Acting ethically without approval:

venturing into online research

Dr Rachael Adlington

Introduction – online research

▶ Research that uses the Internet pushes many ethical boundaries and challenges traditional models of ethical research

▶ This is especially so for research that involves human-like subjects

　▶ Avatars, bots,

　▶ the products of humans located in the public domain

▶ It is important to consider ethics in any research

　▶ many issues and approaches pertinent to offline studies hold true online

▶ The online realm requires fresh thinking on what it means to act in an ethical manner

▶ Rethinking occurs around three points of tension:

　▶ human subjects, notions of 'personhood', and definitions of privacy

▶ Ethical principles:

　▶ consideration of people indirectly involved, harm minimization, rights of subjects vs social benefits, proportional obligation of the researcher to protect 'human element'

Introduction – online research

► Use **Association of Internet Researchers (AoIR) Ethics Working Committee recommendations** as a framework for acting ethically when undertaking online research (Markham & Buchanan, 2012)

► What constitutes *online research*

► Explore problematic area: argument for the need to act ethically to honour and protect the 'human element' in research that does not directly involve humans or human contact

► Focus: online texts

► Different approaches to text analysis -> view and manage ethics

► Acting ethically throughout the inquiry process

Online research

▶ *The widespread use of the Internet provides*

    ▶ *new vantage points from which to observe conventional behavior,*

    ▶ *views of new kinds of behavior, and*

    ▶ *new tools with which to observe it all.*

▶ *Accompanying these opportunities come two specific concerns about research approaches:*

    ▶ *how new research methods using the Internet may or may not affect the ethical protections to which human subjects are entitled, and*

    ▶ *the validity of data collected using the Internet.*

Walther, 2002, p. 205.

Online research

- Research on/about the happenings of the Internet (*observing conventional and new behavior*)
  - Involving/observing interactions, texts, spaces
  - To understand how people interact, produce texts, shape spaces
  - In order to build identity, influence, communicate
  - And the impact of these spaces, texts on people
- Research using tools of the Internet (*new tools with which to observe it all*)
- Located increasingly with a foot in both camps
  - Internet-based research (where research occurs; on or with)
  - Internet-mediated research (how it occurs; through)

Online research

▶ Key methodological concerns – validity and ethical protections

▶ Data collected in online spaces and/or using online tools and methods must be valid, or becomes an ethical issue > social benefit

▶ Chief is identity deception

  ▶ Creation of virtual persona and alter ego is key to functioning in some online spaces; potential obstacle in such environments (Gaiser & Schreiner, 2009)

▶ "… probably highly inflated in public perception" and "despite the fact that one *can* misrepresent his or herself online, it is useful to ask why someone *would* misrepresent his or herself online, in order to consider how widespread the phenomenon is and whether it would take place in research settings" (Walther, 2002, p. 211).

  ▶ Verification of identity is problematic for offline data collection tools too - phone polls, anonymous surveys.

▶ Ethical protections for offline research must extend to online contexts.

Offline texts – dilemmas, dilemmas

- ▶ Publicly available texts
- ▶ Unknown/uncontactable authors
- ▶ Possibly juvenile
- ▶ Of questionable legality
- ▶ http://thetoiletstudy.com/info/

Rethinking around points of tension:
    human subjects, notions of 'personhood', privacy

Ethical principles:
    consideration of people involved, harm minimization, rights
    of subjects, proportional obligation of the researcher to
    protect 'human element'

THE TWO THINGS I HATE ARE GRAFFITI & IRONY

Human subjects

- Determination of 'human subjects' and requirement of ethical review has origins in bio-medical research, broadly maintained as yardstick for social research (Markham & Buchanan, 2012).

- At UNE "It is the responsibility of the researchers to ensure that all facets of human research meet the requirements of the National Statement on Ethical Conduct in Human Research 2007 (updated May 2015)"

  - Resides at National Health & Medical Research Council (NHMRC)

- The notion of what constitutes a 'human subject' has long been a matter for debate

- Online, underpinning research with concerns for **harm**, **vulnerability** and **privacy** may better serve to protect individuals than definitions of humans as subjects (Markham & Buchanan, 2012).
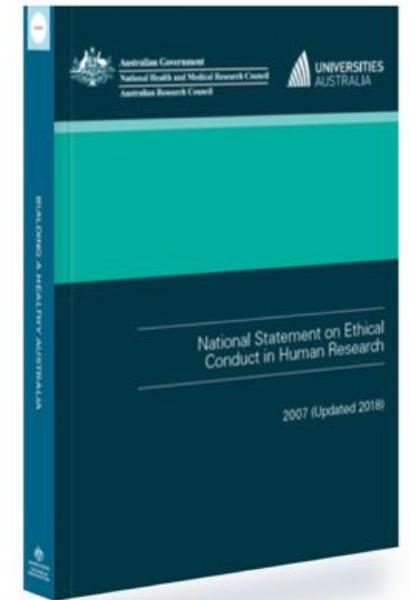
Human subjects

▶ Variety in interactions between researchers, texts and humans

▶ Range of ethical approaches is evident across a field. e.g., text production - blogs

▶ linguistics-orientation

  ▶ rhetorical genre (e.g., Miller & Shepherd, 2004; Herring, 2010), discourse (e.g., Myers, 2010) and media studies (e.g., Davis, 2011) traditions

  ▶ texts as the object of study and use content analytical approaches

  ▶ blogs are analysed as texts, and researchers do not interact with blog authors

  ▶ descriptions of data collection and analysis and other aspects of methodology do not refer to human subjects

Human subjects

- Ethnography e.g., Fowley (2011)
  - Interacts directly with blog authors and analyses creative works
  - Object of study is people
  - Follows human ethics protocols
- Efimova (2009)
  - Distinguishes between blogs as artefacts and authors as human participants
  - Blogs – publicly authored and available texts, attributes cited materials to the bloggers who created it (not anonymized)
    - Honours cultural norms of blogging community (public intellectuals)
  - Interactions, interviews with blog authors – follow ethics protocols

Human subjects – protect 'human element'

Fowley

- Blogs authored in a password-protected sharing space

- Authors are adolescents

    - care "... is even more necessary when the bloggers are young people, and could be seen as vulnerable..." (Fowley, 2011, p. 79).

Efimova

- Blogs are publicly available (so can/should be shared)

- Authors are adults

- Attend to the privacy status ascribed to online texts by human authors – access, dissemination
- Consider the status of the 'human element' - increased obligations to protect more vulnerable authors (Markham & Buchanan, 2012)

Human subjects – protect 'human element'

- Examples of protecting human element
    - Align stance with the ethical expectations of the research community in which practice is based.
        - Object of study – texts
    - Use publicly available texts
    - Honour the cultural norms of creative community
        - Share content, share source – increase traffic (desirable expected outcome for authors)
    - Take conservative approach to **protect more vulnerable authors** e.g., children

- While a study of text might free the researcher of certain ethical considerations and processes (Serfaty, 2004), it does not exempt scholars from acting in an ethical manner.

Notions of personhood

- Concerned with representation of individuals in datasets and the ease with which individuals may be identified (Markham & Buchanan, 2012).

- Connection between data and the person
  - Clear when collected from individuals – follow ethics protocols
  - Less clear in huge datasets e.g., tweets (Zappavigna, 2012)
  - Less clear for human-like entities e.g., avatars or bots

- Easy to "… forget that there was ever a person somewhere in the process" (Markham & Buchanan, 2012, p. 7).

- Robustness of anonymization to protect privacy is becoming less reliable, although still legally & ethically 'watertight' (Narayanan & Shmatikov, 2009).

Notions of personhood

- Issue of personhood is one that scholars are yet to resolve, although many working on the issue (e.g., Narayanan & Shmatikov, 2009).

- Assume that texts of all kinds are indeed written by human beings (at some level), and adhere to the fundamental ethical principle of **harm minimization** (Markham & Buchanan, 2012).

  - Apply to large- and small-scale data pools, public and private

- Example: anonymise aggregated data, **protect privacy**:

"While weblog text is public and the blogging patterns could be easily discovered from it, aggregating and visualising those patterns adds an additional layer of information and it is not necessarily in the interests of the participant to share it publicly" (Efimova, 2009, p. 37).

Definitions of privacy

- Privacy is a highly malleable concept; two key dimensions for consideration

    - Notions change depending on the privacy parameters of an online space, and the cultural norms and shared understandings negotiated within and between groups of users (Markham & Buchanan, 2012).

    - Viewpoints vary on that which constitutes 'private' and inappropriate to share – subject matter and identity

Definitions of privacy – privacy parameters

- Web 2.0 social realms instate various privacy parameters
  - Entirely public
  - Offer degrees of privacy
  - Self-nominate levels for differing content
- 'Perceived privacy' (Markham & Buchanan, 2012) - expectation
  - LiveJournal – password protected, users share with specific groups within and may choose to make content publicly available (Fowley, 2011).
  - Publicly available, search engine indexed blogs (Efimova, 2009).

Definitions of privacy – privacy parameters

▶ Authors of online texts can change the privacy status of outputs over time

▶ Authors decision to retract a once-public blog complicates perceived privacy

    ▶ Impacts on data analysis, publication, dissemination

▶ Early scholars sought to overcome fickle habit of disappearance by capture and storage (e.g., McMillan, 2000), and practice persists in web content analysis (e.g., Herring, 2010)

- Attend to the privacy status ascribed to online texts by human authors – analysis, publication, dissemination
- Need to respect expressions of privacy remains, even if and when expressions alter.

Definitions of privacy – viewpoints on privacy

- HUGE variation - appropriate to share vs private or taboo
  - We have always shared personal info (Zimmer & Hoffman, 2012)
  - Increasing trend to divulge once-private info (Miller & Shepherd, 2004)
  - Contrasting standards of appropriateness are evident; 'oversharers' (Zappavigna, 2012), attention seekers (Marwick, 2013)
  - Some users adopt pseudonyms (Fowley, 2011), some use real and full names and addresses etc (Serfaty, 2004)
- What should be kept private by children is more clearly defined e.g.,
  - Full name, address, phone numbers, DOB, school, email, username, bank details (Office of the Children's eSafety Commissioner, n.d.)

Definitions of privacy – viewpoints on privacy

▶ Only use blogs that are publicly available

▶ Remove blogs from dataset as authors remove them from public view

▶ Remove identifying information from blogs to maintain privacy

▶ Maintain the privacy of child authors, especially when they exercise poor judgement

   ▶ Some child authors include both first and surnames in blog URLs. Withhold these URLs from publication

- Increased obligations to protect more vulnerable authors (Markham & Buchanan, 2012) – analysis, publication, dissemination

Concluding remarks

"The internet is a social phenomenon, a tool, and also a (field) site for research. Depending on the role the internet plays in the research project or how it is conceptualized by the researcher, different epistemological, logistical and ethical considerations will come into play."

"At its most fundamental level, we recognize that ethical decision-making interweaves one's fundamental world view (ontology, epistemology, values, etc), one's academic and political environment (purposes), one's defining disciplinary assumptions, and one's methodological stances.

Decision making occurs at many junctures in the cycle of inquiry, including research design, research conduct, and research production and dissemination.

Because of the complexity of ethical decision making .., [we focus] on general principles".

(Markham & Buchanan, 2012, p. 3)

References

Efimova , L. (2009). Passion at work: blogging practices of knowledge workers. (PhD), Utrecht University, Enschede, Netherlands.

Fowley, C. (2011). Publishing the confidential: an ethnographic study of young Irish bloggers. (PhD), Dublin City University, Dublin.

Herring, S. C. (2010). Web content analysis: expanding the paradigm. In J. Husinger, L. Klastrup, & M. Allen (Eds.), International Handbook of Internet Research (pp. 233-249). Netherlands: Springer.

Gaiser, T. J., & Schreiner, A. E. (2009). *A Guide to Conducting Online Research.* **London: SAGE Publications.**

LiveJournal. (1999). About LiveJournal [webpage]. Retrieved from http://www.livejournal.com/about

McMillan, S. J. (2000). The microscope and the moving target: the challenges of applying content analysis to the World Wide Web. Journalism and Mass Communication Quarterly, 77(1), 80-98. Retrieved from http://web.utk.edu/~sjmcmill/Research/research.html

Miller, C. R., & Shepherd, D. (2004). Blogging as social action: a genre analysis of the weblog. In A. Gurak, S. Antonijevic, L. Johnson, C. Ratliff, & J. Reyman (Eds.), Into the Blogosphere: Rhetoric, Community, and Culture of Weblogs: University of Minnesota.

Markham, A., & Buchanan, E. (2012). *Ethical decision-making and Internet research 2.0: Recommendations from the AoIR ethics working committee.* **Retrieved from** http://www.aoir.org/reports/ethics2.pdf

Marwick, A. E. (2013). Status update: celebrity, publicity, and branding in the social media age. New Haven, CT: Yale University Press.

Narayanan, A., & Shmatikov, V. (2009). De-anonymizing social networks. IEEE Symposium on Security & Privacy. Oakland, CA. Retrieved from http://www.cs.utexas.edu/~shmat/shmat_oak09.pdf

Office of the Children's eSafety Commissioner. (n.d.). eSafety - Enhancing online safety [website]. Retrieved from https://www.esafety.gov.au/

Senft, T. M. (2013). Microcelebrity and the branded self. In J. Hartley, J. E. Burgess, & A. Bruns (Eds.), A Companion to New Media Dynamics (pp. 346-354). Somerset, NJ: Wiley-Blackwell.

Serfaty, V. (2004). The Mirror and the Veil: An Overview of American Online Diaries and Blogs. Amsterdam, Netherlands: Rodopi.

Walther, J. B. (2002). Research ethics in Internet-enabled research: Human subjects issues and methodological myopia. *Ethics and Information Technology, 4(3), 205-216.*

Zappavigna, M. (2012). The Discourse of Twitter and Social Media: How We Use Language to Create Affiliation on the Web. London: Continuum

Zimmer, M., & Hoffman, A. (2012). Privacy, context, and oversharing: reputational challenges in a Web 2.0 world. In H. Masum, M. Tovey, & C. Newmark (Eds.), The Reputation Society: How Online Opinions are Reshaping the Offline World (pp. 175-184). Cambridge, Massachusetts: MIT Press.