



A call for an ethical framework when using social media data for artificial intelligence applications in public health research

Jean-Philippe Gilbert^{1*}, Victoria Ng², Jingcheng Niu³, Erin E Rees²

Abstract

Advancements in artificial intelligence (AI), more precisely the subfield of machine learning, and their applications to open-source internet data, such as social media, are growing faster than the management of ethical issues for use in society. An ethical framework helps scientists and policy makers consider ethics in their fields of practice, legitimize their work and protect members of the data-generating public. A central question for advancing the ethical framework is whether or not Tweets, Facebook posts and other open-source social media data generated by the public represent a human or not. The objective of this paper is to highlight ethical issues that the public health sector will be or is already confronting when using social media data in practice. The issues include informed consent, privacy, anonymization and balancing these issues with the benefits of using social media data for the common good. Current ethical frameworks need to provide guidance for addressing issues arising from the use of social media data in the public health sector. Discussions in this area should occur while the application of open-source data is still relatively new, and they should also keep pace as other problems arise from ongoing technological change.

This work is licensed under a [Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/).



Affiliations

¹ Université Laval, Québec, QC

² Public Health Agency of Canada, Ottawa, ON

³ University of Toronto, Toronto, ON

*Correspondence:

jean-philippe.gilbert.5@ulaval.ca

Suggested citation: Gilbert J-P, Ng V, Niu J, Rees EE. A call for an ethical framework when using social media data for artificial intelligence applications in public health research. *Can Commun Dis Rep* 2020;46(6):169–73. <https://doi.org/10.14745/ccdr.v46i06a03>

Keywords: ethics, ethical research, social media, artificial intelligence

Introduction

Rapid technological advancements in artificial intelligence (AI), and more specifically, natural language processing (NLP) using machine learning techniques, are enabling easy access and use of open-source big data. NLP allows computers to analyze datasets of natural language discourse (i.e. text not structured for quantitative analysis).

In public health, digital epidemiology has emerged as a new field that focuses on using non-public health sector data such as open-source internet data (e.g. Google Trends, news media) and social media data (e.g. Twitter and Facebook posts), whereas traditional epidemiology uses data collected for the purposes of health care, such as reporting of notifiable diseases by healthcare professionals to contribute to data for the surveillance of disease cases.

Researchers and policy makers recognize the potential of digital epidemiology data for advancing early warning of public health threats (1–3). Odium & Yoon (4) used NLP to assess Twitter data and reported that Tweets related to Ebola increased in the days

leading up to the official alert of the 2014 Ebola outbreak in Africa. Yousefinaghani *et al.* (5) showed that 75% of real-time outbreak notifications of avian influenza were identifiable from Twitter; one-third of outbreak notifications were reported on Twitter earlier than official reports. These observations support using Twitter volumes to predict the occurrence of outbreaks, and even forecast expected case counts, has also been shown with Google Trends data (1,6). Furthermore, refinement of social media data into various disease-relevant categories, by using NLP to classify Tweets into symptom types (e.g. fever, vomit), or focusing analysis on specific search terms from Google Trends, helps increase the accuracy in predictions of outbreak occurrence and forecast estimates.

Research that uses data from human participants requires ethical approval. A review process by a government body or university committee independent of the researchers assesses if use of these data ensures the safety, dignity and rights of the participants. Researchers need to demonstrate to the research ethics board (REB) that their study minimizes harm to participants



and respects their autonomy, generates and maximizes benefit (e.g. to society, science, participants) and acts with integrity, fairness and transparency to all stakeholders (e.g. participants, beneficiaries of the research). However, in a systematic review of the utilization of Twitter for health research, only 32% of the studies acquired ethical approval (7).

This is an example of technology moving faster than policy, in that the availability of newer data sources, such as from social media, have outpaced the need to assess the ethics of their use. This has led to studies with questionable ethical actions, which casts a shadow on all fields that use big data. An example is the “Tastes, Ties, and Time” study in 2007, where the researchers published an anonymized dataset of a group of university students and a codebook with information about the dataset; the dataset was identifiable from the codebook (8). Similarly, in 2012, evidence of online emotional contagion was sought, without prior consent, by manipulating the Facebook news feed of thousands of people to see if doing so changes sentiments in individuals’ posts (9).

In this article, we explore issues to do with traditional ethical frameworks in relation to research based on AI, particularly in the field of public health and digital epidemiology. We then present ethical frameworks that allow scientists and policy makers to use data from social media and their applications.

Contemporary ethics

In contemporary science, researchers need ethical approval for the use of human data. This very criterion is the main problem in big data-based research. It raises a seemingly simple question: Does a post or a Tweet represent human data or text data? (10). Several issues and points of view arise from this question, leading to a necessary debate given that the popularity of using social media data is increasing in several scientific fields, including digital epidemiology.

Currently, studies that use social media data are usually perceived as outside the scope of ethics committees’ evaluation because these data are commonly not considered to be human data (11,12). Many researchers, policy makers and practitioners assume that they can use open-source data, for example, Tweets, public posts on Facebook, public photos on Instagram and Google Trends queries, which do not require passwords to access (8,13). However, for many users of social media, posting publicly does not equate with giving their consent for the post to be used for research (8,11,12). This issue is not covered by existing ethical review mechanisms (14).

Furthermore, the ease of access to social media data (in the absence of ethical regulations and using rapid data capture via AI) means that the number of data points is often much larger than from traditional epidemiological datasets. Therefore, decisions about the use and implications of social media data

can potentially affect more people (14). For example, the number of people accidentally or maliciously reidentified in a Twitter database is only limited by the resources used to compile and analyse the database, which is far less than traditional surveillance systems (14).

Informed consent

Informed consent in the way it exists in contemporary ethics fits poorly with social media data. Firstly, it is almost impossible to obtain the informed consent of people whose data contribute to digital epidemiology because there are often insufficient resources to contact such high numbers of people who can be living anywhere (15).

Secondly to obtain informed consent, scientists need to confirm the identity of the social media users (16). There is no way to ensure that the person behind the social media profile is who they claim to be or to confirm whether the social media post was not generated by a bot (i.e. “robot” responsible for computer-generated social media posts). Because of this complication, some researchers consider consent to the terms and services of a social media platform, which users must give to use the platform, to be a surrogate for informed consent (16). However, users often do not read the terms and services or understood them well (17–19); nor do these stipulate the terms and conditions under which the data will be used for research, which calls into question the legitimacy and integrity of using terms and services as a surrogate for informed consent. Many “participants” in digital epidemiology are not aware that their data were collected or used (20).

Privacy and anonymization issues

We are becoming increasingly reliant on technology to structure and analyze the data proliferating in our digital societies. Data mining helps researchers find complex and unintuitive data patterns. However, data mining methods can also reveal confidential information from seemingly harmless social media data, for example, political affiliations (12,21). In addition, Wang *et al.* (22) reported being able to identify people’s sexual orientation by processing pictures of people from a dating website.

An anonymized dataset is the minimal requirement to protect the identity of subjects in social science (23) or in traditional epidemiology (20). According to the Common Rule, also known as 45 CFR 46 Subpart A, the principal regulation for human research from the Department of Health and Human Services of the United States (24), 17 identifiers need to be removed to consider a dataset anonymized. These include, among others, name, location of residence, all dates except the year and biometric identifiers (25). The Canadian Institutes of Health Research (CIHR), the Natural Sciences and Engineering Research Council of Canada (NSERC) and the Social Sciences and Humanities Research Council (SSHRC), identify similar identifiers (26). However, removing the 17 Common Rule identifiers is often not enough to ensure a dataset is anonymized.



This is because social media data are highly complex (i.e. have high dimensionality). Many non-traditional attributes can enable identification, such as reidentification from assessing the structure of the social networks (i.e. human connections) from multiple social media platforms (15,27). The advancements in AI algorithms and computational power to extract information and assess patterns means it is no longer possible to have anonymous databases (28,29). Many examples in the scientific literature demonstrate this issue by reidentifying an anonymized and subsequently published dataset (12,21).

The common good

The common good takes roots in the utilitarian vision of ethics. In this vision, the common good that research can do is considered versus the potential harm to individuals. A certain level of harm can be tolerated if the result is "positive morality". In the context of social media, the harm is mostly an invasion of privacy (30). People are more willing to sacrifice their privacy if they perceive that usage of their data will benefit the common good (31,32). For the most enthusiastic social media users in the Mikal *et al.* study (31), "it's cool when it's stuff [...] like the flu, because then that's how [public health decision-makers] know to get the vaccines to a place." Similarly, for the social media users in the Golder *et al.* study (32), it "could give a voice to patients and others groups, uncover true prevailing issues, and improve patient care." Factors that influence people's compliance in sharing their data for the common good include the type of research and the researchers affiliations (i.e. university, company, government) (32–34).

Ultimately, while the majority of people agree with the concept of the common good, there is no agreed-upon threshold for which an invasion of privacy can, and should, be tolerated for public health research.

New ethical frameworks

New frameworks that respond to new ethical challenges regarding the use of AI for research have been proposed by the Association of Internet Researchers (AoIR) (35) and Zook *et al.* (36) (Table 1).

Following a framework can help to legitimize research for the population (37). Since the AoIR framework (35) is accepted in the scientific literature, with the Association being one of the most cited organizations in terms of ethics and big data, scientists may want to use this framework rather than the lesser-known Zook *et al.* framework. However, the Zook *et al.* (36) framework is less restrictive and easier to follow.

Many points in these guidelines are already considerations that public health scientists have to address (e.g. protection of the vulnerable population, the potential harms of the study, the anonymization process). Public health scientists already frequently use highly confidential data. The main difference between social media data and traditional data is the way the data are accessed; the original intent for which the data

Table 1: Proposed ethical frameworks

Authors	Guidelines
AoIR (35)	<ol style="list-style-type: none"> 1. Protect vulnerable populations 2. Assess potential harm from research studies on a case-by-case basis 3. Consider data from humans to be human 4. Balance the rights of all involved parties (i.e. the right of privacy for the subject and the right to do research for the scientist) 5. The temporal variability of ethical considerations must be resolved when it occurs 6. Discuss ethical problems with qualified professionals when these arise
Zook <i>et al.</i> (36)	<ol style="list-style-type: none"> 1. Acknowledge that data are people and can do harm 2. Recognize that privacy is more than a binary value 3. Guard against the reidentification of your data 4. Practice ethical data sharing 5. Consider the strengths and limitations of your data; big does not automatically mean better 6. Debate the tough, ethical choices 7. Develop a code of conduct for your organization, research community or industry 8. Design your data and systems for auditability 9. Engage with the broader consequences of data and analysis practices 10. Know when to break these rules

Abbreviation: AoIR, Association of Internet Researchers

are produced; and the limited ability for social media users to provide informed consent. The data still represent humans, and can result in unintentional consequences such as identifying the individual behind their social media content. Public health scientists have an obligation to protect the individuals behind their data while balancing this with the common good; this subjective decision is extremely difficult to agree upon.

Discussion

As technology advances rapidly and more research is done with AI and social media data, an established ethical framework is essential to prevent improper use of social media data in public health applications. Researchers in public health, computer science and ethics need to come together to develop a framework that will help scientists conduct responsible research. In general, existing frameworks have been developed for use in every scientific field. Public health-related decisions can have an important impact on the population, however, going as far as to restrict the freedom of movement of persons in the case of a highly infectious disease, as an example (20).

The REB is an important part of the process to ensure the research is within the ethical framework. Inherent in using open-source social media data is that people do not know, or do not have the opportunity to consent, with their data being used. Thus, the REB provides the means to defend the safety, dignity and rights of the participants as stipulated through the ethical framework.



The REB and ethical framework are also needed to address the limitations of social media data. Many social media platforms are available, and the predominance in their use can differ by location. For example, Twitter and Facebook are used extensively in Western countries but banned in the People's Republic of China; the Chinese government authorizes the use of Sina Weibo and WeChat as the respective Twitter and Facebook equivalents. Furthermore, the demographics of use can vary among applications. Older generations tend to use Twitter and Facebook, while younger generations tend to use Snapchat, Instagram and TikTok. This is known as the digital divide (38). Some profiles may be underrepresented (e.g. children and elderly), depending of the social media platforms.

Conclusion

The ethical issues to do with using social media data for AI applications in public health research centre around whether these data are considered human. Current ethical frameworks are inadequate for public health research. To prevent further misuse of social media data, we argue that considering social media to be human would facilitate an REB process that ensures the safety, dignity and rights of social media data providers. We further propose that there needs to be more consideration towards the balance between the common good and the intrusion of privacy. Collaboration between ethics researchers and digital epidemiologists is needed to develop ethics committees, guidelines and to oversee research in the field.

Authors' statement

JPG — Writing—original draft, project administration, conceptualization

VN — Writing—reviewing & editing, conceptualization, supervision

NJ — Writing—reviewing & editing, conceptualization, supervision

EER — Writing—reviewing & editing, conceptualization, supervision

Conflict of interest

None.

Acknowledgements

The authors would also like to acknowledge S de Montigny, N Barrette and P Gachon for their comments.

Funding

This work is supported by the Public Health Agency of Canada.

References

1. Ginsberg J, Mohebbi MH, Patel RS, Brammer L, Smolinski MS, Brilliant L. Detecting influenza epidemics using search engine query data. *Nature* 2009;457(7232):1012–4. [DOI](#) [PubMed](#)
2. Barboza P, Vaillant L, Mawudeku A, Nelson NP, Hartley DM, Madoff LC, Linge JP, Collier N, Brownstein JS, Yangarber R, Astagneau P; Early Alerting Reporting Project Of The Global Health Security Initiative. Evaluation of epidemic intelligence systems integrated in the early alerting and reporting project for the detection of A/H5N1 influenza events. *PLoS One* 2013;8(3):e57252. [DOI](#) [PubMed](#)
3. Salathé M. Digital epidemiology: what is it, and where is it going? *Life Sci Soc Policy* 2018;14(1):1–5. [DOI](#) [PubMed](#)
4. Odlum M, Yoon S. What can we learn about the Ebola outbreak from tweets? *Am J Infect Control* 2015;43(6):563–71. [DOI](#) [PubMed](#)
5. Yousefinaghani S, Dara R, Poljak Z, Bernardo TM, Sharif S. The assessment of Twitter's potential for outbreak detection: avian influenza case study. *Sci Rep* 2019;9(1):18147. [DOI](#) [PubMed](#)
6. Rangarajan P, Mody SK, Marathe M. Forecasting dengue and influenza incidences using a sparse representation of Google trends, electronic health records, and time series data. *PLOS Comput Biol* 2019;15(11):e1007518. [DOI](#) [PubMed](#)
7. Sinnenberg L, Buttenheim AM, Padrez K, Mancheno C, Ungar L, Merchant RM. Twitter as a tool for health research: a systematic review. *Am J Public Health* 2017;107(1):e1–8. [DOI](#) [PubMed](#)
8. Zimmer M. "But the data is already public": on the ethics of research in Facebook. *Ethics Inf Technol* 2010;12(4):313–25. [DOI](#)
9. Joughi J, Lauk E, Penttinen M, Sormanen N, Uskali T. Facebook's emotional contagion experiment as a challenge to research ethics. *Media Commun* 2016;4(4):75–85. [DOI](#)
10. Buchanan E. Considering the ethics of big data research: A case of Twitter and ISIS/ISIL. *PLoS One* 2017;12(12):e0187155. [DOI](#) [PubMed](#)
11. Fiesler C, Proferes N. "Participant" perceptions of Twitter research ethics. *Social Media Soc* 2018;4(1):14. [DOI](#)
12. Ienca M, Ferretti A, Hurst S, Puhan M, Lovis C, Vayena E. Considerations for ethics review of big data health research: A scoping review. *PLoS One* 2018;13(10):e0204937. [DOI](#) [PubMed](#)
13. Gehner M, Oughton D. Ethical challenges in social media engagement and research: considerations for code of engagement practices. *J Radiol Prot* 2016;36(2):S187–92. [DOI](#) [PubMed](#)



14. Lee EC, Asher JM, Goldlust S, Kraemer JD, Lawson AB, Bansal S. Mind the scales: harnessing spatial big data for infectious disease surveillance and inference. *J Infect Dis* 2016;214 suppl_4:S409–13. [DOI PubMed](#)
15. Lipworth W, Mason PH, Kerridge I, Ioannidis JP. Ethics and epistemology in big data research. *J Bioeth Inq* 2017;14(4):489–500. [DOI PubMed](#)
16. Webb H, Jirotko M, Stahl BC, Housley W, Edwards A, Williams ML, Procter R, Rana OF, Burnap P. The Ethical Challenges of Publishing Twitter Data for Research Dissemination. Proceedings of the 2017 ACM Web Science Conference. 2017, New York: Assoc Computing Machinery; 339–48.
17. Fiske ST, Hauser RM. Protecting human research participants in the age of big data. *Proc Natl Acad Sci USA* 2014;111(38):13675–6. [DOI PubMed](#)
18. Saqr M. Big data and the emerging ethical challenges. *Int J Health Sci (Qassim)* 2017;11(4):1–2. [PubMed](#)
19. Hesse A, Glenna L, Hinrichs C, Chiles R, Sachs C. Qualitative research ethics in the big data era. *Am Behav Sci* 2019;63(5):560–83. [DOI](#)
20. Vayena E, Salathé M, Madoff LC, Brownstein JS. Ethical challenges of big data in public health. *PLOS Comput Biol* 2015;11(2):e1003904. [DOI PubMed](#)
21. Mooney SJ, Pejaver V. Big data in public health: terminology, machine learning, and privacy. *Annu Rev Public Health* 2018;39:95–112. [DOI PubMed](#)
22. Wang Y, Kosinski M. Deep neural networks are more accurate than humans at detecting sexual orientation from facial images. *J Pers Soc Psychol* 2018;114(2):246–57. [DOI PubMed](#)
23. Phillips A, Borry P, Shabani M. Research ethics review for the use of anonymized samples and data: A systematic review of normative documents. *Account Res* 2017;24(8):483–96. [DOI PubMed](#)
24. Shade J, Coon H, Docherty AR. Ethical implications of using biobanks and population databases for genetic suicide research. *Am J Med Genet B Neuropsychiatr Genet* 2019;180(8):601–8. [DOI PubMed](#)
25. Rothstein MA. Is deidentification sufficient to protect health privacy in research? *Am J Bioeth* 2010 Sep;10(9):3–11. [DOI PubMed](#)
26. Canadian Institutes of Health Research. Natural Sciences and Engineering Research Council of Canada, Social Sciences and Humanities Research Council, Tri-Council policy statement: ethical conduct for research involving humans – TCPS 2 (2018). Ottawa (ON): Government of Canada; 2018 Dec. <https://ethics.gc.ca/eng/documents/tcps2-2018-en-interactive-final.pdf>
27. Vayena E, Haeusermann T, Adjekum A, Blasimme A. Digital health: meeting the ethical and policy challenges. *Swiss Med Wkly* 2018;148:w14571. [DOI PubMed](#)
28. El Emam K, Jonker E, Arbuckle L, Malin B. A systematic review of re-identification attacks on health data. *PLoS One* 2011;6(12):e28071. [DOI PubMed](#)
29. de Montjoye YA, Radaelli L, Singh VK, Pentland AS. Identity and privacy. Unique in the shopping mall: on the reidentifiability of credit card metadata. *Science* 2015;347(6221):536–9. [DOI PubMed](#)
30. Herron M, Sinclair M, Kernohan WG, Stockdale J. Ethical issues in undertaking internet research of user-generated content: a review of the literature. *Evid Based Midwifery* 2011;9(1):9–15.
31. Mikal J, Hurst S, Conway M. Ethical issues in using Twitter for population-level depression monitoring: a qualitative study. *BMC Med Ethics* 2016;17(1):22. [DOI PubMed](#)
32. Golder S, Ahmed S, Norman G, Booth A. Attitudes toward the ethics of research using social media: a systematic review. *J Med Internet Res* 2017;19(6):e195. [DOI PubMed](#)
33. Beninger K, Fry A, Jago N, Lepps H, Nass L, Silvester H. Research using social media; users' views. London (UK): NatCen Social Research; 2014. https://www.researchgate.net/publication/261551701_Research_using_Social_Media_Users%27_Views
34. Golder S, Scantlebury A, Christmas H. Understanding public attitudes toward researchers using social media for detecting and monitoring adverse events data: multi methods study. *J Med Internet Res* 2019;21(8):e7081. [DOI PubMed](#)
35. Markham A, Buchanan E. Ethical decision-making and internet research: recommendations from the AoIR Ethics Working Committee (Version 2.0). Association of Internet Researchers; 2012.
36. Zook M, Barocas S, Boyd D, Crawford K, Keller E, Gangadharan SP, Goodman A, Hollander R, Koenig BA, Metcalf J, Narayanan A, Nelson A, Pasquale F. Ten simple rules for responsible big data research. *PLOS Comput Biol* 2017;13(3):e1005399. [DOI PubMed](#)
37. Shilton K, Sayles S. "We aren't all going to be on the same page about ethics:" Ethical practices and challenges in research on digital and social media. In: Bui TX, Sprague RH, editors. Proceedings of the 49th Annual Hawaii International Conference on System Sciences. 2016, IEEE Computer Soc: Los Alamitos. 1909–18.
38. Hargittai E. Digital Na(t)ives? Variation in internet skills and uses among members of the "Net Generation". *Sociol Inq* 2010;80(1):92–113. [DOI](#)