.235	0.470	0.237	3.926	1.60
Postpersonal photos	0.132	0.27	0.237	1.14	0.119	0.259	0.209	0.259	0.133	0.261	0.259	1.14
Knowledge of ToS consent	-0.465*	0.244	3.626	0.63	-0.246	0.234	1.104	0.234	-0.289	0.235	1.505	0.75
Net use (years)	-0.041	0.071	0.335	0.96	-0.059	0.068	0.746	0.068	0.059	0.068	0.731	1.06
Net use (hours per day)	0.092*	0.043	4.478	1.10	0.089*	0.042	4.598	0.042	0.062	0.042	2.255	1.06
Net skill	0.078	0.071	1.202	1.08	0.149*	0.068	4.783	0.068	0.135*	0.068	3.91	1.14
Sex	-0.659**	0.246	7.193	0.52	-0.291	0.232	1.569	0.232	-0.353	0.233	2.304	0.70
Age	0.003	0.014	0.038	1.00	0.072**	0.015	21.995	0.015	0.066**	0.015	19.497	1.07
Sexual orientation	-0.27	0.29	0.866	0.76	-0.752**	0.281	7.143	0.281	-0.653*	0.282	5.373	0.52
Ethnicity	0.055	0.383	0.021	1.06	-0.311	0.377	0.679	0.377	-0.422	0.378	1.247	0.66
Relationship status	-0.414	0.216	3.677	0.66	-0.206	0.206	1.006	0.206	-0.363	0.207	3.072	0.70
Income	-0.045	0.034	1.774	0.96	-0.045	0.032	1.921	0.032	-0.053	0.033	2.66	0.95
Has child under 16	0.846*	0.416	4.137	2.33	0.27	0.408	0.437	0.408	0.498	0.409	1.481	1.65
Model Fit												
-2 Log Likelihood	790.730			944.497					920.106			
Model Chi-square	31.182			65.712					66.789			
df	15			15					15			
sig.	0.00			0.00					0.00			
Cox and Snell Pseudo R ²	0.08			0.17					0.17			
Nagelkerke Pseudo R ²	0.09			0.18					0.18			

Note.^aBootstrapped (BCa) standard errors reported; Reduction in sample size due to listwise deletion of cases necessary for regression.

of 0.62; see Table 3). Early adopters of the Internet were likely to hold the same view, but to a lesser degree. Female Twitter users and those who post personal photos were more likely to expect anonymity in publishing (odds increase of 1.47 and 1.61, respectively). By far, the most striking result was that of BME tweeters who were much more likely (odds increase of 3.90) to want anonymity compared to white tweeters. These findings lend support to the position that for Twitter-based research to be conducted in an ethical manner it is possible to rely on terms of service to harvest data, but personal information (e.g., extreme opinion, photo, demographic information, location) should not be directly quoted in publication without some form of informed consent.

Table 3. Ordered Regression Predicting Expectation of Request for Informed Consent and Anonymity in Social Media Research in University Settings.

	Informed Consent				Anonymity			
	B	SE ^a	Wald	Exp(B)	B	SE ^a	Wald	Exp(B)
Frequency of posts	-0.05	0.095	0.275	0.95	-0.097	0.11	0.771	0.91
Post personal activity	0.034	0.253	0.018	1.03	0.311	0.314	0.979	1.36
Post personal photos	-0.272	0.277	0.961	0.76	0.471*	0.33	2.037	1.61
Knowledge of ToS consent	-0.478*	0.262	3.315	0.62	0.115	0.318	0.131	1.12
Net use (years)	-0.155*	0.074	4.388	0.86	-0.105	0.091	1.321	0.9
Net use (hours per day)	0.055	0.045	1.480	1.06	0.049	0.056	0.758	1.05
Net skill	-0.063	0.075	0.710	0.94	-0.109	0.093	1.363	0.9
Sex	-0.241	0.244	0.974	0.79	-0.385*	0.299	1.656	0.68
Age	-0.020	0.014	2.004	0.98	0.017	0.02	0.735	1.02
Sexual orientation	-0.167	0.298	0.316	0.85	0.004	0.356	0.001	1.00
Ethnicity	0.160	0.394	0.165	1.17	-1.369*	0.318	10.13	3.90
Relationship status	-0.019	0.222	0.008	0.98	-0.129	0.275	0.222	0.88
Income	-0.021	0.034	0.380	0.98	-0.004	0.044	0.009	1.00
Has child under 16	0.243	0.431	0.318	1.27	0.052	0.298	0.030	1.05
Model fit								
-2 log likelihood	788.767				526.805			
Model chi-square	24.762				18.68			
Df	15				15			
sig.	0.00				0.00			
Cox and Snell pseudo R ²	0.09				0.09			
Nagelkerke pseudo R ²	0.09				0.09			

DISCUSSION

The NatCen qualitative study and the Social Data Science Lab survey provided a first look at users' concern over being included in social media research in various settings and their expectations regarding consent and anonymity. While both the qualitative and quantitative data showed a general lack of concern from social media users over their information being used for research purposes (with university research attracting least concern), the majority of respondents stated that they would want to be asked for consent and to remain anonymous in publications reporting social media research. These patterns reflect those found in the Eurobarometer Survey (2011) that showed three-quarters of Europeans accepted that disclosing personal information was now a part of modern life, but only a quarter of respondents felt that they had complete control over their social media information and 70 percent were concerned that their personal data may be used for a purpose other than for which they were archived. A clear majority of Europeans (75 percent) want be able to delete personal information on a website whenever they decide to do so, supporting the "right to be forgotten" principle. Taken together, these findings show that there may be a disjuncture between the current practices of researchers in universities, government departments, and commercial organizations in relation to the harvesting and representation of social media data, and users' views of the appropriate use of their online information and their rights as research subjects.

Informed Consent

The Social Data Science Lab survey showed that nearly 80 percent of respondents agreed to some extent that they should be contacted for their informed consent before their posts are published in academic outputs. Participants in the qualitative study expressed a range of views about the extent to which researchers should seek informed consent when observing or collecting data from social media platforms. Some participants did not think consent needed to be gained because "there is no such thing as privacy online," and by posting content you automatically consent to its wider use – if you did not wish your data to be used you should adjust your privacy settings. However, others believed consent should always be sought due to common courtesy and to protect the "intellectual property" rights of users, to the extent that some felt that using content without permission may be illegal. It was recognized that gaining consent may be impractical, but for those that felt consent should be gained this was not viewed as a justification not to ask permission.

Anonymity

The survey showed that over 90 percent of respondents agreed to some extent that they would want to remain anonymous in publications stemming from social media research (in particular female and BME tweeters and those posting personal photographs). Anonymity for participants in the qualitative study meant not having their name, or username, used in any research outputs alongside any content they posted online. It was felt anonymity should be upheld for two reasons: to avoid judgment from others and to prevent reputational risk. Of course, this goal of anonymity does not balance well with ideas of intellectual property and the importance of proper referencing, also discussed by participants. For participants who disagreed with the need for anonymity, the reason was similar to those who did not think informed consent was needed, that is, it is up to the user to manage their identity when online. There was also a view that some responsibility should fall on platform owners to educate users about the potential risks of sharing content online. Regardless of these various viewpoints, anonymity is not possible in the case of Twitter given their Developer Agreement (see Appendix). Furthermore, even if the user name was removed, this is not enough to preserve anonymity as tweet text is searchable.

Avoiding Undue Harm

Participants in the qualitative study were wary about how they could be sure of what researchers were saying, and how difficult it would be online to decide if they were “legitimate” researchers. Closely related to anonymity, participants felt that being identifiable in research could lead to unsolicited attention online and, more seriously, “abuse.” This might be from people they knew, or from organizations that could “exploit” them. For others it meant use by the police or courts, for purposes of prosecution. These concerns relate to the abundance of sensitive information about users available on social media and generated by algorithms based on account meta data and communication patterns. The survey found associations between sexual orientation, ethnicity, and gender and feelings of concern and expectations of anonymity. Such characteristics may be considered sensitive under the UK Data Protection Act. A principle ethical consideration in most learned society guidelines on digital social research is to ensure the maximum benefit from findings while minimizing the risk of actual or potential harm during

data collection, analysis, and publication (interpreted as physical or psychological harm, including discomfort, stress, and reputational risk). Potential for harm in social media research increases when sensitive data are estimated and harvested. These data can include personal demographic information (such as ethnicity and sexual orientation), information on associations (such as memberships to particular groups or links to other individuals known to belong to such groups), and communications of an overly personal or harmful nature (such as details on morally ambiguous or illegal activity and expressions of extreme opinion). In some cases, such information is knowingly placed online, whether or not the user is fully aware of who has access to this information and how it might be repurposed (Dicks, Mason, Williams, & Coffey, 2006). In other cases sensitive information is not knowingly created by users, but it can often come to light in analysis where associations are identified between users (not everything can be known about another user before connecting nor can changes in affiliation be monitored on a routine basis) and personal characteristics are estimated by algorithms (van Dijck, 2013; Sloan et al., 2015).

CONCLUSION AND RECOMMENDATIONS

This chapter outlines users' views on the use of their social media data in social research. Over four in five reported expecting to be asked for their consent and nine in ten reported expecting anonymity ahead of publication of their qualitative social media posts. However, some researchers consider most social media communications as public, and data from networks such as Twitter are routinely harvested and published without anonymisation and without direct consent from users. Contrary to this practice, we argue that while it may be reasonable for researchers to rely on the T&Cs of social media companies as an indication of informed consent with respect to data collection, a more reflexive approach to research ethics is needed in the later stages of the research process. Researchers should consider users' views and expectations, in addition to legal data protection requirements, when setting out to analyze (including the use of algorithms), publish, and store social media data.

With respect to publishing social media data, given the unique nature of this online public environment and what we know about users' views pertaining to informed consent, anonymity, and harm, we recommend that researchers seeking to embark on social media research should conduct the following

risk assessment process to determine likely privacy infringement and potential user harm from publishing nonanonymised data (a necessity given Twitter's Developer Agreement and the ability to search content):

- Are social media posts from individual private users, with no public profile (i.e., not celebrities or public figures¹⁰)? If so, in almost every case researchers should seek opt-out consent¹¹ to publish content without anonymisation.
- Are social media posts from organizational accounts (such as government departments, businesses, charities) where users are tweeting on behalf of the organization (and not personally) or public figures/celebrities? If so, consider as truly "public data," and in almost every case publish without informed consent.
- Are posts identified as sensitive (e.g., posts about criminal activity, financial problems, mental health issues and feelings of suicide, extramarital sexual activity, controversial political opinions and activism; Townsend & Wallace, 2016)? If so, researchers should seek opt-in consent.
- Are private users identifiable as vulnerable (e.g., children, learning disabled, and those suffering from an illness)? If so, opt-in consent should be sought from the user and/or guardian or proxy.
- Have tweets been deleted (either individually or via a deleted account) at the time of writing? If so, consider content removed from the public domain, and do not publish (unless they are from an organizational or public figure account – see above).
- Are tweets from identifiable bots? If so, consider user as non-human subject and publish without consent.

Codes of ethical conduct that were first written over half a century ago are being relied upon to guide the representation of digital data. This risk assessment process updates these frameworks allowing researchers to move beyond the permissions granted by "legal" accounts of the use of these new forms of data (e.g., T&Cs) to a more nuanced and reflexive ethical approach that puts user expectations, safety, and privacy rights central stage.

NOTES

1. Although we acknowledge communications are mediated by technology in ways that are sometimes obfuscated to the user and analyst.
2. Nsmnss.blogspot.com
3. socialdatalab.net

4. APIs are Internet interfaces that facilitate programmatic access to online data feeds.

5. The DPA states personal data means data which relate to a living individual who can be identified – (a) from those data, or (b) from those data and other information which is in the possession of, or is likely to come into the possession of, the data controller, and includes any expression of opinion about the individual and any indication of the intentions of the data controller or any other person in respect of the individual.

6. We assume read in whole or in part.

7. Low users were defined as people who did not use social media websites, or used them once a week or less, medium users as those who used websites from twice a week up to once a day, and high users as those who used social media websites several times a day.

8. See <http://www.survey.bris.ac.uk/>

9. A nonparametric resampling procedure used to empirically estimate the sampling distribution of the indirect effect, thus reducing problems with type I errors and low statistical power endemic to analyses that rely on assumptions of sampling distribution normality.

10. Researchers should consult existing ethical guidelines that provide definitions of public figures (e.g., politicians and celebrities who aim to communicate to a wide audience).

11. In the case of opt-out consent, researchers may wish to set a reasonable time window for a reply (e.g., 2 weeks to 1 month), and repeat consent requests several times should a timely response not be forthcoming. If the tweeter is no longer active (i.e., has not tweeted in last 3-6 months), consider the account as inactive and do not publish (as we can reasonably assume the tweeter has not seen the request and therefore cannot take up the option of opting-out).

REFERENCES

- AoIR (Association of Internet Researchers). (2012). Ethical decision-making and internet research. Retrieved from <http://www.aoir.org/reports/ethics2.pdf>. Accessed on 15 January 2016.
- Barocas, S., & Nissenbaum, H. (2014). Big data's end run around procedural privacy protections. *Communications of the ACM*, 57(11), 3–33.
- Beninger, K., Fry, A., Jago, N., Lepps, H., Nass, L., & Silverster, H. (2014). *Research using social media. Users Views*. London: NatCen.
- boyd, d (2014). *It's complicated: Social lives of networked Teens*. New Haven, CT: Yale University Press
- Burnap, P., Gibson, R., Sloan, L., Southern, R., & Williams, M. L. (2016). 140 characters to victory?: Using Twitter to predict the UK 2015 General Election. *Electoral Studies*, 41(2), 230–233.
- Burnap, P., & Williams, M. L. (2015). Cyber hate speech on Twitter: An application of machine classification and statistical modeling for policy and decision making. *Policy & Internet*, 7(2), 223–242.
- Burnap, P., & Williams, M. L. (2016). Us and them: Identifying cyber hate on Twitter across multiple protected characteristics. *EPJ Data Science*, 5, article number: 11. doi:10.1140/epjds/s13688-016-0072-6

- Burnap, P., Williams, M. L., Rana, O., Edwards, A., Avis, N., Morgan, J. et al. (2013) Detecting tension in online communities with computational Twitter analysis. *Technological Forecasting & Social Change*.
- Burnap, P., Williams, M. L., Rana, O., Edwards, A., Avis, N., Morgan, J., et al. (2014). COSMOS: Towards an integrated and scalable service for analysing social media on demand. *International Journal of Parallel, Emergent and Distributed Systems*.
- Burnap, P., Williams, M. L., & Sloan, L. (2014) Tweeting the terror: modelling the social media reaction to the Woolwich terrorist attack. *Social Network Analysis and Mining*, 4, 206.
- Dicks, B., Mason, B., Williams, M., & Coffey, A. (2006). Ethnography and data reuse: Issues of context and hypertext. *Methodological Innovations*, 1(2), 33–46.
- Dorofeev, S. & Grant, P. (2006). Statistics for real-life sample surveys: Non-simple random samples and weighted data. Cambridge: Cambridge University Press.
- Edwards, A., Housley, W., Williams, M. L., & Sloan, L. (2013). Digital social research, social media and the sociological imagination: Surrogacy, augmentation and re-orientation. *International Journal of Social Research Methodology*, 16, 245–260.
- Eurobarometer Survey 359. (2011). *Attitudes on data protection and electronic identity in the European Union*. Brussels.
- Evans, H., Ginnis, S., & Bartlett, J. (2015). SocialEthics: A guide to embedding ethics in social media research. London: Ipsos MORI.
- Ginsberg, J., Mohebbi, M. H., Patel, R. S., Brammer, L., Smolinski, M. S., & Brilliant, L. (2009). Detecting influenza epidemics using search engine query data. *Nature*, 457 (7232).
- International Visual Sociology Association (2009). IVSA code of research ethics and guidelines. *Visual Studies*, 24(3), 250–257.
- Iphofen, R. (2011). Ethical decision making in qualitative research. *Qualitative Research*, 11, 443–446.
- Joinson, A. N. (1998). Causes and effects of disinhibition on the Internet. In J. Gackenbach (Ed.), *The psychology of the Internet* (pp. 43–60). New York, NY: Academic Press.
- Lazer, D., King, G., & Vespignani, A. (2014). The parable of Google Flu: Traps in big data analysis. *Science*, 343(6176), 1203–1205
- Markham, A. (2012). Fabrication as ethical practice: Qualitative inquiry in ambiguous Internet contexts. *Information, Communication and Society*, 15(3), 334–353.
- Metcalf, J. & Crawford, K. (2016). Where are human subjects in big data research? The emerging ethics divide. *Big Data and Society*, 3(1), 1–14.
- NatCen. (2014). *Research using social media: Users' views*. London: NatCen.
- Nissenbaum, H. (2009). Privacy in context: Technology, policy, and the integrity of social life. Stanford, CA: Stanford University Press.
- Salmons, J. (2014). New social media, new social science... and new ethical issues! Retrieved from <http://nsmnss.blogspot.co.uk/2014/02/new-social-media-new-social-science-and.html>
- Sloan, L. (2017). Who tweets in the United Kingdom? Profiling the Twitter population using the British social attitudes survey 2015. *Social Media and Society*, 3(1).
- Sloan, L. & Morgan, J. (2015). Who Tweets with their location? Understanding the relationship between demographic characteristics and the use of geoservices and geotagging on Twitter. *PLoS ONE*, 10(11).
- Sloan, L., Morgan, J., Burnap, P., & Williams, M. L. (2015). Who Tweets? Deriving the demographic characteristics of age, occupation and social class from Twitter User Meta-Data. *PLOS ONE*.

- Sloan, L., Morgan, J., Williams, M. L., Housley, W., Edwards, A., Burnap, P. et al. (2013). Knowing the Tweeters: Deriving sociologically relevant demographics from Twitter', *Sociological Research Online*, 18(7).
- Stewart, K. F., & Williams, M. L. (2005). Researching online populations: The use of online focus groups for social research. *Qualitative Research*, 5(4), 395–416.
- Townsend, L., & Wallace, C. (2016). *Social media research: A guide to ethics*. Scotland: University of Aberdeen.
- Twitter. (2015). Developer agreement, San Francisco: Twitter. Retrieved from <https://dev.twitter.com/overview/terms/agreement-and-policy>
- Twitter. (2016). Broadcast guidelines, San Francisco: Twitter. Retrieved from <https://about.twitter.com/en-gb/company/broadcast>
- van Dijck, J. (2013). *The Culture of Connectivity*. Oxford: Oxford University Press.
- Williams, M. L. (2015). Towards an ethical framework for using social media data in social research. Paper presented at Social Research Association Workshop, Institute of Education, UCL, 15 June 2015.
- Williams, M. L., & Burnap, P. (2016). Cyberhate on social media in the aftermath of Woolwich: A case study in computational criminology and big data. *British Journal of Criminology*, 56(2), 211–238.
- Williams, M. L., & Burnap, P. (2017). Crime sensing with big data: The affordances and limitations of using open source communications to estimate crime patterns. *British Journal of Criminology*. Online Advance Access.
- Williams, M. L., Edwards, A., Housley, W., Burnap, P., Rana, O., Avis, N., ..., Sloan L. (2013). Policing cyber-neighbourhoods: Tension monitoring and social media networks. *Policing and Society*, 23, 461–481.

APPENDIX

Twitter Terms of Service

The Twitter Developer Agreement (Twitter 2015) terms of service states that all users of its API must:

Maintain the integrity of Twitter's products by ensuring:

- i) Usernames are always displayed (and name if possible) with tweet text;
- ii) Users of the API respond to content changes such as deletions or public/private status of tweets; and
- iii) Content of tweets is not modified, translated or deleted (in part or in whole).

Respect Users' Privacy and get the user's express consent before they do any of the following:

- i) Take any actions on a user's behalf, including posting content and modifying profile information;
- ii) Store nonpublic content such as direct messages or other private or confidential information;
- iii) Share or publish protected content, private or confidential information.

In the case of the reproduction of tweets (exhibition, distribution, transmission, reproduction, public performance, or public display of Tweets by any and all means of media) the Twitter guidelines state broadcasters & publishers should:

- i) Include the user's name and Twitter handle (@username) with each Tweet;
- ii) Use the full text of the Tweet. Editing Tweet text is only permitted for technical or medium limitations (e.g., removing hyperlinks);
- iii) Not delete, obscure, or alter the identification of the user. Tweets can be shown in anonymous form in exceptional cases such as concerns over user privacy;
- iv) In some cases, seek permission from the content creator, as Twitter users retain rights to the content they post.