

Secondary data and ethical issues

Dr Muir Houston

17/07/19

Disclaimer: this presentation is a general overview – it is the responsibility of the researcher to ensure that they conform to legal and ethical requirements

Types or forms of secondary data

- Secondary data containing personal and/or sensitive personal data
- Open access secondary data
- Purchased and/or licensed secondary data
- Restricted secondary data

Secondary data containing personal and/or sensitive personal data

- Research will require to comply with both GDPR and subject to ethical oversight and review
- Access through safe havens for 'official data' and subject to disclosure control – data points
- Social media and social networking data – blurs boundaries between primary and secondary data and also between public and private

Open access secondary data

- Dataset in the public domain and often considered free from restrictions
- Datasets freely available on the internet – some may require registration – but often as bot protection
- May still be subject to an ‘Open data’ or ‘Creative Commons’ licence and as such may still be subject to certain restrictions and limitations on use
- Generally would not be required to obtain ethical approval – but journals may require a statement to explain why approval not required

Purchased and/or licensed secondary data

- The terms and conditions of the license or contract from the data supplier or source must be adhered to; and, data should not be used, processed, shared, or distributed beyond the limits agreed or covered by the license
- House price data; other commercial datasets;
- Certain datasets in the UK Data Archive and other repositories
- Ethical approval if personal and/or sensitive personal data included

Restricted secondary data

- Administrative data
 - Health/medical data
 - Social security data
 - Criminal justice system data
 - Educational data
- Generally required to access in 'Safe Haven' can be physical or electronic; and subject to strict disclosure controls
- Ethical oversight – may result in confirmation that ethical approval is not required

Big Data, sharing and ethics

- Libby Bishop (2017). Big data and data sharing: Ethical issues. UK Data Service, UK Data Archive. https://www.ukdataservice.ac.uk/media/604711/big-data-and-data-sharing_ethical-issues.pdf
- Often differ from traditional research data in that they have not been generated specifically by researchers for research purposes. As a result, the usual ethical protections that are applied at several points in the research data life cycle have not taken place.
- *big data most commonly used for social research as administrative data, records of commercial transactions, social media and other internet data, geospatial data and image data. (OECD, 2013)*
- Potential issues of privacy and disclosure – ethical oversight and DPIIO input

Social media and online ethics

- What constitutes 'privacy' in an online environment?
- How easy is it to get informed consent from the participants in the community being researched?
- What does informed consent entail in that context?
- How certain is the researcher that they can establish the 'real' identity of the participants?
- When is deception or covert observation justifiable?
- How are issues of identifiability addressed?
- How do country-specific legal requirements (eg for data protection) apply in internet-mediated research that crosses national boundaries?

ESRC – internet mediated research

Online ethics- general

- Commonly cited sources:
 - Association of Internet Researchers (AOIR) -
 - 2012: [Ethical decision-making and Internet research 2.0: Recommendations from the AoIR ethics working committee](#) [PDF]
 - 2012: [This chart](#) provides a useful starting point for internet researchers to consider ethics.
 - 2002: [Ethical decision-making and Internet research: Recommendations from the AoIR ethics working committee](#) [PDF]

Issue to be addressed

- 1 Please provide details of the data you wish to collect or access – please include details of the platform, app, data archive, API, etc.
- 2 You should have consulted the specific Terms and conditions of the specific platform or data source; please answer the following:
 - 2.1 What do the terms and conditions say about retention of datasets?
 - 2.2 What are the rules regarding publishing or re-sharing collected data?
 - 2.3 Are there specific provisions within the terms and conditions that permit research usage of data collected?
 - 2.4 What are the explicit limits on usage that may be relevant for planned research work?
- 3 Have you consulted the relevant legal guidelines disciplinary, funder or institutional guidelines in relation to the specific ethical concerns research of this nature can raise? E.g. copyright/Intellectual Property Rights/contracts/licensing/ privacy/GDPR.

An example

- In 2015 Dan Gray, at the University of Cardiff, used Twitter to study misogynist speech. He encountered numerous legal and ethical challenges with consent and anonymisation when considering how to fairly represent research participants. He collected some 60,000 Tweets in 2015 by filtering on keywords of hateful speech and needed to be able to publish selected quotations of Tweets to support his arguments.

Issues, constraints and decisions

- Twitter's Terms and Conditions prohibit modifying content, meaning that tweets could not be anonymised.
- Gray had to decide if the Tweets could be considered public, and moreover, would their public status be sufficient to justify publishing without consent.
- Survey analysis done at the Social Data Science Lab at Cardiff, where Gray was connected, showed that Tweeters did not want their content used, even for research, if they were identifiable.
- If he did decide to seek consent, there was no way to do so as private communication to the Tweeter. This would have been possible only if the Tweeters were following him, and they were not.
- Mutual following was not possible as a way of contacting Tweeters because the Research Ethics Committee required that he use an anonymised profile.

- **Outcomes**

- He opted to contact by direct Tweet, though this risked allowing tweeters to find him, and also to contact other tweeters of hateful discourse.
- “Consent by Tweet” severely constrained his ability to explain risks and benefits of the research.
- Consent was successfully obtained for a number of tweets, enabling sharing of selected unanonymised tweets in publications.
- Gray was able to draw upon the UK’s COSMOS Risk Assessment for guidance, but points out that its rigorous attention to harm and privacy can become a barrier, shielding hateful discourse from critical scrutiny

Ethics in social media research: where are we now?

- Early on in the research we quickly realised that many of the learned society ethical resources were of little guidance, given their focus on non-digital data. Where addendums on using Internet data were written, they had little to say about social media. Papers were being published in reputable journals with tweets quoted verbatim, with unacceptable and ineffective methods of anonymisation, and without informed consent from users¹. Research on users' views of the repurposing of their social media data consistently shows that the majority wish to be asked for informed consent if their content is to be published outside of the platform which it was intended for².

Resources – do your research

- BPS (2017) Ethics Guidelines for Internet-mediated Research [https://www.bps.org.uk/sites/bps.org.uk/files/Policy/Policy%20-%20Files/Ethics%20Guidelines%20for%20Internet-mediated%20Research%20\(2017\).pdf](https://www.bps.org.uk/sites/bps.org.uk/files/Policy/Policy%20-%20Files/Ethics%20Guidelines%20for%20Internet-mediated%20Research%20(2017).pdf)
- BPS (nd) Supplementary guidance on the USE OF SOCIAL MEDIA <https://www.bps.org.uk/sites/bps.org.uk/files/Policy/Policy%20-%20Files/Supplementary%20Guidance%20on%20the%20Use%20of%20Social%20Media.pdf>
- BSA (2017) Ethics Guidelines and Collated Resources for Digital Research https://www.britsoc.co.uk/media/24309/bsa_statement_of_ethical_practice_annexe.pdf
- Social Media Analytics: a survey of techniques, tools and platforms. Bogdan Batrinca and Philip C Treleaven *P.C. AI & Soc* (2015) 30: 89. <https://doi.org/10.1007/s00146-014-0549-4>
- Big data and data sharing: Ethical issues https://www.ukdataservice.ac.uk/media/604711/big-data-and-data-sharing_ethical-issues.pdf
- Richards, N.M. and King, J.H (2014) Big Data and Ethics, *Wake Forest Law Review* https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2384174#
- Zwitter, A. (2014) Big Data Ethics *Big Data & Society* July–December 2014: 1–6 <https://journals.sagepub.com/doi/pdf/10.1177/2053951714559253>
- Council for Big Data, Ethics, and Society <https://bdes.datasociety.net/>
- Uršič, Helena. 2019. “The Right to be Forgotten or the Duty to be Remembered? Twitter data reuse and implications for user privacy.” *Council for Big Data, Ethics, and Society*. Accessed March 26, 2019. <https://bdes.datasociety.net/council-output/the-right-to-be-forgotten-or-the-duty-to-be-remembered-twitter-data-reuse-and-implications-for-user-privacy/>
- Herschel, R. and Miori, V.M. (2017) Ethics & Big Data, *Technology in Society*, Vol 49, pp31-36.

- Ryan, L. (2016) Navigating ethics in the big data democracy *The Visual Imperative: Creating a Visual Culture of Data Discovery* 2016, Pages 61-84 – <https://www.sciencedirect.com/science/article/pii/B9780128038444000042>
- Internet based research: Best practice Guidance. University of Oxford
https://researchsupport.admin.ox.ac.uk/sites/default/files/researchsupport/documents/media/bpg_06_internet-based_research_v.5.2.pdf
- Research ethics and data sharing: governance and integrity Louise Corti 2018 <http://ukrio.org/wp-content/uploads/Louise-Corti-UK-Data-Service-Research-ethics-and-data-sharing.pdf>
- Unpacking the Value of Social Media Work Package 3: Ethics – Ginnis, S. et al., 2015
<https://www.ipsos.com/sites/default/files/ct/publication/documents/2018-01/wisdom-of-the-crowd-social-media-ethics.pdf>
- #SocialEthics a guide to embedding ethics in social media research Harry Evans; Steve Ginnis; Jamie Bartlet (2015)
<https://www.ipsos.com/sites/default/files/migrations/en-uk/files/Assets/Docs/Publications/im-demos-social-ethics-in-social-media-research-summary.pdf>
- Wisdom of the Crowd: Attitudes and expectations towards social media data and its uses
<https://www.slideshare.net/IpsosMORI/wisdom-of-the-crowd-attitudes-and-expectations-towards-social-media-data-and-its-uses>
- Digital ethnography and ethics in the context of web 2.0 Y. Morey and A. Bengry-Howell (2009) <https://core.ac.uk/display/1349757>
- Williams, M.L. et al (2017) Towards an Ethical Framework for Publishing Twitter Data in Social Research: Taking into Account Users' Views, Online Context and Algorithmic Estimation, *Sociology* Vol. 51(6) 1149–1168
- Michael Zimmer, Nicholas John Proferes, (2014) "A topology of Twitter research: disciplines, methods, and ethics", *Aslib Journal of Information Management*, Vol. 66 Issue: 3, pp.250-261, <https://doi.org/10.1108/AJIM-09-2013-0083>
- Zimmer, M. (2018). Addressing Conceptual Gaps in Big Data Research Ethics: An Application of Contextual Integrity. *Social Media + Society*, 4 (2). Available online at <https://doi.org/10.1177/2056305118768300>
- Zimmer, M. (2010). "But the data is already public": On the ethics of research in Facebook. *Ethics and Information Technology*, 12(4), 313-325.
- Van Baalen, S. (2018) 'Google wants to know your location': The ethical challenges of fieldwork in the digital age. *Research Ethics*, Vol. 14(4) 1–17.



“Social scientists do not have an unalienable right to conduct research involving other people (Oakes, 2002). That we continue to have the freedom to conduct such work depends on us acting in ways that are not harmful and are just. Ethical behaviour may help assure the climate of trust in which we continue our socially useful labours (AAAS, 1995; Jorgensen, 1971; Mitchell and Draper, 1982; PRE, 2002; Walsh, 1992). If we act honestly and honourably, people may rely on us to recognize their needs and sensitivities and consequently may be more willing to contribute openly and fully to the work we undertake.” (Israel and Hay, 2006: p3)

Israel, M. and Hay, I. (2006) *Research Ethics for Social Scientists*, London: Sage.