

Insights from a Workshop on Social Media Analysis and Mental Health: Putting People at the Centre of Human Data

Elvira Perez Vallejos¹, Ansgar Koene¹, Chris James Carter¹, Ramona Statache¹, Svenja Adolphs¹, Claire O'Malley¹, Tom Rodden¹, Derek McAuley¹, Jens Binder², Karen Douglas³, Ilka Gleibs⁴, Monica Whitty⁵, and Victoria Betton⁶

HORIZON Digital Economy Research, University of Nottingham¹

Nottingham Trent University²

University of Kent³

London School of Economics⁴

University of Leicester⁵

Leeds & York Partnership NHS Foundation Trust & Leeds Community Healthcare NHS Trust⁶

elvira.perez@nottingham.ac.uk

ansgar.koene@nottingham.ac.uk

christopher.carter@nottingham.ac.uk

ramona.statache@nottingham.ac.uk

svenja.adolphs@nottingham.ac.uk

claire.omalley@nottingham.ac.uk

tom.rodde@nottingham.ac.uk

derek.mcauley@nottingham.ac.uk

jens.binder@ntu.ac.uk

K.Douglas@kent.ac.uk

i.h.gleibs@lse.ac.uk

mw229@leicester.ac.uk

victoria.betton@nhs.net

Abstract: In collaboration with the Institute of Mental Health and the Centre for Advanced Studies at the University of Nottingham, CaSMa (Citizen-Centric Approaches to Social Media Analysis) held a thought-provoking workshop to reconsider and reflect upon concerns about social media data, especially in relation to vulnerable adults and minors who may inadvertently be part of the user group. Held over a single day, the workshop included subject experts such as Monica Whitty (University of Leicester), Karen Douglas (University of Kent), Jens Binder (Nottingham Trent University) and Ilka Gleibs (London School of Economics). The speakers engaged with the audience to illustrate a series of relevant ethical issues (e.g., anonymity and privacy) and their implications not only on Internet-mediated research aspects but for day-to-day Internet-related activities. During the workshop, two hands-on sessions focused on the ethics of personal data in both theory and practice, and how social media data can be accessed and analysed to: (1) understand the ways people use social media and what this means for individuals and society, (2) understand social phenomena and events expressed in social media by drawing upon social media as a critical, and timely, source of information, (3) develop facilities and approaches that are sensitive to the personal nature of human data: Citizen-centric approaches, (4) promote responsible innovation in the capture, analysis and use of human data.

The ultimate goal of the workshop was to further understand the issues that social scientists encounter when accessing and handling personal data, and possible ethical solutions in developing sustainable and reliable 'Citizen Social Science'. This paper presents the most relevant insights as formulated from the workshop discussions on social media data and mental health. These insights were provided by clinicians, researchers, and policy makers interested in understanding the ethical constraints inherent to digital human data generated through the use of online platforms by vulnerable users (e.g., children, young people and adults

experiencing mental distress). These groups were of particular interest given that they may often not be fully aware of the terms and conditions accepted when accessing such online services.

Keywords: Social Media, Mental Health, Privacy, Informed consent, Anonymity, Ethics, Internet Mediated Research

Introduction

The risks and challenges of personal data are numerous and often underestimated, ranging from a lack of transparency about the quantity and quality of personal information currently collected and analysed by large institutions, potential imbalances in the trust relationship between service users who generate the data and those who collect and analyse it (e.g., making new uses of already collected data), and insufficient privacy safeguards in the handling of personal data. The Citizen-centric Approaches to Social Media Analysis (CaSMa) research team at the Horizon Digital Economy Institute, University of Nottingham, aims to promote social media literacy by informing the general public, in addition to academics, clinicians and policy makers, about ethically sound ways in which to collect, access, and analyse digital human data. One of CaSMa's main objectives is to engage with wide audiences to explore the ethical challenges that are associated with unprecedented access to human data for the purpose of social science analysis.

In January 2015, CaSMa conducted a workshop with the Institute of Mental Health and the Centre for Advance Studies at the University of Nottingham, gathering experts on cyber psychology and social sciences to present on a shared interest in social media research ethics, and to discuss the challenges of using social media data for research with a citizen-centric approach. In the current paper, we summarise the main points to have emerged from the key notes delivered on the day of the workshop, followed by a discussion of some of the most consistent issues raised during two hands-on, parallel sessions titled 'Social Media Access and Analysis: Citizen-centric Approaches' and 'Ethics of Personal Data: Consent, Anonymity Issues, Privacy, and Trust'.

Keynote Speakers

Opening the workshop was Dr Jens Binder, Assistant Professor of Psychology at Nottingham Trent University, who effectively illustrated a series of relevant aspects to take into consideration when conducting research that includes social media data. These points included the intercultural and global aspects of social media research, including cultural and institutional differences in securing consent, creating / implementing research regulations and confronting issues of digital literacy and awareness; the growing responsibility to offset the unreliability of anonymisation techniques by finding new ways to minimise risk to user confidentiality; and the breach of trust and loss of reputation caused by the increasing identification risk when combining and aggregating large datasets.

The second speaker, Karen Douglas, Professor of Social Psychology, at the University of Kent, further developed the topic of anonymity by exploring the he pros and cons inherent to anonymity and de-individuation (e.g., national security vs. cyber hate) and their relation to social identity theories (e.g., purpose of anonymous behaviour by Spears and Lea, 1994). Topics of discussion focused on the uses and misuses of social media and how anonymity can enable multiple ways of harassment and online hostility to groups or individuals; and the sociological aspects of social media and how online communication and social networking can affect individuals, groups and society (e.g., friendships, activism and social change).

Ilka Gleibs, Assistant Professor in Social and Organisational Psychology, at the London School of Economics, dealt with the lack of clear ethical guidelines for social science research, especially for large-scale online

studies on social network sites. Dr Gleibs, focused on the ethical question of getting informed consent when data is collected from social network sites (SNS). She argued that data from SNS are not necessarily perceived as public by the user and research based on these data should not be exempt from the ethical standard that informed consent must be obtained from participants. Based on the concept of privacy in context (Nissenbaum, 2010), she further proposed that the norms of distribution and appropriateness are violated when researchers manipulate online contexts and collect data without consent. Finally, suggestions were made for existing and possible future practices for large-scale online studies such as ensuring that the methods and processes used remain rooted in long-lasting ethical practices.

The reality of obtaining fully informed consent from research participants and the dangers of identifying anonymised individuals were also exposed by Monica Witty, Professor at the Department of Media and Communication, at The University of Leicester, together with the complexity and inaccessibility of terms and conditions and the need to be more aware of all the additional information researchers can access when connecting (i.e., aggregating) single data points. With a focus on cybersecurity, the superidentity model (i.e., understanding identity when merging psychological, biological, biographical, and cyber world data), and detection of corporate insider threats, her engaging talk invited the audience to reflect on the consequences of large organisations sharing personal data to third parties to modify people's behaviour (e.g., increase sales).

Hands-on Sessions

During the hands-on session titled 'Social Media Access and Analysis: Citizen-centric Approaches', discussions primarily focused on data ownership, terms and conditions and consent. The public debate in the Gmail case about account access in case of user death (James and Magee 2013) was discussed to illustrate that data might be personal to the user, but the right to access and use it is given to the platform.

A second issue was the complexity of terms and conditions and the no-opt-out which was perceived as a form of coercion. The majority of terms and conditions for online services include a section indicating that the data can be used for research; most attendees agreed that this implies that the users are aware that research is being done and consent to it. However, most users are likely to implicitly assume that the research mentioned in terms and conditions refers primarily to research aimed at improving the product and conducted by or in the interest of the platform owners. The user can also be said to presuppose that 'research' means research for economic benefit by the platform and, therefore, it should transparent when research refers to academic or highly personal issues.

A third topic for debate focused on participant consent. It was debated that when the type of research justifies it, it is possible to ask for participant consent after the research has been performed. The use of post-hoc consent raises many problems. It exposes the researcher and his/her institution to possible legal action for damages, invasion of privacy, etc. It raises legal and ethical questions about data ownership: the participant owns their personal data and can request to withdraw it from the study; but the study has already been performed, and the resulting data may actually belong to the funding body that paid for the researcher's time, expertise and equipment. Attendees concluded that consent should not be treated as a dichotomy but a more fluid, dynamic concept. We interpret this idea about withdrawal of consent and actual meaning of consent as a departure from the classic dichotomised view of it. The myth of informed consent was clearly presented and it was acknowledged by most of the attendees that even the people asking for consent sometimes do not fully understand the implications of what they propose. Consent fatigue (i.e., a problem that may appear when the subject is constantly asked to review consent or when the consent is broken down into minute subsections in the interest of full coverage/ disclosure) was viewed as an undesirable consequence that must be taken into consideration. A final theme that emerged from these discussions was the increasing trend among employers to contractually regulate the private, off-work social media activity of employees under the banner of ensuring professional conduct. The ethics around this kind of practise were questioned as they could breach privacy rights.

The points raised at the hands-on sessions titled 'Ethics of Personal Data: Consent, Anonymity Issues, Privacy, and Trust' included the need to raise confidentiality standards when participants with mental health issues could be – intentionally or unintentionally - part of the data sample. Participants agreed that online personal data should be treated differently than offline data to ensure anonymity and privacy as the risks were more difficult to predict and therefore control. Attendees agreed that privacy should be assessed under a certain temporal, spatial, application/usage, functionality/identity and social contexts (i.e., contextual privacy). The urgency to consider legal and ethical guidelines was an aspect to be considered taking into account the recent mood manipulation experiments by Facebook (Kramer et al 2014).

An interesting distinction was made between 'being public' and 'being in the public', and a similar distinction was made between 'public data' and 'open data' as not all public data is actually open data (i.e., data that is made available by organisations, businesses and individuals for anyone to access, use and share) and consequently explicit consent should be obtained before the 'public-not-open' data is used. Concerns were also raised about the dangers and potential harm that could be triggered in vulnerable users when social media research involved manipulation of the user experience (e.g., Kramer et al 2014). A final intervention brought awareness on how to communicate data to avoid misunderstanding and misuse.

Conclusions

Insight from this workshop will guide the development of methodologies for internet mediated research and improving the guidelines for research ethics related to social media data. The recurrent concerns on privacy, informed consent, anonymity, and global/cultural differences that have emerged from this workshop urge the need for citizen centric approaches sensitive to the personal nature of human data and designed for the promotion of responsible innovation in the capture, analysis and use of human data.

References

Kramer, A., Guillory, J., Tancock, J. (2014) "Experimental evidence of massive-scale emotional contagion through social networks", *PNAS*, Vol. 111, pp 8788-8790.

Nissenbaum, H. (2010) *Privacy in context. Technology, policy, and the integrity of social life*. Stanford, CA: Stanford University Press.

Spears R., and Lea, M. (1994) "Panacea or panopticon? The hidden power in computer-mediated communication" *Communication Research*, Vol. 21, pp 427–459.

James P., and Magee S. (2013) "[Google RIP: What Inactive Account Manager means for your will](http://www.telegraph.co.uk/technology/google/9997105/Google-RIP-What-Inactive-Account-Manager-means-for-your-will.html)", [online]. *The Telegraph*. 16 April 2013. <http://www.telegraph.co.uk/technology/google/9997105/Google-RIP-What-Inactive-Account-Manager-means-for-your-will.html>