



## Report on workshop: Ethical issues in social media research: current thinking

15th May, 1-4pm, King’s College London

Funded by a [Wellcome Trust award](#) exploring ethical decision-making and governance in the field of social media research. For more information, and for a list of our publications, please visit <https://wp.lancs.ac.uk/social-media-research-ethics/>. In collaboration with Wasim Ahmed (#NSMNSS) and Curtis Jessop (NatCen Social Research)

### Contents

Contents .....	1
Introduction .....	2
Speakers.....	3
Emerging themes.....	4
1. Moving from human participant research to data subjects .....	4
2. Ethics is not discrete, it is a process.....	4
3. The perspective of social media users.....	6
Specific discussions and the negotiation of blurred boundaries .....	7
1. GDPR.....	7
2. Ethical approval .....	8
3. Consent.....	8
3. Data privacy, anonymisation and data security .....	9
4. Academic versus industry social media research .....	10

## Introduction

There has been a massive rise in the number of academic researchers using information from social media sites, such as Facebook, Twitter, blogs and online forums, as sources of data for research ('social media research'). Researchers have been accessing, processing and analysing social media data using a combination of both traditional and innovative methodological approaches, including data modeling, large quantitative ('big data') studies and smaller in-depth qualitative analyses.

Social media research, whilst exciting and offering great potential for public benefit, can be ethically complex. This complexity stems from the way in which social media research blurs the boundaries around traditional ethical requirements such as consent and data anonymisation. Key ethical questions well-versed in the social media research ethics literature include:

- Should social media data be considered as belonging to human participants research or should it be considered as published data? How does this influence practices of receiving consent from social media users? Extrapolating from this, when consent is preferred, how should academic researchers act when this may not be practically feasible?
- Do social media users have perceived expectations of privacy when using social media platforms (and in which contexts)? If so, how can we maintain this privacy, thereby protecting these expectations when conducting academic research? Is anonymising data the answer?

Much progress has been made in trying to address these and similar questions, and the issues emerging from them. There are a plethora of discussions within the literature and various publications citing case studies of ethical best practice; professional bodies have published a range of guidelines to aid academic researchers as they negotiate this ethical complexity; and a number of international forums have been developed to share best practice amongst researchers.

During these discussions, there has come a realisation that many of the answers to the above questions are context specific, and require deep thought on behalf of each scholar embarking on their work about how best to approach their research ethically. There are no black and white answers, and no blanket ethical approaches. Whilst this may be the case, social media research ethics has come a long way over the past few years, and there has been immense progression in

terms of being able to give advice, drawing on a range of expertise and evidence in terms of how best to approach research practice.

It was this advice, expertise and evidence which we wanted to 'tap into' for this current workshop. The aim of the workshop was to bring together a range of expert speakers to discuss current thinking for social media research ethics. It was an effort to pass on expertise to those embarking on social media research, those reviewing such research (as in the case of research ethics committees) and those experienced in this area but who wanted to improve their understanding of current ethical thinking. Though the workshop was also just as much about acting as a research project itself - to decipher how the field of social media research ethics is progressing in terms of developing ethical best practice. The workshop of over 40 participants was well attended by researchers and research ethics committee members alike. The interactivity and engagement of the audience, and the diversity and range of questions asked, provides evidence that bringing this community together in these forums is essential to promote a shared understanding of ethical best practice for social media research.

## Speakers

The workshop programme brought together a diverse range of speakers from the field:

- **Helen Kara:** Ethics Lead for the Social Research Association
- **Gabrielle Samuel:** Research Fellow, Lancaster University/King's College London
- **Itzelle Medina Perea:** PhD student at the Information School at the University of Sheffield
- **Sarah Quinton:** Senior lecturer at Oxford Brookes University
- **Curtis Jessop:** Research Director at NatCen Social Research
- **Steven Ginnis:** Research Director, Deputy Head of Public Communication Research
- **Wasim Ahmed:** Lecturer, Northumbria University
- **Libby Bishop:** Coordinator for International Data Infrastructures at GESIS, Leibniz Institute for the Social Sciences
- **Sally Sanger:** PhD student at the Information School at the University of Sheffield
- **Brittany Kelley:** Teaching Fellow at King's College London

## Emerging themes

Three predominant themes emerged during the afternoon.

### 1. Moving from human participant research to data subjects

Much focus of social media research ethics discussion has been associated with whether social media data should be classified as human participant research. This is because journal editors and research ethics committee members often ask this question to determine whether ethical approval/consent is required for a specific research project.

In contrast to the literature, during the workshop, speakers rarely touched upon the issue of human participants, rather, moving beyond it. For the speakers, there seemed to be an implicit assumption that all researchers – no matter what data was being analysed - should be concerned about social media users and think carefully about how to act ethically when collecting and analysing social media data (see theme 3 below). One specific speaker, Brittany, explained how, in spite of the fact she came from a discipline which would ordinarily describe the written word (including social media data) as ‘text’, and therefore not subject to the ethical norms of human participant research, her journey though analysing fan fiction online forums allowed her to reflect on the nature of social media data, and the need to think carefully about respecting social media users when analysing that data. Through her journey Brittany learned the importance of receiving consent from social media users to use the data they had posted online.

As such, there was a realisation that the ethical boundary between human participant versus other research may no longer be useful in the field of social media research because instances arose in which social media research may be classified as ‘text’ but consent was still viewed as ethical best practice. I use this opportunity to make a recommendation for research ethics committees and journal editors to incorporate the term ‘data subjects’ to encapsulate this research field. Some committees have started to move towards this approach already.

### 2. Ethics is not discrete, it is a process

A major over-arching theme of the workshop was to view ethics as much broader than through the lens of questions around consent and anonymisation (as is often the case). This did not mean that such questions went unmentioned during the workshop (see below). Rather, these issues were not

always the central focal point of discussion, instead being situated within the context of broader discussions around the notions of ‘how to act ethically’. Nearly all of speakers spoke about ethics in this capacity.

Helen explained the importance of being aware that research ethics intersects many other ethics in our lives – our individual ethics, those of our professional , our institution, and our society, and also politically, and that these ethics may not always align. When we think about ethics we must consider this bigger picture.

Moreover, research ethics is relevant, said a number of speakers, at all stages of the research process – from when we choose a research question right through to the aftercare (which could be years later when, for example, a participant needs to contact a researcher to remove a picture they originally consented to being online). It’s important that researchers consider all of these aspects of ethics before they embark on their research as well as throughout the research process. Curtis described this as a *“holistically ethical approach to managing people’s data”*.

At the very beginning of the research process, says Sarah, when choosing our research question, we must ensure that the question has value or is of public benefit. Sarah explained that with so much data now available to researchers, we need to take the stance of, *‘just because it is there, doesn’t mean we should use it’*. Though Libby rightly questions how such value or benefit could be evaluated long-term.

Particularly at the presentation and dissemination stage, said Helen (drawing on work done with her colleague Lucy Pickering), researchers must consider how to report their research through an “ethics of engagement”, whereby we have an ethical duty to engage our audience with understandable findings. Sarah and Libby, as well as a number of other speakers also explained that when presenting our findings, it is imperative that are methods are transparent and explained. As ethical researchers, our transparent explanations of analytical methods and assumptions, of our algorithms, and of how we collected the data, ensure to the best of our abilities that our findings are not interpreted as the whole ‘truth’ but one version dependent on our biases.

Again, I use this opportunity to make a recommendation to re-evaluate how research ethics committees can accommodate this broader view of ethics. Helen explained that one of the issues

with viewing ethics as acting along as well as across the research process is that it does not ‘fit’ with our current research ethics committee structure. These committees inherited the ‘do no harm’ principle from the biomedical field where the potential for harm is focused on the data collection stage. Such a structure is inadequate for social media research, in which data collection accounts for only a minor aspect of the research project, and the possibility of ‘harm’ is enacted in different ways.

### 3. The perspective of social media users

Intersecting the above themes was the absolute requirement to consider social media users’ perspectives when embarking on research. This view has been discussed in the literature previously, but was strengthened dramatically at the workshop, with nearly all speakers pointing to the importance of such considerations for good ethical practice. Steve, Gabby, Sally and others related this back to the imperative of ensuring academics have a reputation as trustworthy researchers.

Social media users, the speakers explained, must be respected not just at the data collection stage, but right through to dissemination of the research. As Sarah remarked, and resonant of discourses in other research ethics fields, social media users must not be viewed as objects from which we extract data, but rather as collaborators in research. Sarah stressed, for example, that we do not consider social media users enough, especially in our research outputs, with little acknowledgement in our publications, nor in terms of developing outputs which are transparent and understandable (in terms of methodologies) for the social media user. Sally also asked, *“if you were a user of the [social media] group would it put you off to know that there was a researcher hovering around looking at everything you are writing?”* Some of Sally’s participants, and also those mentioned by other speakers such as Brittany’s, still have the perception that these spaces are private and researchers must respect this.

Sally provides a case example of the above perspective. She explained that typically researchers attempt to categorise the social media platforms they are researching as either private or public so as to help decide how to act ethically. However, she says this is unhelpful, not least because of the blurred boundaries - for example, some online spaces make members’ posts freely available but require registration to post a message. Rather, when making ethical decisions, we should ‘put ourselves in the shoes’ of social media users – this was something Brittany also spoke about in her

journey through researching online fan fiction writers. For Sally, a better question to ask is “*not is this group public or private, but what does this group mean to them [the social media user]?*” and from this, what is the risk of harm to the users? (including, what will be the impact of your research on the social media user?). This is especially important for sensitive and health online forums, says Sally, which may be seen as a ‘lifeline’ for users, or at least a very personal space. Brittany’s talk also illustrated similar concerns with social media users on fan fiction webpages, and Brittany highlighted the power of ‘vulnerability’ as her guiding ethical principle in these instances.

Steve provided some powerful data on social media user views. He explained that whilst in some instances, researchers may be legally permitted to use social media data, surveys reveal that large proportions of social media users actually do not want their data used for these purposes. Perspectives were context specific - dependent on the type of research being conducted and the institution conducting it. Public institutions conducting research for public benefit (Steve used the example of health) had the most positive support. This is unsurprising and is seen in other areas in which research tied to a public good such as health sees increased public support. A quick look at the literature reveals that other surveys have shown similar findings.

## Specific discussions and the negotiation of blurred boundaries

### 1. GDPR

The General Data Protection Regulation (GDPR) will affect all personal data (i.e., all data) collection and processing, including for social media research, from May 2018 (and will still apply post-Brexit). It requires that all data is processed in a lawful and transparent manner. Consent should be received, and where it cannot, data can only be collected if there is a legitimate interest to do so. If researchers are unsure about what this means, there is a legitimate interest assessment form which can be explored. This assessment includes, as just an example, that the rights of an individual (for example, social media user) are not overridden by the rights of the researcher; opt-out mechanisms are in place; data security is adhered to; and children’s data are excluded. Itzelle and Steve explained how discussions were still underway in terms of exactly what this regulation will mean for researchers and how these assessments impact social media research. Though Itzelle reminded us that most of the conditions of the GDPR are already implemented in academic institutions. Key aspects include only collecting data for a purpose (a

researcher must not collect data with the provision that it might be useful at some point), data security and data anonymisation.

## 2. Ethical approval

During the workshop a number of questions were asked regarding whether ethical approval was required for research using social media data. Whilst Gabby used her research to argue the need for all social media research to pass through an ethical review stage (though noted that this is not necessarily synonymous with requiring consent), both Gabby and Wasim noted that currently different institutions have very different practices in terms of whether ethical approval is required or not.

## 3. Consent

The question of whether researchers need to receive consent from social media users is well versed in the literature. The sense coming from the workshop was that receiving consent should be the default position, but in many instances consent may be unfeasible, in which case steps should be taken to ethically account for this.

Helen reminded us of the need to think about what we are receiving consent for. Will we want to use the data in years to come, and if so, how can we ensure the consent process adequately allows this? (Sarah described this as the blurred boundary between past and present.) It is not always sufficient to add the information to a consent form, which we know participants rarely read (a vast amount of research in many fields has shown how consent is often more related to trust in the researcher than an individual being 'informed' about the project and making a rational choice). We must allow for this and set systems in place to remain re-contactable for periods of time after the research project has ended in case social media user participants wish to withdraw from the research years later (for example, having information about themselves removed from an online medium).

Curtis suggested, and other speakers explained, that layered online consent forms can be helpful. Curtis's group's consent form had a short introductory paragraph explaining the project, with a number of links below this, supplying more information to the potential participant about, for example, how to opt out of the research, a withdrawal protocol, and detailed explanations of the



methodological approaches used during the research (as per the ‘transparency’ point discussed above). In respect of what Helen discussed above, it may be important to think about the length of time we need to leave these consent forms online.

For Wasim, who processed and analysed a large number of tweets in his research, consent could not be received for all Twitter users. Rather, he used a Google page to request consent only from those social media users’ tweets who he wished to quote in his research. However, because his project was historical this became difficult, so rather than quoting tweets directly in his thesis, he changed their wording. He reminded us that whilst Twitter does not permit the re-wording of tweets in publications, as yet, Twitter has taken no action against this.

Sally explained that online forums have smaller audiences and consequently a greater perception of privacy. It is here that consent is crucial, especially given that in some forums, particularly in, for example, the health and support arena, these sites may be the only places their users feel that they can openly and honestly discuss personal matters not discussed elsewhere in the ‘real world’. Though Sarah raised questions about from whom we should receive consent. Discussions in the literature suggest that moderator consent in online forum research is acceptable. However, Sarah questioned whether it is ethical for a site moderator to act as a representative for all users of said site.

### 3. Data privacy, anonymisation and data security

Curtis’s discussion of his research exemplified that data anonymisation as a default to protect the privacy of social media users is unhelpful. Other speakers concurred, explaining how anonymised data can often still be traced back to an individual. Anonymisation is still vitally important, but in line with other ethical approaches. For example, Sally raised questions about whether a researched online forum needs to be anonymised or just the individual social media users. Sally describes that the answer to this question will depend on which online forum is being researched and that it is a “*balancing act of principles rather than hard and fast rules*”. For Wasim and Sally, rather than just anonymising quotes they used from social media platforms, they re-word them too - though Sally noted this gets difficult when you have a number of quotes from one individual, in which case a description of those quotes may be more appropriate. Curtis explained how data security was an approach he had taken to ensure social media user privacy in his research, in which the data was not fully anonymised.

#### 4. Academic versus industry social media research

Wasim explained that industry Twitter research is often related to the profiling of individuals for marketing campaigns, and for industry, Twitter's terms and conditions are only relevant to ensure they remain within the boundaries of law. In contrast, academics are more concerned with ethical and privacy implications of social media research. When an audience member questioned whether this extra layer of academic ethical governance is necessary, all speakers unanimously agreed that it was, and that the argument 'they do it so we can' was unhelpful and problematic. Speakers re-iterated the importance of universities acting with integrity, having high standards of ethical practice, so as to ensure the trust of society. Having said this, there was also a realisation of the blurry lines between academic and industry research. Libby explained that academics are increasingly required to form research collaborations with industry and/or can often wear two hats (academic and industry). This makes institutional distinctions sometimes difficult to apply.