Some Challenges for Ethics in Social Network Research

Luke Hutton
School of Computer Science
University of St Andrews
St Andrews, Fife, UK
lh49@st-andrews.ac.uk

Tristan Henderson
School of Computer Science
University of St Andrews
St Andrews, Fife, UK
tnhh@st-andrews.ac.uk

ABSTRACT

Social network sites (SNSes) comprise one of the most popular networked applications of late, with hundreds of millions of users. Collecting and analysing data from such systems creates myriad ethical issues and challenges for researchers both in networked systems and other fields, as highlighted by recent media sensitivity about research studies that have used data from Facebook. In our workshop contribution we discuss recent work that we have been carrying out in the area of responsible SNS research, revolving around themes of reproducibility, consent, incentives, and creating ethical workflows.

Keywords

social networks sites; research ethics; ethics; methodology; reproducible research

1. INTRODUCTION

One of the most popular networked systems of late is the social network site (SNS) [3]. With millions of people using Facebook, Twitter, Weibo, Instagram and the like on a daily basis, these systems have provided an important source of data for researchers in computer science, social science, psychology and more [4]. With sensitive and private data being shared, sometimes inadvertently, it is paramount that researchers treat such studies with care. Two recent studies using Facebook data, one studying emotion [10], and another studying political opinion [2], led to many news stories (e.g., [5]) raising concerns that data were collected without appropriate consent, and these are not the first to attract such attention [19].

We have been conducting some work in the area of responsible SNS research that we believe would be of relevance to this workshop, and we would welcome the opportunity to discuss the various challenges in this forum.

2. CHALLENGES

2.1 Reproducibility

An important part of the nascent interest in ethical and responsible research is the dissemination of best practices. Events such as this one, and similar events at CSCW, ICWSM and so forth do this. But another way to disseminate information about what researchers do is to improve the documentation and sharing of information (data collection practices, data analysis practices, and even better the data themselves) that relate to a particular research artefact. There has been a huge push in this area recently, with many researchers and new initiatives urging research to be made more reproducible.

In a recent study [8] we looked at 505 papers that used SNS data published between 2011 and 2013, to determine the state of the art in reproducible SNS research. Each paper was checked against ten criteria, looking at the extent of documentation and sharing of the paper's code, methods, and data. Unfortunately we found that only one paper met all ten criteria. Particularly poorly documented was indeed the ethics considerations for each study; this included a lack of information about whether ethics committee or Institutional Review Board (IRB) approval had been obtained, how participants were informed about the procedures of a study or how consent was obtained, and how data were sanitised or otherwise protected. The situation is not all bad; in particular papers in venues such as SOUPS (Symposium On Usable Privacy and Security), which requires an appendix outlining ethical considerations, and WPES (Workshop on Privacy in the Electronic Society), which allows but does not require such appendices, did better than others. But clearly this is an area that might benefit from improvements. One issue in making such information easily accessible to readers and reviewers is the variation between ethics procedures. We looked at this as part of a PhD summer school last year [1], where ten ethics forms from universities in the UK, US, EU and Asia were studied. We found that the ten forms featured 145 different attributes, only two of which were common to all ten forms (the name of the principal investigator, and whether informed consent was sought), making it difficult to generalise or compare these procedures.

We are keen to discuss whether our requirements for encoding the reproducibility of SNS studies are shared by the community. Is meeting our ten metrics sufficient to reproduce any SNS study, or are there types of experiments which use SNS data which introduce new reproducibility challenges?

2.2 Consent

One of the common elements in the aforementioned ethics forms was informed consent, and indeed this is something that is considered the cornerstone of ethical research. Many SNS studies do not ask for consent, with some arguing that it is unnecessary [18]. But for those studies that do obtain consent, one common way is what has been termed "secured consent" [11], where a checkbox at the beginning of an experiment indicates that a participant is willing to share all their SNS information with a researcher. We explored whether this assumption holds true in a study that compared what Facebook information participants would be willing to share when they were told that they were participating in a health study or in a computer science study; our hypothesis was that participants would opt to share information that was more relevant to the purposes of the study [13]. We found that secured consent was inappropriate in that there were various types of information (in particular photos and photo albums) that participants were unwilling to share with researchers. This has implications for studies which conduct wholesale collection of data from people's SNS accounts.

If secured consent is insufficient, then one alternative might be to ask participants before each piece of data is collected by a researcher. But this can be burdensome in longitudinal or large-scale experiments, where participants may not want to be bothered by requests for information, but at the same time might be uncomfortable with divulging different types of information to researchers. We have been exploring the use of Nissenbaum's model of contextual integrity [16] to reduce the burden of asking for consent in SNS studies, by only asking for consent when we predict that informational norms have been violated [7].

We would like to discuss the challenges other researchers encounter when acquiring consent for such studies. Can we determine best practices for using language which meaningfully conveys to participants how potentially sensitive SNS data are used, and consider whether other means of capturing consent may be more appropriate for such studies?

2.3 Incentives

Consent can also be affected by power relationships between researcher and participant, in particular through the use of incentives. We have explored the use of incentives in location-based advertising [9], but would be interested in discussing how to best use or design incentives for participation in SNS studies, and the ethical challenges associated with introducing incentives.

2.4 Workflow

Various researchers have proposed guidelines for SNS research [12, 17], while others believe that ethics is contextual or situational [15, 14] as concerns may change over the course of study. We believe that regardless of whether fixed or flexible guidelines are in place, it would be use of to researchers if they were able to exchange information about best practices and exactly what had been done for particular experiments. We have built a workflow system for our own SNS experiments [6] that allows the creation of experimental policies, enforces the policies when data are collected or used, and lets researchers exchange these policies to disseminate best practices.

We have developed these tools for our own purposes, but we would like to discuss whether the language we have developed for capturing the workflow of an experiment is sufficiently expressive. We are interested in learning more about the studies others are running, and encourage others to try out our system, helping us steer the development of additional functionality to support studies we had not anticipated in its original design.

3. CONCLUSION

To conclude, our work aims to address some of the ethical challenges in SNS research, specifically, reproducibility, consent, incentives and workflow. We invite feedback about the appropriateness of our approaches, and would welcome the opportunity to contribute to discussion of the myriad other challenges to improve the state of the art in ethics and responsible networked systems research.

4. ACKNOWLEDGEMENTS

5. REFERENCES

[1] S. Arabas, M. R. Bareford, I. P. Gent, B. M. Gorman, M. Hajiarabderkani, T. Henderson, L. Hutton, A. Konovalov, L. Kotthoff, C. McCreesh, R. R. Paul, K. E. J. Petrie, A. Razaq, and D. Reijsbergen. Case studies and challenges in reproducibility in the computational sciences, 11 Sept. 2014. Online at http://arxiv.org/abs/1408.2123.

[2] E. Bakshy, S. Messing, and L. Adamic. Exposure to ideologically diverse news and opinion on Facebook. *Science*, pages aaa1160+, 07 May 2015. doi:10.1126/science.aaa1160.

[3] d. boyd and N. B. Ellison. Social network sites: Definition, history, and scholarship. *Journal of Computer-Mediated Communication*, 13(1):210–230, Oct. 2007. doi:10.1111/j.1083-6101.2007.00393.x.

[4] R. Caers, T. De Feyter, M. De Couck, T. Stough, C. Vigna, and C. Du Bois. Facebook: A literature review. *New Media & Society*, 15(6):982–1002, Sept. 2013. doi:10.1177/1461444813488061.

[5] K. Hill. Facebook Added 'Research' To User Agreement 4 Months After Emotion Manipulation Study. *Forbes*, 30 June 2014. http://onforb.es/15DKfGt.

[6] L. Hutton and T. Henderson. An architecture for ethical and privacy-sensitive social network experiments. *ACM SIGMETRICS Performance Evaluation Review*, 40(4):90–95, Apr. 2013. doi:10.1145/2479942.2479954.

[7] L. Hutton and T. Henderson. "I didn't sign up for this!": Informed consent in social network research. In *Proceedings of the 9th International AAAI Conference on Web and Social Media (ICWSM)*, pages 178–187, Oxford, UK, May 2015. Online at http://www.aaai.org/ocs/index.php/ICWSM/ICWSM15/paper/view/10493.

[8] L. Hutton and T. Henderson. Towards reproducibility in online social network research. *IEEE Transactions on Emerging Topics in Computing*, 2015. To appear.

[9] L. Hutton, T. Henderson, and A. Kapadia. "Here I am, now pay me!": Privacy concerns in incentivised location-sharing systems. In *Proceedings of the 2014 ACM Conference on Security and Privacy in Wireless & Mobile Networks (WiSec)*, pages 81–86, Oxford, UK, July 2014. doi:10.1145/2627393.2627416.

[10] A. D. I. Kramer, J. E. Guillory, and J. T. Hancock. Experimental evidence of massive-scale emotional contagion through social networks. *Proceedings of the National Academy of Sciences*, 111(24):8788–8790, 17 June 2014. doi:10.1073/pnas.1320040111.

[11] E. Luger. Consent reconsidered; reframing consent for ubiquitous computing systems. In *Proceedings of the 2012 ACM Conference on Ubiquitous Computing*, pages 564–567, Pittsburgh, PA, USA, 2012. doi:10.1145/2370216.2370310.

[12] R. McKee. Ethical issues in using social media for health and health care research. *Health Policy*, Mar. 2013. doi:10.1016/j.healthpol.2013.02.006.

[13] S. McNeilly, L. Hutton, and T. Henderson. Understanding ethical concerns in social media privacy studies. In *Proceedings of the ACM CSCW Workshop on Measuring Networked Social Privacy: Qualitative & Quantitative Approaches*, San Antonio, TX, USA, Feb. 2013. Online at http://tristan.host.cs.st-andrews.ac.uk/pubs/mnsp2013.pdf.

[14] C. Munteanu, H. Molyneaux, W. Moncur, M. Romero, S. O'Donnell, and J. Vines. Situational ethics: Re-thinking approaches to formal ethics requirements for human-computer interaction. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, Seoul, South Korea, 2015. doi:10.1145/2702123.2702481.

[15] F. Neuhaus and T. Webmoor. Agile ethics for massified research and visualization. *Information, Communication & Society*, 15(1):43–65, Feb. 2012. doi:10.1080/1369118x.2011.616519.

[16] H. Nissenbaum. *Privacy in Context: Technology, Policy, and the Integrity of Social Life*. Stanford Law Books, Stanford, CA, USA, 2009.

[17] C. M. Rivers and B. L. Lewis. Ethical research standards in a world of big data. *F1000Research*, 6 Feb. 2014. doi:10.12688/f1000research.3-38.v1.

[18] L. Solberg. Data mining on Facebook: A free space for researchers or an IRB nightmare? *University of Illinois Journal of Law, Technology & Policy*, 2010(2), 2010. Online at http://www.jltp.uiuc.edu/works/Solberg.htm.

[19] M. Zimmer. "But the data is already public": on the ethics of research in Facebook. *Ethics and Information Technology*, 12(4):313–325, Dec. 2010.