

# Social Media Research Ethics: Preliminary Considerations

Emily Van Haren, Andrea Zeffiro  
Isaac Pratt & Jay Brodeur

Do More with Digital Scholarship Workshop Series  
January 18, 2021



McMaster University sits on the traditional Territories of the Mississauga and Haudenosaunee Nations, and

*McMaster University sits on the traditional Territories of the Mississauga and Haudenosaunee Nations, and within the lands protected by the “Dish With One Spoon” wampum agreement.*

# Module Goals

*Welcome to the Social Media Research Ethics: Preliminary Considerations module!*

Today we aim to:

- introduce you to formative considerations for integrating social media platforms and data into your research
- follow a question-based approach, with accompanying case studies, in order to connect the methodological and ethical considerations of social media research to specific examples
- draw out your own insights into, and questions about, social media research ethics

# Case Studies

*Why explore social media research ethics through case studies?*

Every social media research context is unique, which means that it is impossible to adopt a 'one size fits all' approach to social media research ethics.

Case studies therefore allow researchers to explore the methodological and ethical considerations that arise in particular research and social media contexts.

This situated exploration is not a one-off review, but an iterative and deliberative process.

Worksheet available here: <https://cutt.ly/ZjvtmLQ>

**Please pause this video and review  
the case studies in the worksheet!**

# Methodological Considerations

*What are social media, and what types of social media platforms and data might I work with?*

Broadly defined as “online, often mobile platforms that support the creation and exchange of user-generated content” (Taylor & Pagliari 2018).

- Networking, content and information sharing (e.g. Facebook, Twitter, LinkedIn)
- Community-specific forums (e.g. SteadyHealth)
- Private collaboration tools (e.g. Slack, Asana, Trello)
- Crowdsourcing platforms (e.g. GoFundMe)

Data can be big or small and types include text-based, video, audio, photographs, and digital trace data.



# Methodological Considerations

*What kinds of analysis might I conduct?*

- **Computational** (e.g. text mining, natural language processing, sentiment analysis). Automated collection and analysis, can include human analysis.
- **Quantitative** (e.g. statistical analysis, social network analysis). Human and machine collection and analysis.
- **Qualitative** (e.g. auto/ethnography, discourse analysis, content analysis, close reading). Human data collection and interpretation.
- **Mixed** (e.g. discourse analysis and social network analysis). Human/machine data collection and analysis.

# Methodological Considerations

*What should I consider about the availability, quality, and provenance of social media data?*

- **Representativeness:** the self-selecting nature of social media users.
- **Access:** unequal access to social media platforms and data.
- **Non-traditional sampling:** observing and interpreting what is “out there”.
- **Heterogeneity:** data of variable quality and provenance.
- **Terms and Conditions:** conditions of use laid out by various social media service providers.

(Taylor & Pagliari 2018; Munson et al., 2013)



# Methodological Considerations

*How can I organize, secure, and backup that social media data?*

It all starts with a good plan: a Data Management Plan is a **living document** that outlines:

- how data will be collected, documented, formatted, protected, and preserved
- how existing datasets will be used and what new data will be created
- whether and how data will be shared, and where data will be deposited
- roles and responsibilities for managing the data, including succession plans
- ethical, legal and commercial constraints of the data
- methodological considerations that support or preclude data sharing

DMPs can be written using standardized tools like the [Portage Network's DMP Assistant](#).

- a web-based, bilingual data management planning tool
- available to all researchers in Canada
- a guide for best practices in data stewardship
- exportable data management plans

<https://www.youtube.com/watch?v=zgLaJpJfehQ>



<https://assistant.portagenetwork.ca/>



# Methodological Considerations

*How can I organize that social media data?*

- Tabular data should be organized following the **Tidy Data principles**:

- Each variable forms a column.
- Each observation forms a row.
- Each type of observational unit forms a table.

**Table 1: Cats**

id	name	colour
1	floof	gray
2	merlin	orange
3	panda	calico
4	cat	brown

# Methodological Considerations

*How can I organize that social media data?*

Files and folders should be named using a clear, consistent format which can include metadata such as:

- a summary of contents, such as GrantApplication
- date of collection/creation, formatted as YYMMDD
- file version, written as v1 or v2

Avoid using special characters that may not be consistent across different operating systems (MacOS, Windows, iOS) and data storage platforms:

- & , \* % # \* ( ) ! @\$ ^ ~ ' { } [ ] ? < > -

# Methodological Considerations

*How can I organize that social media data?*

Research Data metadata should be recorded as well, including:

- variable names and descriptors
- data collector
- date of collection

Various standardized metadata schemes exist such as Dublin Core or DDI.

# Methodological Considerations

*How can I secure and backup that social media data?*

Data storage options range from local hard drives and on-device storage to cloud based file systems with automatic backups and versioning.

- Who needs to be able to access your data? And from where and what kind of device?
- Should you be encrypting your data?
- How much storage space do you need?
- How will you manage data versions?

**3-2-1 rule for backups:** you should have 3 copies of your data, in 2 different storage types, with 1 being off-site.

# From methods to ethics (and back again) ...



# Ethical Considerations

*What are the ethical considerations and challenges raised by social media research contexts?*

Common ethical considerations centre upon four key, and interconnected, areas of concern:

- **Private vs. public**
- **Informed consent**
- **Anonymity**
- **Risk of harm**

# Ethical Considerations

*What **expectations of privacy** might be associated with this data?*

Understandings and expectations of privacy are deeply contextual and must be assessed on a case-by-case basis.

The usage norms of a platform can point to privacy expectations.

- e.g. password protection, pseudonyms, hashtags, @-mentions
- platform Terms of Service

**Recommendation:** the availability of social media data should not be conflated with its ethical use.

(boyd and Crawford 2012)

# Ethical Considerations

*What expectations of privacy might be associated with this data?*

An emphasis on expectations of privacy prompts researchers to consider the potential reluctance of users to have their online activities analyzed by a researcher.

- Do users believe or assume that their community/data is private or unobserved?
- Does the research require deception on the part of the researcher?

Data posted by public figures are intended to reach a wide public audience and therefore may anticipate engagement and critique.

- How is 'public figure' status determined or defined?

(Convery & Cox 2012)

# Ethical Considerations

*How do expectations of privacy relate to the question of **informed consent**?*

Achieving informed consent may be difficult or impossible due to the scale or anonymous nature of the data.

- informed consent vs. terms of service

However, informed consent can still be necessary within social media research.

- contextual approach to decision-making, based on expectations of privacy—but also anonymity and risk of harm

(Townsend & Wallace 2016)

# Ethical Considerations

*In what ways, and to what degree, can I **anonymize** this data?*

Strategies for de-identifying data are made more complex by social media contexts.

- some platforms may acquire full attribution
- searchability of content (e.g. anonymizing individual elements and quoted material is typically moot)
- different forms of identifiable data (e.g. text-based, image, video, audio)

Protecting the identify of unknowing participants is paramount, especially when dealing with sensitive subject matters.

(Townsend & Wallace 2016)

# Ethical Considerations

*What are the potential **risks** of this research for participants, and how can I communicate and mitigate those risks?*

Risk of harm can be increased when privacy is breached and data of a sensitive nature is made public.

Social media requires a re-scaling of risk assessment:

- communities rather than simply individuals
- networked or tagged users rather than individual accounts (e.g. photographs, @mentions or commenters)

(Townsend & Wallace 2016)

# Recommendations and Next Steps

*What are the next steps for my own social media research project?*

- **Review:** carefully read the terms and conditions of any specific platforms you will be using to obtain data, and review the institutional guidelines of your university, research organization, funding bodies, and publication/presentation venues.
- **Consult:** relationship with ethical boards like [McMaster Research Ethics Board \(MREB\)](#) is iterative and generative (rather than a box-ticking step).
- **Integrate:** ethical considerations should drive research questions and methodological decision-making, and be integrated into all stages of the project.



# Please pause this video and revisit the case studies in the worksheet!

- Is the data private?
- Can the subject matter be considered sensitive?
- Are any of the subjects vulnerable? What is the risk of harm?
- Is informed consent necessary? Has it been, or how will it be, obtained?
- How (if at all) should source information be presented in publications?
- How (if at all) should the data be shared?
- Should the researcher identify themselves?
- Is the research exploitative?
- Is the data representative?
- Is there a need to account for bots, trolls, and spam?
- Is ethics approval necessary?
- What recommendations might follow from an REB consultation?
- Are there other ethical and methodological considerations?

(Zeffiro & Brodeur 2020)

# Questions, Responses, and Feedback?

To share your thoughts on the methodological and ethical considerations of social media research, please complete our short feedback form:

- <https://forms.gle/rTbpa44FLa7A6APG7>

Please also join us for the follow up module——Social Media Research Ethics: Project Design——on February 22, 2021.

- <https://libcal.mcmaster.ca/calendar/scds/dmds-smre-project-design>

# Works Cited

- boyd, d. and Crawford, K. (2012). Critical questions for big data. *Information, Communication and Society* 15(5), 662-679.  
<https://doi.org/10.1080/1369118X.2012.678878>.
- Convery, I. & Cox, D. (2012). A review of research ethics in internet-based research. *Practitioner Research in Higher Education* 6(1) 50-57.  
<https://files.eric.ed.gov/fulltext/EJ1130546.pdf>.
- Munson SA, Cavusoglu H, Frisch L, Fels S, Sociotechnical Challenges and Progress in Using Social Media for Health. *J Med Internet Res* 2013;15(10):e226. DOI: [10.2196/jmir.2792](https://doi.org/10.2196/jmir.2792)
- Taylor, J. & Pagliari, C. (2018). Mining social media data: How are research sponsors and researchers addressing the ethical challenges? *Research Ethics* 14(2), 1-39. <https://doi.org/10.1177/1747016117738559>.
- Townsend, L. & Wallace, C. (2016). Social media research: A guide to ethics. *The University of Aberdeen*.  
[https://www.gla.ac.uk/media/Media\\_487729\\_smxx.pdf](https://www.gla.ac.uk/media/Media_487729_smxx.pdf).
- University Research Ethics Committee. (2016). The ethics of Internet-based and social media research. *University of Sheffield*.  
[https://www.sheffield.ac.uk/polopoly\\_fs/1.644904!/file/Report\\_Ethics\\_of\\_Social\\_Media\\_Research\\_Jul16.pdf](https://www.sheffield.ac.uk/polopoly_fs/1.644904!/file/Report_Ethics_of_Social_Media_Research_Jul16.pdf).
- von Benzon, N. (2018). Informed consent and secondary data: Reflections on the use of mothers' blogs in social media research. *Area* 51(1), 182-189. <https://doi-org.libaccess.lib.mcmaster.ca/10.1111/area.12445>.
- Zeffiro, A. & Brodeur, J. (2020, March 5). *DMDS: Social Media Research Ethics and Data Management* [Workshop Powerpoint slides]. Sherman Centre for Digital Scholarship. <https://macsphere.mcmaster.ca/handle/11375/25327>.

# Thank you!