THE ETHICS OF ONLINE RESEARCH

ADVANCES IN RESEARCH ETHICS AND INTEGRITY

Series Editor: Dr. Ron Iphofen FAcSS, Independent Consultant, France

Recent Volumes:

Volume 1: FINDING COMMON GROUND: Consensus in Research Ethics Across the Social Sciences. *Edited by Ron Iphofen*

ADVANCES IN RESEARCH ETHICS AND INTEGRITY

Series Editor

Dr. Ron Iphofen *FAcSS, Independent Consultant, France*

Editorial Advisory Group

Professor Robert Dingwall *FAcSS, Dingwall Enterprises Ltd and Nottingham Trent University, UK*

Dr. Nathan Emmerich *Queens University Belfast, UK*

Professor Mark Israel *University of Western Australia, Australia*

Dr. Janet Lewis AcSS, *Former Research Director, Joseph Rowntree Foundation, UK*

Professor John Oates *FAcSS, Open University, UK*

Associate Professor Martin Tolich *University of Otago, New Zealand*

This page intentionally left blank

ADVANCES IN RESEARCH ETHICS AND INTEGRITY VOLUME 2

THE ETHICS OF ONLINE RESEARCH

VOLUME EDITOR

KANDY WOODFIELD

Samaritans, UK

United Kingdom – North America – Japan
India – Malaysia – China

Emerald Publishing Limited
Howard House, Wagon Lane, Bingley BD16 1WA, UK

First edition 2018

Copyright © 2018 Emerald Publishing Limited

Reprints and permissions service
Contact: booksandseries@emeraldinsight.com

No part of this book may be reproduced, stored in a retrieval system, transmitted in any form or by any means electronic, mechanical, photocopying, recording or otherwise without either the prior written permission of the publisher or a licence permitting restricted copying issued in the UK by The Copyright Licensing Agency and in the USA by The Copyright Clearance Center. No responsibility is accepted for the accuracy of information contained in the text, illustrations or advertisements. The opinions expressed in these chapters are not necessarily those of the Editor or the publisher.

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library

ISBN: 978-1-78714-486-6 (Print)
ISBN: 978-1-78714-485-9 (online)
ISBN: 978-1-78743-446-2 (Epub)

ISSN: 2398-6018 (Series)

ISOQAR certified
Management System,
awarded to Emerald
for adherence to
Environmental
standard
ISO 14001:2004.

Certificate Number 1985
ISO 14001

INVESTOR IN PEOPLE

This volume is dedicated to the members of the #NSMNSS community across the world who continue to push the boundaries of social science and internet mediated research by sharing their experiences and challenges as technology continues to transform how we live, and understand, our everyday lives.

This page intentionally left blank

CONTENTS

DEDICATION *vii*

ABOUT THE AUTHORS *xi*

ABOUT THE SERIES EDITOR *xvii*

ABOUT THE VOLUME EDITOR *xix*

INTRODUCTION TO VOLUME 2: THE ETHICS OF ONLINE RESEARCH *1*
Kandy Woodfield and Ron Iphofen

1. THE ETHICAL DISRUPTIONS OF SOCIAL MEDIA DATA: TALES FROM THE FIELD *13*
Susan Halford

2. USERS' VIEWS OF ETHICS IN SOCIAL MEDIA RESEARCH: INFORMED CONSENT, ANONYMITY, AND HARM *27*
Matthew L. Williams, Pete Burnap, Luke Sloan, Curtis Jessop and Hayley Lepps

3. THE CHANGING ROLES OF RESEARCHERS AND PARTICIPANTS IN DIGITAL AND SOCIAL MEDIA RESEARCH: ETHICS CHALLENGES AND FORWARD DIRECTIONS *53*
Sarah Quinton and Nina Reynolds

4. USING TWITTER AS A DATA SOURCE: AN OVERVIEW OF ETHICAL, LEGAL, AND METHODOLOGICAL CHALLENGES *79*
Wasim Ahmed, Peter A. Bath and Gianluca Demartini

5. GETTING TO YES: INFORMED CONSENT IN QUALITATIVE SOCIAL MEDIA RESEARCH *109*
Janet Salmons

6. THE TROUBLE WITH TINDER: THE ETHICAL COMPLEXITIES OF RESEARCHING LOCATION-AWARE SOCIAL DISCOVERY APPS *135*
Jenna Condie, Garth Lean and Brittany Wilcockson

7. ETHICAL CHALLENGES OF PUBLISHING AND SHARING SOCIAL MEDIA RESEARCH DATA *159*
Libby Bishop and Daniel Gray

8. THE ETHICS OF USING SOCIAL MEDIA DATA IN RESEARCH: A NEW FRAMEWORK *189*
Leanne Townsend and Claire Wallace

9. WHERE NEXT FOR #SOCIALETHICS? *209*
Steven Ginnis

10. CONCLUSION: GUIDING THE ETHICS OF ONLINE SOCIAL MEDIA RESEARCH – ADAPTATION OR RENOVATION? *237*
Ron Iphofen

INDEX *243*

ABOUT THE AUTHORS

Wasim Ahmed is a Doctoral Researcher at the Information School, University of Sheffield. His PhD looks at understanding pandemics and epidemics on social media platforms. He has many years of experience in social media research working on a number of commercial and industry projects across academia, government, industry for companies such as Manchester United. He has published peer reviewed articles on social media, and has delivered a number of prestigious talks across the world. He runs his own analytics blog, and regularly posts on social media research for the LSE Impact blog with a readership in at least 136 countries. Wasim has an interdisciplinary background with an MSc in Information Systems, and a BA in Philosophy.

Peter Bath is Professor of Health Informatics, Head of the Information School, and Head of the Health Informatics Research Group at the University of Sheffield. As Chair of the University Research Ethics Committee (UREC), Peter was responsible for the review of University of Sheffield's research ethics policy, which included new policy on the ethical issues relating to social media research. Peter is the Principal Investigator of the 'A Shared Space and a Space for Sharing' project, a transdisciplinary project examining how people in extreme circumstances share information online. He is supervising a number of students who are utilising social media data as part of their PhDs. Peter has published over 200 papers/abstracts, including 116 articles in peer-reviewed journals.

Libby Bishop is currently working 80% at GESIS-Leibniz Institute for the Social Sciences in Cologne (Germany) where she is the Leader of the Data Security and Data Linkage Team in the Data Archive. The team is enhancing capacities for the Secure Data Center, including remote access to sensitive data and curation of new forms of data such as social media. With her 20% at the UK Data Service at the University of Essex (UK), Libby is a Manager in the Research Data Management section. She has expertise in the ethics of re-using data and informed consent. She also develops and delivers training on legal and ethical issues in sharing big data and on secondary analysis of qualitative data. During 2017, she is leading the Legal/Ethics strand of a CESSDA (Consortium of European Social Science Data Archives) Research Data Management project.

Pete Burnap is Reader in the School of Computer Science and Informatics, Cardiff University. He is an applied computer scientist with a principal focus on data and computational methods such as machine learning to model Web-enabled human and software behaviour – contributing to the academic fields of Social Computing, Web Science and Cybersecurity. His research outcomes, which include more than 60 academic articles – stemming from funded research projects worth over £8million, are organised and disseminated via the Social Data Science Lab, within which he is a director and the computational lead.

Jenna Condie is a Lecturer in Digital Research and Online Social Analysis at Western Sydney University, and a World Social Science Fellow on transformations to sustainable urbanisation for the International Social Science Council. Her research focuses upon how everyday experiences and social connections are mediated by digital technologies in the places we live, work, play and stay. As an experienced qualitative researcher, Jenna use methods that are participatory, critical, creative, and digital. Current projects include examining the use of social media by non-profit organisations, particularly within the social housing sector; how location-aware technologies are reworking our social relations; and how digital technologies can be harnessed for urban governance, particularly within urban planning and community safety initiatives.

Dr. Gianluca Demartini is a Senior Lecturer in Data Science at the University of Queensland. His main research interests are Information Retrieval, Semantic Web, and Human Computation. His research has been supported by the UK Engineering and Physical Sciences Research Council (EPSRC) and by the EU H2020 framework program. He received the Best Paper Award at the European Conference on Information Retrieval (ECIR) in 2016 and the Best Demo Award at the International Semantic Web Conference (ISWC) in 2011. He has published more than 70 peer-reviewed scientific publications including papers at major venues such as WWW, ACM SIGIR, VLDBJ, ISWC, and ACM CHI. He has given several invited talks, tutorials, and keynotes at a number of academic conferences (e.g., ISWC, ICWSM, WebScience, and the RuSSIR Summer School), companies (e.g., Facebook), and Dagstuhl seminars. He is an ACM Distinguished Speaker since 2015.

Steven Ginnis is a Research Director at Ipsos MORI, leading both Education, Children and Families research, and innovation in research methods across the UK business. With nine years of experience across central and local government research, Steven is responsible for pioneering the use of new and emerging technologies in public sector research, including mobile research, text analytics, biometrics, neuroscience, and social media research.

He directed the Wisdom of the Crowd project, working with Demos, University of Sussex and Centre for Analysis of Social Media to help embed rigour and ethics at the heart of the social media research process. More widely, he recently led the Dialogue on Data Science Ethics commissioned by the Government Data Science partnership looking at public attitudes towards government use of data science.

Daniel Gray is a PhD student in the Economic and Social Research Council Wales Doctoral Training Partnership in the School of Social Sciences at Cardiff University. His research investigates misogynistic hate speech found on Twitter and other social media, employing big data collection and critical discourse analysis. He has worked with the Social Data Science Lab at Cardiff, using its guidelines for social media research for his projects, including his M.A dissertation on misogynist speech, which is included as a case study in this volume.

Susan Halford is a Professor of Sociology, Director of the Web Science Institute at the University of Southampton and Deputy Director of the EPSRC Centre for Doctoral Training in Web Science Innovation. She was a founding Director of the University's Work Futures Research Centre and has over 25 years' experience of research on organizational change, including digital innovation, employee driven innovation and the voluntary workforce. .

Curtis Jessop is a Research Director in the Longitudinal Surveys team at NatCen Social Research where he manages the NatCen Panel, an open mixed-mode random-probability research panel, and works on large-scale longitudinal surveys such as Understanding Society and the British Birth Cohort Studies. Curtis is the lead for the 'New Social Media, New Social Science' collaborative network which facilitates discussion and knowledge sharing between researchers using social media in their research, including a particular focus on research ethics. He has a particular interest in the ethics and practicalities of linking social media and social survey data to enhance the research potential of both.

Garth Lean is a Lecturer in Geography and Urban Studies in the School of Social Sciences and Psychology at Western Sydney University. His research and writing primarily investigate experiences of travel in a modern, mobile world. He is the lead researcher of the Transformative Travel Research Project (www.transformativetravel.com) and co-lead of the TinDA (Travel in the Digital Age) Project (www.tindaproject.com). He has published a variety of papers on travel, tourism and mobilities, along with the monograph *Transformative Travel in a Mobile World* (CABI Books, 2016), and the edited volumes *Travel and Representation* (Berghahn, 2017), *Travel and Imagination* (Ashgate, 2014) and *Travel and Transformation* (Ashgate, 2014). He is Vice

President of the Geographical Society of New South Wales and a member of the Geographies of Leisure and Tourism Research Group with the Royal Geographical Society.

At the time of contributing to this chapter: Hayley Lepps was a Researcher in the Health Policy team at NatCen Social Research, having previously worked in the Questionnaire Development and Testing Hub at NatCen. Hayley is a mixed-methods researcher with experience in conducting research through depth interviews, focus groups and observations; developing and testing survey questions through cognitive interviews; and managing web and telephone surveys. Hayley is a member of the #NSMNSS (New Social Media, New Social Science) network and is particularly interested in how ethical guidelines for social research apply to research using social media data, following her work on the NatCen project entitled 'Research using Social Media; Users' Views'. Through this work, Hayley was invited to contribute towards 'Social Media Research: A Guide to Ethics', a document intended to provide clear guidelines on the ethical use of social media data in research.

Sarah Quinton is currently the Chair of Research Ethics at Oxford Brookes University where she teaches and researches in the Business School. Her particular area of research interest is how the digital environment is shaping behaviour change across society, business, consumers and citizens. Her recent publications include the *Journal of Industrial Marketing Management*, the *International Journal of Management Reviews*, and the *Journal of Marketing Management*. She believes that digitalisation is impacting on the tools and perspectives within social science research methods and that the new complex research environment that we inhabit requires further consideration.

Nina Reynolds is currently the Professor of Marketing at the University of Wollongong where she teaches and researches in the School of Management, Operations and Marketing. She has a longstanding interest in research methods, originally focusing on design and analysis issues related to international survey research, but now focusing on how technological changes impact on how researchers gain insights into individuals' behaviours. Her other research currently focuses on how we, as consumers, manage and use our personal resources in today's consumption environment, and on the influence of consumption experiences on our wellbeing. Her work has been published in a number of journals including the *Journal of International Business Studies*, the *Journal of Service Research*, the *European Journal of Marketing*, the *Journal of Business Research* and *International Marketing Review*.

Janet Salmons, PhD, is an independent scholar through Vision2Lead. She has authored numerous books and articles about online research approaches

and related ethical issues. Her most recent book is *Doing Qualitative Research Online*. Dr. Salmons is based in Boulder, Colorado.

Luke Sloan is Deputy Director of the Social Data Science Lab (http://socialdatalab.net/) and Reader at the School of Social Sciences, Cardiff University. His research is concerned with increasing the utility of Twitter data for the social science community, through the development and evaluation of demographic proxies, to establish who is and who is not represented. He also directs the Cardiff the Q-Step Centre.

Leanne Townsend is a senior researcher based at the University of Aberdeen. Her work explores how technologies can transform rural communities and economies. Her interests include the ethics of social media research; the rural creative economy; rural food producers; marginalised groups; crafting communities; rural economic in-migration and the transformational potential of social media for communities of practice and place. She has led various projects and brought together new networks. She is on the board of directors at the Scottish Sculpture Workshop and on the editorial board at Local Economy. She works collaboratively to co-produce research which seeks positive change and new academic understandings.

Claire Wallace is Professor of Sociology at the University of Aberdeen. She was Leader of the Enterprise and Culture theme of the RCUK dot.rural Digital Economy Hub at the University of Aberdeen. She also worked on the ESRC project on Social Media at the same University on which this report is based. Claire Wallace has published widely on digital communications, quality of life, community and other topics. Her most recent book is *The Decent Society* together with Pamela Abbott and Roger Sapsford and published by Routledge in 2015.

Matthew Williams is Professor of Criminology at the School of Social Sciences, Cardiff University. He was a lead researcher on the Cardiff Online Social Media Observatory (COSMOS) programme (2012–15), and now continues this work as the Director of the Social Data Science Lab (http://socialdatalab.net/), part of the ESRC Big Data Network. He has published extensively on the use of social media data in crime and security research. He was appointed to the ESRC's Big Data Network Phase 3 working group and is Principal Investigator on the ESRC's New and Emerging Forms of Data Policy Demonstrator Grant Centre for Cyberhate Research and Policy.

Brittany Wilcockson is a research assistant and PhD candidate (psychology) within the School of Social Sciences and Psychology at Western Sydney University, Australia. Her research is situated within social and political psychology and explores the cognitive underpinnings of language about 'outsider' groups (groups which are positioned by themselves or

government as outside the law, e.g., outlaw motor cycle gangs). Her work also explores the intersection of online research methods and ethics. She is more broadly interested in the use of language data in forensic and investigative settings, and the security and privacy implications of online information storage and use.

ABOUT THE SERIES EDITOR

Dr. Ron Iphofen FAcSS is Executive Editor of the Emerald book series *Advances in Research Ethics and Integrity* and edited Volume 1 in the Series: *Finding Common Ground: Consensus in Research Ethics Across the Social Sciences* (2017). He is an Independent Research Consultant, a Fellow of the UK Academy of Social Sciences, the Higher Education Academy and the Royal Society of Medicine. Since retiring as Director of Postgraduate Studies in the School of Healthcare Sciences, Bangor University, his major activity has been as an adviser to the European Commission (EC) on both the seventh framework programme (FP7) and Horizon 2020. His consultancy work has covered a range of research agencies (in government and independent) across Europe. He was Vice Chair of the UK Social Research Association and now convenes their Research Ethics Forum. He was scientific consultant on the EC RESPECT project (establishing pan-European standards in the social sciences) and chaired the ethics and societal impact advisory board for SECUR-ED (a European Demonstration Project on passenger transport security). He has advised the UK Research Integrity Office; the National Disability Authority (NDA) of the Irish Ministry of Justice; and the UK Parliamentary Office of Science and Technology among many others. Ron was founding Executive Editor of the Emerald gerontology journal *Quality in Ageing and Older Adults*. He published *Ethical Decision Making in Social Research: A Practical Guide*, with Palgrave Macmillan (2009/2011) and coedited with Martin Tolich *The SAGE Handbook of Qualitative Research Ethics* (to appear in 2018).

This page intentionally left blank

ABOUT THE VOLUME EDITOR

Kandy Woodfield is a social researcher and learning & development professional. She is Head of Learning and Development at Samaritans and her previous roles include Head of Social Sciences at the Higher Education Academy and Director of Learning & Enterprise at NatCen Social Research. Kandy is a Chartered Fellow of the Chartered Institute of Personnel and Development (CIPD), a Fellow of the Learning & Performance Institue (LPI) and has worked across the academic, public and third sectors. Kandy spent over twenty years in social policy research working across a range of topics including welfare reform and crime & justice before moving into learning & development. She is the founder of the New Social Media New Social Science? (#NSMNSS) community of practice and most recently edited a crowd sourced book of blogs on 'Social Media in Social Research'.*

* Woodfield, K. Ed. 2014 *Social Media in Social Research: Blogs on Blurring the Boundaries* (NatCen Social Research; London)

INTRODUCTION TO VOLUME 2: THE ETHICS OF ONLINE RESEARCH

Kandy Woodfield and Ron Iphofen

HOPES AND PROMISES

The Internet, the World Wide Web and social media – indeed all forms of online communications – have been seen as attractive fields of research since their inception for many reasons. Some of the earliest discussion and commentary were eager in terms of the opportunities for methodological initiatives and innovations in research and the 'attractiveness' of easily accessed, massive amounts of primary and secondary data sources.

Using e-mail as a research tool was seen to potentially offer researchers advantages such as easy access to world-wide samples, low administration costs (both financially and temporally), ready-transcribed data and its unobtrusiveness, and 'friendliness' to respondents was also valued (Robson & Selwyn, 1998). Once you became an experienced Web 'surfer,' you could have access to a wealth of valued and authoritative information sources (Peters, 1998). Instant messaging was seen as a cost- and time-effective method for in-depth interviewing (Fontes & O'Mahoney, 2008); online participants might

The Ethics of Online Research
Advances in Research Ethics and Integrity, Volume 2, 1–12
Copyright © 2018 by Emerald Publishing Limited
All rights of reproduction in any form reserved
ISSN: 2398-6018/doi:10.1108/S2398-601820180000002013

be better able to 'tell their story' in a way that suits them and so may even be ethically sounder than conventional methods of narrative data collection. Traditional survey methods have long suffered from increasing costs and declining responses rates – Web surveys offered an attractive alternative and have the advantage of ease in collecting more sensitive data: people have been shown to respond to Web surveys on sex and health more readily and openly than in face-to-face interviews (Burkill et al., 2016). Glaser, Dixit, and Green (2002, pp. 177–193, 189–190) argue the case for better access to 'hard to reach' groups: 'the anonymity of the internet permits research into marginal groups for whom self-disclosure may have costs, and where participants may be suspicious of researchers and outsiders.' And Illingworth (2001) suggests that the Internet affords an efficient way of recruiting specialist participants.

Now we have moved on considerably from the 'simpler' form of the Internet and the World Wide Web. The term 'social media' refers to websites, online platforms, or applications that allow for one-to-one, one-to-many, or many-to-many synchronous or asynchronous interactions/dialogue between users who can create, archive, and retrieve user-generated content. In social media, the user is producer; communication is interactive and networked with fluid roles between those who generate and receive content (see Bechmann & Lomborg, 2013; Salmons, 2014). Online social media cross cultures, communities, populations, and continents. They have the methodological potential of access to large sample sizes and diverse populations for multidisciplinary, multimethod, and multinational purposes. The last decade has seen an explosion in Internet-based social science research and in research which uses social media to attract participants, generate data, disseminate findings, and engage in dialogue with audiences for that research. As indicated above online research has been conducted since the early days of the Internet itself, but the ubiquity of online social platforms in the last decade, such as Twitter or Instagram, has placed the means to deliver social media and internet based research into the hands of all researchers not just those with sophisticated technical skills. It is no longer necessary to be a technical specialist to gather data from online platforms and the computing power required to undertake Internet searches or draw down social media data is now available to the many not the few.

Just as the democratization of the Internet has meant millions of people worldwide now see online interaction as a normal part of their everyday lives, it is increasingly acceptable that social scientists consider the role that Internet data could or should play in their research studies. Of course, to talk of the democratization of the Internet is to underplay the very real power issues that are evident in the use of Internet platforms. Who owns and

controls data and who has or does not have access to participate in online discourse and communities is highly contested. As we have seen in Volume 1 in this book series, ethical debates have remained a core element of the social science discourse: what is ethical or 'right' in how we gather and use data from participants remain a continuing thread of lively discussion; the growth in the use of the Internet to recruit participants, gather data, and analyze online experiences, behaviors, and viewpoints is not immune to the same ethical decisions or quandaries. But the most contentious and ongoing debate is about whether 'conventional' research methods and ethical codes/guidelines apply to online research equally well or whether new methods and new codes need to be sought (Eynon, Schroeder, & Fry., 2009). Indeed, some researchers have even suggested online research as a way of evading the bureaucratic 'obstructions' of the formal ethics review process.

CHALLENGES

Most of the ethical issues have considerable overlaps – perhaps even more so than conventional research. It is hard to separate the seeking of informed consent from confidentiality and anonymity, and all have implications for data ownership and the publication of findings. It may be fair to suggest that all of the research ethics issues cohere around the issue of the public and/or private nature of this modern and developing 'research site' and how effective is an 'expectation of privacy' in an essentially public environment?

The concern may be most readily illustrated in research into 'virtual environments' and 'online worlds' that are used for gaming and/or socializing. People experience these spaces as 'other than the ones they are physically in' (Eynon, 2009, p. 189). Within these spaces, researchers can be engaging in experiments in interpersonal behavior, gathering large amounts of data and/or participant observation. The kinds of ethical issues to be addressed here include how much virtual danger should participants be subjected to? How close to the participants 'real' physical attributes can avatars be permitted to be? Thus, for example, Slater et al. (2006) conducted a *virtual* replication of Stanley Milgram's notorious obedience to authority experiments. Questions arise then about whether it is ethical even to 'abuse' avatars and, though the authors don't mention this, what negative effects such an experiment might still have on the participants – which was the main concern of the after effects of the Milgram studies.

The availability of large amounts of data leads to the temptation to gather them – for social scientific purposes. Such opportunities have been rarely

available in the past without a great deal of effort. The question here is 'even if one's behavior is public and can be observed and captured, is it nevertheless always allowable for researchers to use data from these worlds?' (Eynon, 2009, p. 192). Issues involve the possibility of gaining consent, of gathering data without consent/permissions, and the possibilities of linking data to identifiable individuals thereby failing to preserve their anonymity whether they sought it or not. The same position applies to participant observation – just because a researcher engages 'as if' they were a genuine participant, the other interactants may not be aware of that, and may assume the researcher to be a genuine participant unless they are formally informed otherwise. Necessarily, while the latter option is more ethical, the methodological consequence is that participants' behavior might not be quite so authentic.

It might be less challenging simply to avail oneself of the mass of self-evidently publicly available data. But is it ethical to 'harvest' public twitter accounts without users' consent? Looking at the ongoing online debate opinions vary: some say consent is not necessary since the tweets are public, a conscious choice made by the user to allow their activity to be seen by whoever chooses to see it. If you don't restrict access to your account, there cannot be an expectation of privacy. Perhaps one hopes that one's contribution is less public since it might be 'obscured' by the millions of other tweets – 'hidden in plain sight' – but that's an unrealistic, even naive, assumption that, as a consequence, it is *slightly* private! It is public by definition. Some argue that this is not a formal consenting as to how the contributions may be used – collected, stored and analyzed. Obviously, the same could be said of any public text or statement.

The UK research carried out by NatCen (National Centre for Social Research; Beninger, Fry, Jago, Lepps, Nass, & Silvester, 2014) suggests that a first principle of online usage of any kind is that the user must bear the responsibility for choosing where to post and how privately to post. Of course, they can't do that alone since the site owner shares a mutual responsibility – site owners must make clear just how public the data are and who is able to access them. Given all of that then there are still times when informed consent should be sought. There seems to be a great deal of consensus around the moral requirement to seek consent for data access and analysis even if it is not legally required. In such a way, at least some measure of trust is promoted between researchers and online participants.

Such trust is all too easily breached if the implications of a promise of, say, confidentiality are not fully considered – by both researcher and researched. Eynon et al. (2009), for example, advocate an enhanced sensitivity to context on the grounds that the distance between researcher and research is greater

than in conventional research and people have a range of different expectations about the nature of online interactions, and one cannot always know directly what those understandings are. Moreover, once you describe a computer-mediated community it is relatively easy to find out which one it is and who is on it. Complete anonymity is impossible since it is almost impossible to remove the trace of computer-transmitted messages. The 'personal' is not necessarily 'private' (Robson & Robson, 1999; Robson & Selwyn, 1998; Zimmer, 2010).

MEETING THE CHALLENGES

During 2017, there was widespread discussion about social media and Internet-mediated research, this was not the case 5 or 10 years ago. These debates have become part of mainstream methodological debate, and researchers are developing new tools and approaches for exploring social media data and understanding the social media dimension of contemporary life. It is hard to find any sector of life where the promise and potential of 'big data' have not been touted as the next 'big thing.' However, this new frontier of social science research has posed an increasingly knotty set of challenges for researchers.

We have faced, and continue to face, methodological challenges around the quality, quantity, and representativeness of social media datasets. Conducting research in this new domain has forced us to consider questions of representation, power, and authenticity in our research. In fact, most social media data is *not quantitative data, rather qualitative data on a quantitative scale* (Francesco D'Orazio on https://www.pulsarplatform.com), and we have yet to fully address the high proportion of social media traffic that consists of pictures, not text. The social science of images and visual data are not well served by current social science approaches and tools that focus on text and numerical data. We continue to have much to learn from colleagues in the digital humanities, computer science, and artificial intelligence (AI) disciplines.

There is also a collaborative challenge. As a new, developing field, it is clear that the most powerful insights from social media and Internet-mediated research will come from transdisciplinary efforts drawing on the varied insights and skills of, for example, statisticians, qualitative researchers, digital curators, information scientists, machine learning experts, and human geographers. We have a window of opportunity to forge a new shape and rhythm for our research methods and epistemologies. We are progressing but not yet fulfilling the potential transformative nature of this moment.

In a world where technology moves fast, we are also faced with a capability challenge. Are we conversant with the social worlds we research, are we sufficiently confident and competent to understand the complex algorithms and processes that increasingly define our interactions? Which brings us to the connective or contextual challenge how can we research what we do not understand or use? Discussions at #NSMNSS network meetings, online discussions, and conferences have revealed that many research methods lecturers, research supervisors, research commissioners, and research ethics review committee members do not feel adequately equipped to make rounded, informed decisions about the quality, ethics or value of social media research projects and proposals.

We also face a synthesis challenge, how if at all can new forms of research and findings map onto, elaborate or further inform conventional social research data?

This volume focuses on perhaps the knottiest of the challenges facing researchers – the ethical dimensions of Internet-mediated and social media research. It largely concentrates on the use of social media data and platforms for social science research although chapters also look more broadly at the ethical challenges of 'big data.'

BACKGROUND TO THIS VOLUME

The origin of this volume was the formation of an online community of practice called 'New social media, New Social Science?' (or #NSMNSS for short) in 2011. #NSMNSS was established with a 12-month grant from the National Centre for Research Methods in the UK (NCRM) to provide a year of facilitated dialogue and knowledge exchange for researchers using social media and Internet forums or platforms in their social science research projects. It aimed to provide a safe, collaborative place to explore the challenges and issues that researchers were facing in trying to explore these new online spaces for research purposes. Membership was not limited to those using certain methods or approaches and we actively sought to encourage participation from all sectors in the research community including those working in the academic sphere, not-for-profit sector, in government and public services, and those working in market research. The only limits placed on membership were that participants needed to be investigating social science questions and making use of the Internet and social media in their work. Throughout the life of the network, the dominant issue has persistently been an uncertainty over the ethical boundaries of Internet mediated research, a sense of a lack of guidance over what is and is not acceptable ethically and a growing sense

of the need to debate these issues. At the time the network began, there was limited guidance for researchers in existing ethical codes of practice relating to Internet-mediated research.

We might feel that our social science research is a benign endeavor contrasted to the commercial *harvesting* of customer insight data, but we all face similar ethical and legal challenges, as we always have done: Whose data? Whose consent? Whose ownership? All are complex issues about how researchers collect and use the data of people using social media and other Internet platforms. We have only just begun to scrape the surface of these discussions, and meanwhile, data is being mined, harvested, analyzed, and reported in increasing volume. The critical moments that will shape and define the ethical and legal frameworks for the use of social media data may not come from social science research but from the use of social media data in the commercial world or media realm, these industries practices may shape our future access to research data. Are we engaging enough with these sectors and issues? Have the voices of researchers in the social sciences been loud enough in ongoing legislative debates around the use of personal data?

These challenges are hard things to tackle, but they also give us great opportunities to push the boundaries of our practice as social scientists. Social media research needs social science as much as it does data science, it needs anthropology and ethnography as well as big data analytics, it needs to reflect, explore, and understand the context and communities that anchor and shape social media data.

This volume presents nine chapters each of which addresses the ethical challenges of particular research issues, social media platforms, or approaches. The authors describe the research they have conducted and provide an insight into the ethical state of play for Internet-mediated research.

Many interesting questions are raised including just how different should ethical guidance be for internet-mediated research? Are the issues new or is it merely a matter of change in context – in the research site? Familiar issues are addressed such as informed consent, anonymity, confidentiality, privacy, protection from harm, social benefits; but there are no easy answers. Do we really need a new ethical code of research practice, or do we just need to be agile in applying our existing social science ethics to this new area of operations?

OUTLINING THE VOLUME

Susan Halford's opening chapter addresses the inadequacy of the current ethics review process at meeting many of the challenges referred to above.

Methodological innovation is a necessary feature of responding to the rapid development of digital technology and online communications. The process of research ethics review could obstruct progress in this field since 'it' does not know how to handle the 'disruptions' caused by researchers keen to enter these new and attractive research sites. Halford's account draws on her experiences of teaching doctoral students how to engage in Web science and their collective attempts to seek ethical approval for innovative online research. The five 'disruptive' concerns that Halford addresses are: the data are in essence necessarily secondary – existing already and not 'created' by the research engagement; that means the data are beyond the control of the researcher – already public and so not so easily 'protected' by the researcher; datasets are potentially infinite, not 'bounded' for targeted care and attention, and only limited by the time the researcher has available to collect and study them; the scale of the data alters the nature of relationships between researcher and researched; and, the attractiveness of access to such data blurs boundaries between professional research domains, almost 'requiring' interdisciplinary cooperation in order to more fully understand what their analysis entails. Halford suggests further consideration of the 'situational' ethics approach that has grown in recent years and which is returned to in subsequent chapters in this Volume. Unless both ethics codes and guidelines are updated and ethics review committee topic 'checklists' are amended to take account of such disruptions, valuable innovative research opportunities into rich and meaningful modern data sources will be lost. Worse, unless responsible social science researchers manage the means of online data access and use there are others with different political and commercial aims who could contaminate the field for all future research attempts.

In the second chapter, Matthew Williams and colleagues directly address users' expectations about how their online activities should be exploited for research purposes. Whether realistic or not, there is an expectation of privacy, anonymity, and confidentiality. Such expectations place perhaps an even greater ethical burden on online researchers to respect participants' wishes and a sustained ethical reflexivity as a consequence of the 'blurred' relationship between researcher and researched in online environments. Williams and colleagues rehearse the legal, professional, and moral guidelines that researchers should follow in ensuring research integrity and ethical practice and show the relevance of these principles in a study of their own. Readers will find many useful, practical suggestions for handling online research ethics in the lessons learned from this particular engagement.

Sarah Quinton and Nina Reynolds focus more particularly on the changed relationship between researcher and researched in online environments.

Qualitative researchers have long held concerns for the nature of this relationship – how it is developed and managed. Issues of sensitivity and vulnerability are linked to refining questions of consenting and anonymity, and all are connected to research impact. There has been an assumption that the necessary physical distance between the 'actors' in digital research engagements removes the more delicate aspects of their relationship. This is far from the truth and, if anything, the relationship is more complex and requires more understanding of consequences than has previously been assumed – or, rather, presumed by those not regularly engaged in online research.

Wasim Ahmed and colleagues form a research postgraduate supervisory team and so have a particular interest in assuring their Ph.D. students get the ethics of online research right. Their interest is even further specific in targeting social media use, in this case Twitter, during infectious disease outbreaks. It may be particularly important to understand how people make use of social media during times of crisis and emergency – so in the focused study what they report can help understand how people manage their online social networking during other times of crisis and incident. Any researcher wishing to target Twitter as a research site will do well to start here and learn the legal and ethical issues this team has benefitted from understanding and is able to share in preparation for a research project and ethics review. The case studies drawn upon are especially illustrative of the issues that arise and the care that needs to be taken to both review and conduct studies of this nature.

Janet Salmons is well known for her writing and training for some years in this field so her contribution here will be particularly instructive for newcomers to online social media research. She points to the nuanced understandings necessary for gaining consent in and from online communities in terms of the meaning of being 'fully informed' and having the 'capacity to consent.' The detailed coverage of processes and procedures offered here through cases and exemplars almost amounts to a technical manual for conducting qualitative online social research. Finally, Salmons offers some practical templates that will assist researchers presenting their protocol for ethics approval.

In their chapter, Jenna Condie and colleagues raise another set of issues not fully understood by research ethics committees – location-aware social discovery applications for smartphones. The temptations offered for data mining and data extraction are hard to resist for the curious and diligent researcher. Their 'application of interest' is Tinder, a social relationship-building app that uses geolocation to enable people to make connections wherever they may be – geographically. What is particularly insightful from this study is how researchers-as-users drew on their own personal experiences of using Tinder to reflect upon its implications for social research in the move

from a 'social space' to a 'research context.' These insights are enhanced in the ways in which the authors show how conventional formal research ethics approval applications cannot easily make sense of the required transformations of consenting, confidentiality, secondary data access, and so on. Once again, there is much to be done in ensuring the ethics approval process keeps pace with such rapidly developing social media technologies.

Libby Bishop and Daniel Gray address issues that some commentators regard as more related to research integrity than ethics per se – dissemination and data sharing. Clearly, there are overlaps between ethics and integrity. An inevitable consequence of the seeking of open access publishing is a push toward open data access – the opportunity to share data in order to conduct further analyses. Data archiving has long addressed such concerns for secondary access to data and has mainly relied on gaining permissions for durable use. Once again, online social networking research raises additional difficulties connected with how the data were originally derived and the public/private domain debate. While the sharing of data is often seen as 'good in itself' and related to principles of public benefit, transparency, and equitable access to knowledge, there remain concerns with online data that, due to the sophistication of digital technologies, it may be much harder to prevent the revelation of personal data and its associated sensitivities via shared data. With case studies of Twitter and Facebook, Bishop and Gray illustrate directly how these issues play out in specific contexts. They show how the conventional ethical principles can be more severely challenged by the extended linkages made possible with digital technologies. Importantly, they argue that the responsibility for ethical behavior cannot only reside with the researcher 'on the ground' – or rather, in cyberspace – in addition, the institutional structures that both permit and enhance data sharing opportunities must bear some responsibility for how shared data can be used and abused.

The chapter from Leanne Townsend and Claire Wallace offers a practical ethical framework as guidance for online researchers. The framework was developed in collaboration with key experts who have been working in the online research environment in the UK and was 'tested' in application to fictionalized case studies that represent exactly the kinds of contexts in which such research is likely to be conducted. This chapter serves as practical guidance and the cases studies reflect exactly the kinds of ethical dilemmas researchers will face and which they must make judgments on. Solutions to the dilemmas presented by the case studies are offered by Townsend and Wallace and their collaborators, and these, once again, illustrate the enduring problem of research ethics and integrity – solutions may be refined by other commentators and, even, other solutions might be suggested. What is vital

here is how the proposed framework is applied and the research engagement thought through in an ethically robust manner.

The final contributed chapter in this collection is from Steven Ginnis and, as in the previous chapter, draws on collaborative work that seeks to offer practical guidance to best practice in ethical online social media research. In essence, this chapter reports what can best be described as a 'standard-setting' project – ensuring that those who are conducting online social media research have the technical skills and ethical awareness to do it well. When new research opportunities arise, keen researchers are eager to innovate, access the new data, and make contributions while the field is still fresh and fertile. The problem is that such eagerness may not be equally matched by the cautions required to ensure that both data producers and data users are protected from both anticipated and unpredictable consequences. The research drew on a large population sample to assess users' views of the reasons for research on social media, its value, and how it is conducted. A particular valuable insight is taken from including the views of younger social media participants. As a result, the series of recommendations for best practice offered here can be seen as assuredly resting upon public perceptions of and wishes for how research is conducted.

In fact, we are confident that this Volume in the Advances in Research Ethics and Integrity Series will make a much-needed contribution to the quality of online social media research and help ensure the public trust in how researchers are likely to engage them in the future. The contributors are all expert in their field and have conducted research and training with new entrants to online research. Most importantly, in addition to covering the range of ethical issues that need to be borne in mind, they have offered practical suggestions about how to address them. These can only count as recommendations since it is certain that this fertile field is likely to continue to change just as rapidly in the future as it has up to now. And just as the technology evolves, so too will users' use of it – and researchers' eagerness to understand them.

REFERENCES

Bechmann, A., & Lomborg, S. (2013). Mapping actor roles in social media: Different perspectives on value creation in theories of user participation. *New Media & Society*, *15*(5), 765–781. doi:10.1177/1461444812462853

Beninger, K., Fry, A., Jago, N., Lepps, H., Nass, L., & Silvester, H. (2014). *Research using social media: Users' views.* London: National Centre for Social Research.

Burkill, S., Copas, A., Couper, M. P., Clifton, S., Prah, P., Datta, J., …, Erens, B. (2016). Using the web to collect data on sensitive behaviours: A study looking at mode effects on the British National Survey of Sexual Attitudes and Lifestyles. *PLoS ONE*, *11*(2), e0147983. doi:10.1371/journal.pone.0147983

Eynon, R., Schroeder, R., & Fry, J. (2009). New techniques in online research: Challenges for research ethics *Twenty-First Century Society: Journal of the Academy of Social Sciences*, *4*(2), 187–199. Methodological innovation and developing understandings of 21st century society. Published online: 15 Jun 2009.

Fontes, T. O., & O'Mahony, M. (2008). In-depth interviewing by instant messaging. *Social Research Update*, Volume 53. Guildford, UK: University of Surrey.

Glaser, J., Dixit, J., & Green, D.P. (2002). Studying hate crime with the internet: What makes racists advocate racial violence? *Journal of Social Issues*, *58*, 1.

Grinyer, A. (2007). The ethics of Internet usage in health and personal narratives research. *Social Research Update*, Volume 49. Guildford, UK: University of Surrey.

Illingworth, N. (2001). The internet matters: Exploring the use of the internet as a research tool *Sociological Research Online*, *6*(2). Retrieved from http://www.socresonline.org.uk/6/2/illingworth.html

Peters, S. (1998). Finding information on the World Wide Web. *Social Research Update*, Volume 20. Guildford, UK: University of Surrey.

Robson, K., & Robson, M. (1999). Your place or mine? Ethics, the researcher and the internet. In T. Welland & L. Pugsley (Eds.), *Ethical dilemmas in qualitative research* (pp. 94–107). Aldershot: Ashgate.

Robson, K., & Selwyn, N. (1998). Using e-mail as a research tool. *Social Research Update*, Volume 21. Guildford, UK: University of Surrey.

Salmons, J. (2014). *New social media, new social science … and new ethical issues!* Retrieved from https://drive.google.com/file/d/0B1-gmLw9jo6fLTQ5X0oyeE1aRjQ/edit

Slater, M., Antley, A., Davison, A., Swapp, D., Guger, C., Barker, C., Pistrang, N., and Sanchez-Vives, M.V. (2006). A virtual reprise of the Stanley Milgram obedience experiments. *PLoS ONE*, *1*(1): e39. Retrieved from https://doi.org/10.1371/journal.pone.0000039

Zimmer, M. (2010). 'But the data is already public': On the ethics of research in Facebook. *Ethics and Information Technology*, *12*(4), 313–325.