

What ethical responsibilities do social media researchers have to report harmful or illegal content?

For researchers working on social media communities, especially in sensitive subject areas, there is a strong likelihood they will have encountered harmful or illegal content. In this post, Dr Liam McLoughlin discusses the ethical responsibilities researchers face when they encounter this material and considers when and how an ethical requirement to report content might supersede confidentiality.

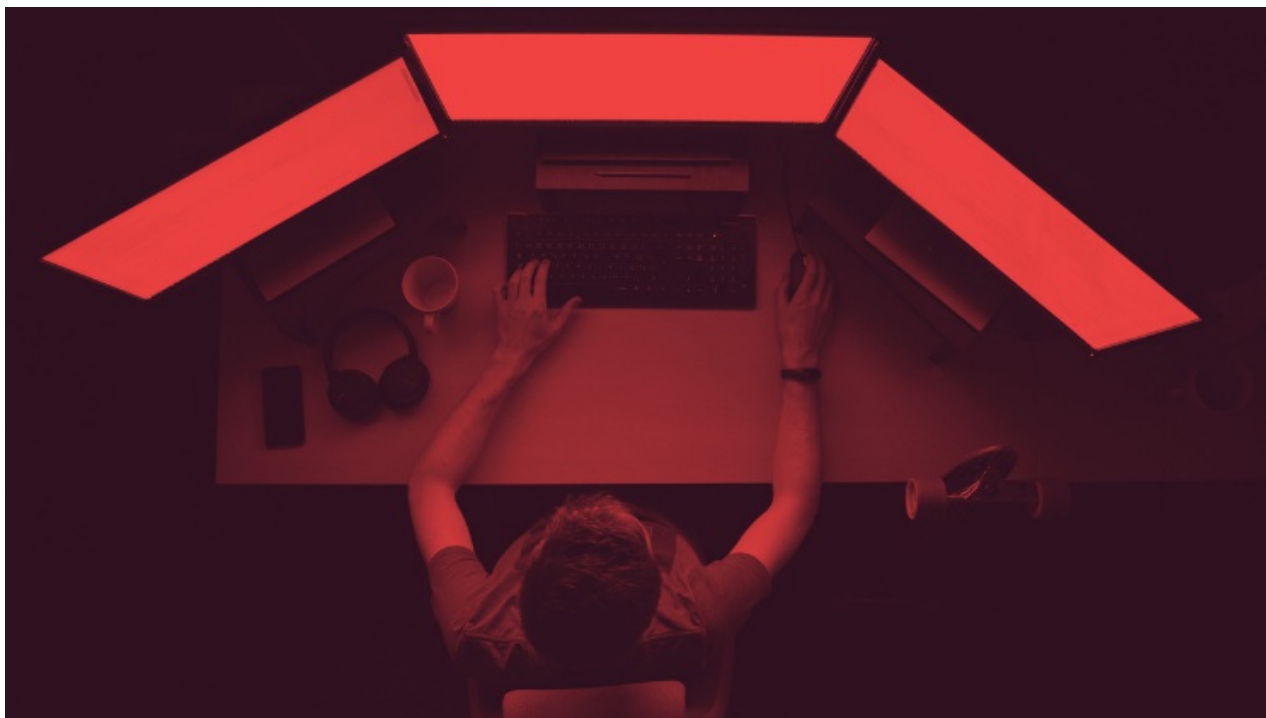
A question you should consider when [undertaking research on social media](#) is, what to do when you come across content which you feel should be reported? Spending enough time scrolling across Twitter, Facebook, or Instagram will likely lead a researcher, especially a qualitative one, to encounter abuse, harassment, illegal content and so on. For instance, [my own research alongside Dr Stephen Ward suggests](#) that researchers looking at tweets towards UK MPs will find around 2.7% of the collected messages contain abuse, hate-speech or even threats of offline violence. Implying that even within political communication, researchers might come across content that warrants at least reporting to platforms, if not highlighting to the police. However, there are certainly more extreme cases to consider. Numerous reports of damaging content such as [revenge porn](#), humiliating or unwanted exposures on [platforms such as Facebook](#) are not uncommon. Likewise, Instagram has had issues in [combating content which promotes self-harm or even suicide](#), which leaves much of this type of content easily assessable. To put it bluntly, not everything on social media is pretty. For some of this content, you might feel a moral obligation to report.

To put it bluntly, not everything on social media is pretty. For some of this content, you might feel a moral obligation to report.

Nevertheless, as an ethical social media researcher, what is your duty to report harmful posts? Is it even ethical to report content from your participants? You should already know about your duty of care towards those people whose posts or profile information might be included within your datasets (consent, privacy). But, there is relatively little information on the reporting of harmful or illegal content. To address these questions, I looked at the major ethical guidelines produced by the likes of the [Association of Internet Researchers \(AoIR\)](#), [British Sociological Association \(BSA\)](#), and others to assess how social media researchers can engage with these ethical dilemmas.

An emphasis on protecting participants

Within the [University of Sheffield's](#), [University of York's](#) & [University of Aberdeen's](#) social media research guidance, there is a significant degree of concern for how researchers should present content people have posted on social media. There is significant focus on how highlighting users posts in research could lead to them losing their jobs, social support and networks, or affect personal relationships. Likewise, the AoIR ethical guidelines and appendix argue that research should benefit those involved, not harm them ([Franzke, 2020: 71](#)). As such, one must appropriately handle social media content. However, to what extent should this be extended? Should researchers sit and watch while some social media users abuse or harass others, or should they report content they see?



On one side, as researchers, we often see ourselves as observers and nothing more. After all, we are viewing the content of others with their consent, either explicitly or implied. A level of consent that would not otherwise be granted if they thought we (as researchers) would report them. Indeed, out of 17 studies on social media research ethics by [Golder et al. \(2017\)](#), most studies found users were concerned about how their data is to be used – and this concern would undoubtedly increase if users thought researchers were also going to report their content to the platforms or authorities. Consequently, it is unlikely that social media users would grant access to researchers to semi-public spaces or private social groups if they thought researchers were to report content. Furthermore, general perceptions by the public that [research is generally good](#) would change – cutting researchers off from implied or explicit consent in many online spaces, and restricting available data.

What about protecting victims of online harms?

However, the above argument takes little account of content which might cause harm to others. Within AoIR guidance, there is some limited suggestion that there could be a responsibility for the researcher to act in scenarios of potential harm – mostly concerning mandatory reporters.

“there is often a professional and/or legal obligation for researchers to report such information to the relevant authorities: but this varies widely, first of all, depending on local and national legislation, specific professional ethical guidelines, and/or specific policies that may apply within a given research context” ([AoIR, 2019:17](#))

This raises serious ethical questions: What legal obligations are researchers under, especially non-medical researchers or those not working with young people? Upon viewing online content which causes harm, should researchers let the harm continue, or seek to take action to stop violence against others? Furthermore, it does not answer the question of content which breaks the law – which, there is often [no legal obligation](#) in most instances (at least in the UK and US) to report in the large majority of cases.

Towards an ethical solution

One suggestion, derived from the visual methodologies, is that images that break the law, “including criminal damage, sexual violence and hate crime do not have the privilege of confidentiality” (BSA 2017). One would assume that this would apply to social media posts and the privilege of confidentiality by their authors. This seems to be an approach agreed upon by Barnes *et al.* (2015) who argue that reporting harmful content should be applied on a case-by-case basis; with researchers making an evaluation based on the likelihood of harm gauging the seriousness and severity of illicit activity.

questions should be raised about the degree to which academics are prepared to encounter online harms and violence

The above solution is certainly one that works in open and public social media platforms, where the data is collected through implied consent, but what about within closed or semi-private social media spaces, such as Facebook Groups or WhatsApp conversations? In this situation, it would be suitable to notify all participants within consent forms that the groups’ contents will be treated with confidentiality and anonymisation for publication, but there are limits. Thus participants should be notified that if illegal and harmful content is found, it will be reported to the lead researcher(s) who will consider if there is the prospect of harm to themselves or others, and if so they would be obliged to report it to the relevant body (Hennell, Limmer & Piacentini, 2019).

Of course, this is not going to be applicable for all contexts. In cases where the researcher is explicitly researching online spaces where online harms or directed hate are expected, such as alt-right communities, this approach obviously will not be workable. Besides, further questions should be raised about the degree to which academics are prepared to encounter online harms and violence. Alongside pre-existing questions regarding the support for researchers who have viewed distressing content, do researchers and ethics committees have access to materials to prepare themselves for making a judgement on these issues? Furthermore, should the research community agree on standards for what action is appropriate when reportable content is identified?

While the above summarises the ethical implications of reporting harmful social media content, with a potential solution, it highlights an ever-persistent issue faced by researchers in new disciplinary areas. Ethical guidance is rarely a one-size fits all solution. New developments in communication technology means it’s up to us to find these creative solutions to the persistent grey areas within social media research. As more of these issues are made apparent, we should seek to share these solutions with others to inform, but also spark debate. At the same time, relevant stakeholders in the ethics process need to be given the time, skills, and resources to navigate the complex issues which surrounding social media research.

Note: This article gives the views of the author, and not the position of the Impact of Social Science blog, nor of the London School of Economics. Please review our [Comments Policy](#) if you have any concerns on posting a comment below.

Image Credit: Adapted from [Max Duzij](#) via Unsplash.
