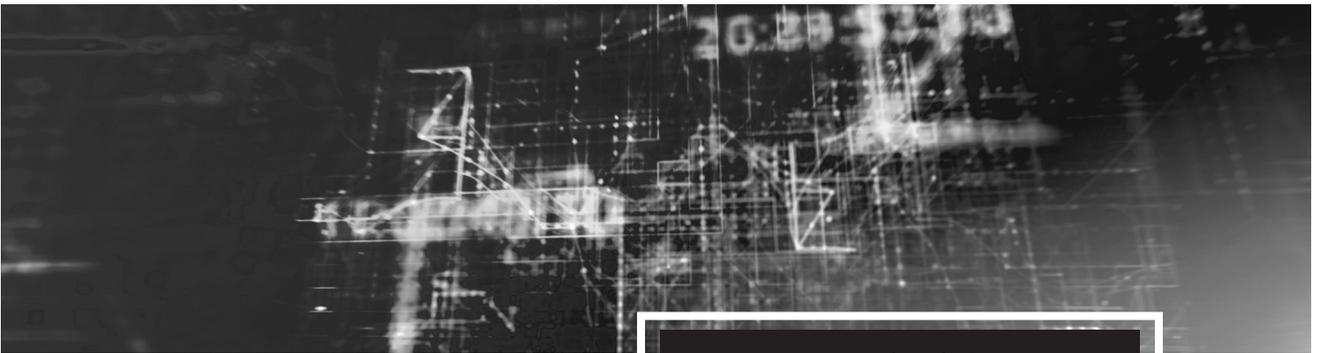


# The SAGE Handbook of Online Research Methods



Second Edition

Edited by  
Nigel G. Fielding,  
Raymond M. Lee  
and Grant Blank

 SAGE reference

Los Angeles | London | New Delhi | Singapore | Washington DC | Melbourne



Los Angeles | London | New Delhi  
Singapore | Washington DC | Melbourne

SAGE Publications Ltd  
1 Oliver's Yard  
55 City Road  
London EC1Y 1SP

SAGE Publications Inc.  
2455 Teller Road  
Thousand Oaks, California 91320

SAGE Publications India Pvt Ltd  
B1/11 Mohan Cooperative Industrial Area  
Mathura Road  
New Delhi 110044

SAGE Publications Asia-Pacific Pte Ltd  
3 Church Street  
#10-04 Samsung Hub  
Singapore 049483

---

Editor: Judi Burger  
Editorial assistant: Matthew Oldfield  
Production editor: Sushant Nailwal  
Copyeditor: Sunrise Setting Limited  
Proofreader: Sue Cleave  
Indexer: Editors  
Marketing manager: Sally Ransom  
Cover design: Wendy Scott  
Typeset by Cenveo Publisher Services  
Printed in the UK

At SAGE we take sustainability seriously. Most of our products are printed in the UK using FSC papers and boards. When we print overseas we ensure sustainable papers are used as measured by the PREPS grading system. We undertake an annual audit to monitor our sustainability.

Introduction & editorial arrangement © Nigel G. Fielding,  
Raymond M. Lee and Grant Blank, 2017

- Chapter 1 © Raymond M. Lee, Nigel G. Fielding and Grant Black, 2017  
Chapter 2 © Rebecca Eynon, Jenny Fry and Ralph Schroeder, 2017  
Chapter 3 © Karsten Boye Rasmussen, 2017  
Chapter 4 © Claire Hewson, 2017  
Chapter 5 © Dietmar Janetzko, 2017  
Chapter 6 © Ayelet Baram-Tsabari, Elad Segev and Aviv J. Sharon, 2017  
Chapter 7 © Martin Innes, Colin Roberts, Alun Preece and David Rogers, 2017  
Chapter 8 © Jonathan Bright, 2017  
Chapter 9 © Vasja Vehovar and Katja Lozar Manfreda, 2017  
Chapter 10 © Ronald D. Fricker Jr, 2017  
Chapter 11 © Vera Toepoel, 2017  
Chapter 12 © Lars Kaczmirek, 2017  
Chapter 13 © Don A. Dillman, Feng Hao and Morgan M. Millar, 2017  
Chapter 14 © Bernie Hogan, 2017  
Chapter 15 © Javier Borge-Holthoefer and Sandra González-Bailón, 2017  
Chapter 16 © Corinna Eisenbroich, 2017  
Chapter 17 © Harko Verhagen, Magnus Johansson and Wander Jager, 2017  
Chapter 18 © Helen Kennedy and William Allen, 2017  
Chapter 19 © Roel Popping, 2017  
Chapter 20 © Mike Thelwall, 2017  
Chapter 21 © Edward Brent, 2017  
Chapter 22 © Nicholas Hookway and Helene Snee, 2017  
Chapter 23 © Christine Hine, 2017  
Chapter 24 © Henrietta O'Connor and Clare Madge, 2017  
Chapter 25 © Katie M. Abrams and Ted J. Gaiser, 2017  
Chapter 26 © Jon Hindmarsh, 2017  
Chapter 27 © Christina Silver and Sarah L. Bulloch, 2017  
Chapter 28 © Louise Corti and Jo Wathan, 2017  
Chapter 29 © Patrick Carmichael, 2017  
Chapter 30 © David Martin, Samantha Cockings and Samuel Leung, 2017  
Chapter 31 © Matthew Zook, Ate Poorthuis and Rich Donohue, 2017  
Chapter 32 © Brian Beaton, David Perley, Chris George, Susan O'Donnell, 2017  
Chapter 33 © William Revelle, David M. Condon, Joshua Wilt, Jason A. French, Ashley Brown and Lorien G. Elleman, 2017  
Chapter 34 © Harrison Smith, Michael Hardey<sup>1</sup>, Mariann Hardey and Roger Burrows, 2017  
Chapter 35 © Michael Fischer, Stephen Lyon and David Zeitlyn, 2017  
Chapter 36 © Grant Blank, 2017

First edition published 2008

Apart from any fair dealing for the purposes of research or private study, or criticism or review, as permitted under the Copyright, Designs and Patents Act, 1988, this publication may be reproduced, stored or transmitted in any form, or by any means, only with the prior permission in writing of the publishers, or in the case of reprographic reproduction, in accordance with the terms of licences issued by the Copyright Licensing Agency. Enquiries concerning reproduction outside those terms should be sent to the publishers.

**Library of Congress Control Number: 2016937391**

**British Library Cataloguing in Publication data**

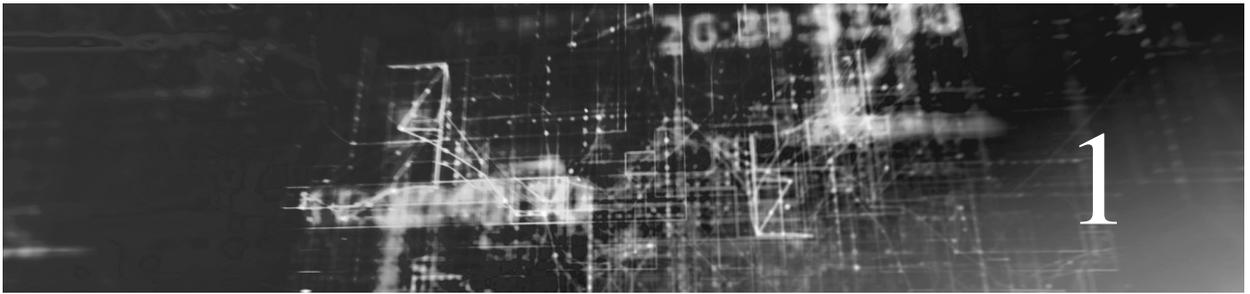
A catalogue record for this book is available from the British Library

ISBN 978-1-4739-1878-8

PART I

# Online Research Methods





# Online Research Methods in the Social Sciences: An Editorial Introduction

Raymond M. Lee,  
Nigel G. Fielding and Grant Blank

Online research methods have come of age as the permeation of everyday life by information and communication technologies has grown ever more ubiquitous. Although substantial digital divides remain by country, and within countries by age, gender and socioeconomic status, the number of Internet users worldwide quadrupled between 2000 and 2014, and the current proportion of the world's population using the Internet is now said to be in excess of 40 per cent (International Telecommunications Union, 2015). Information and communication technologies have had socially transformative effects. They increasingly affect how people make and maintain social relationships, the structure of their social networks, how they go about their work, meet their partners, educate their children, how they shop, take their leisure, present themselves to the world and store their memories. Such things are, of course, of interest to social scientists in and of themselves. However, to study them also requires methods of communication, ways of

harvesting and capturing information, observational strategies and tools for collaboration, not to mention analytic techniques adapted to what are often novel forms and volumes of data, all of which themselves have the capacity to be transformed by new technologies.

Introducing the first edition of *The SAGE Handbook of Online Research Methods*, we emphasised the newness of online methods, and the need for a cautious and critical appraisal of their use and potential. Less than a decade onwards, the terrain occupied by online research methods has changed rapidly, social researchers across a wide range of social science disciplines have become much more familiar with such methods, more adept at their use, and more attuned to the issues and challenges that they pose. As before, our primary purpose in this Handbook is to explore this terrain by highlighting across a wide range of areas the key facets of online research methods and their implications for practice. Given our focus, as was true of the

first edition, we pay relatively little attention to theoretical discourses on the wider social or cultural significance of online environments. While we recognise the historically contingent and socially constructed nature of the changes wrought by development of Internet-based technologies, we leave investigation of such issues to others. So too, we take largely as a given the infrastructural 'substrate' (Star, 1999) that underpins online practices; the standards, protocols, mechanisms, tools and resources without which activity online would be impossible. Neither do we address in any systematic way the drivers of methodological innovation in the social sciences and the social processes that have allowed new online methodologies to be adopted, diffused and used. The Handbook, in other words, retains a pragmatic focus on the current state of the art and on the further potential of online research methods in the social sciences.

## DESIGNING ONLINE RESEARCH

Readers will find in the Handbook comprehensive and detailed coverage of a wide range of online research methods, some possibly more familiar than others. Clearly, though, there are wider issues that crosscut the investigation of particular research problems or the use of particular research methods. For example, in designing a particular study it is necessary to assess how far one's methods and procedures meet the aims and objectives set out for the research, and researchers must attend to the ethical issues surrounding their research.

One debate that emerged early on in relation to online methods was the question of whether the ethical issues they posed were distinct and unique compared to those associated with offline methods. In their chapter on the ethics of online research, Rebecca Eynon, Jenny Fry and Ralph Schroeder argue for the essential continuity between online and

offline methods in relation to research ethics, and although they recognise the importance of ethical governance frameworks they emphasise the importance of the need to make ethical judgements in a context-dependent way. They usefully address the issues that arise in a number of different research situations that include the risks and benefits involved in using online methods to gather data directly from individuals and the challenges involved in obtaining informed consent in such situations, the sometimes novel ethical questions that arise when researchers directly study social interactions in virtual environments, and the increasingly important area of how data generated by social media might be analysed in an ethically responsible way. They bring their chapter to a close by pointing to the challenges posed for online researchers by issues relating to the fluid boundaries between public and private, the potential that arises in some cases for third party reuse of data, the complexities that come with a growing interdisciplinary focus in online research and the implications for ethical practice posed by the existence of digital divides.

Although ethical and legal frameworks provide a largely inescapable context within which a given research project must be conducted, the specific methods used in the study need, of course, to be carefully weighed and considered in relation to its aims. This has, perhaps, not always been the case as far as online methods are concerned. Their relative newness has in the past prompted both unthinking enthusiasm on the one hand, or unreasoned resistance, on the other. There is merit, therefore, in taking a careful, balanced and nuanced approach to the strengths and weaknesses of online approaches.

A putative advantage of researching online is that data can be acquired quickly and often in considerable volumes. The temptation exists, thinks Karsten Boye Rasmussen, to accept the benefits this brings without a parallel commitment to scrutinise the quality of the data so produced. In his chapter Rasmussen argues the need for a systematic

theoretical model of data quality as a basis for assessing the ability of online methods to generate reliable and valid data. Emphasising within this framework the importance of ‘fitness for use’, Rasmussen points to opportunities for assessing data quality that arise as traditional research methods such as the survey move online. In addition, online methods provide novel sources of data with a built-in capacity for quality assessment. In both cases the potential to assess and ensure data quality is enhanced by the ‘documentality’ of online data – in other words its ability to be described via ‘metadata’, as well as the ability to associate it with ‘paradata’ – the data produced as part of the process by which data are collected.

### **ONLINE DATA CAPTURE AND DATA COLLECTION**

One can argue with probably only a little exaggeration that for much of the twentieth century direct elicitation was the method of choice for many social scientists. In other words, it was thought that the way to discover what people thought and did was to ask them directly, usually by means of an interview of one sort or another. A relatively unnoticed aspect of this was that the popularity of the interview as a method depended on a variety of technological developments including in the case of qualitative research the miniaturisation of audio recorders (Lee, 2004) and the advent of long-distance telephone lines that, in the United States at least, fostered the development of telephone survey interviewing. In the twenty-first century, there has been a decisive move away from elicitative methods. This shift has largely been fuelled by a massive extension in the availability of online communication technologies, and by a growing ability to measure more and more aspects of everyday life as and when they occur through the use of data harvested from social media sites. Where even

a few years ago names like Twitter, Facebook, Instagram and the like might only have produced quizzical bemusement if not puzzlement, particularly among older social scientists, social media data available in large volumes now form an increasingly large part of the landscape of social science research.

As this Handbook indicates online research methods are very diverse. They are used across the social science disciplines and produce data, whether directly elicited and not, that manifests itself in numeric, graphical, textual and audio-visual formats. The contexts within which online data are produced range from tightly designed experiments through to looser more naturalistic approaches, the gathering of various forms of non-reactive data, not to mention simulations and games or research in virtual environments. Claire Hewson traverses this terrain in her chapter on designing online research. Emphasising the importance of maximising the trustworthiness, reliability and validity of data produced online, Hewson systematically examines the possibilities, trade-offs, constraints and opportunities researchers need to consider when generating obtrusive and unobtrusive research data online.

The machine-readable traces that our increasingly self-documenting and self-archiving world leaves behind can be thought of as ‘unobtrusive’ or ‘nonreactive’ measures, to use a term popularised by Webb *et al.* (1966) half a century ago. Their now classic monograph was partly meant as a rebuke to the often uncritical use of interviews and questionnaires common at the time they were writing, but it also emphasised the creative appropriation of often quite fleeting behavioural manifestations as sources of data. In his chapter, Dietmar Janetzko attempts in particular to extend the conceptual understanding of nonreactive data by examining ways in which the rather ‘thin’, i.e. non-contextualised, nature of such data can be extended either through triangulating multiple sources of data or the use of newer techniques such as text mining. Janetzko also

enumerates the many different sources of nonreactive data to be found online and provides a detailed guide to the complexities of using such material.

As Ayelet Baram-Tsabari, Elad Segev and Aviv J. Sharon point out in their chapter, the term 'data mining' is relatively new in the social sciences but has become increasingly used in the last decade, fuelled it would seem by the growing popularity of online user-generated content. Data mining involves the automated processes associated with the extraction of knowledge from large-scale databases or online repositories. Baram-Tsabari and her colleagues usefully set out how data mining approaches differ from traditional quantitative methods. They examine the characteristics that make datasets suitable for mining as well as the resources needed to analyse them. In the main part of their chapter they give state-of-the-art examples of data mining techniques in relation to studies of mainstream media, data generated by users of social media and metadata.

As do other contributors to this Handbook, Martin Innes, Colin Roberts, Alun Preece and David Rogers see the need for a discerning approach that cautiously welcomes the opportunities created by the abundance of social media data now available while at the same time critically evaluating the social and technical processes implicated in their production, consumption and use. Innes and colleagues guide readers to an understanding of social media instrumentation, providing in the process an overview of how the data available on various social media platforms might be accessed. They also provide a detailed case study of how they combined to mutually implicative effect analysis of the social media data surrounding a particular event with on-the-spot ethnographic observation taking place at the same time.

In his chapter, Jonathan Bright investigates the issues surrounding the use of 'big data' in the social sciences, the large volumes of data about diverse aspects of social life that have become available as the ability to store and

process such volumes becomes computationally possible. Bright provides an introduction to methods for capturing big data, as well as the processes involved in rendering the material more useful for analytic purposes through proxy variables and data coding. He then goes on to point to some of the complexities surrounding the analysis of big data, taking a somewhat sceptical view of some elements of current practice. Arguing that the methods training currently available to social scientists is seldom sufficiently oriented to the skills needed to work with big data sources, Bright describes some of the specific elements that make up the toolkit that social scientists increasingly need in order to be able to deal adequately with large datasets.

## THE ONLINE SURVEY

Survey researchers have rarely shied away from the latest technological developments available to them and, true to form, were not slow to explore the possibilities for survey deployment opened up by the Internet. In both market and academic research, the use of online surveys is now well established. Nor has development been in any sense static. Researchers have begun to adapt to newer circumstances such as the growth in the use of mobile phones, while looking forward to possibilities that currently remain on the horizon such as the use of smart televisions as survey delivery systems.

Vasja Vehovar and Katja Lozar Manfreda give an overview of the current state of the art in their chapter on online surveys. Conceptually they locate online surveys within a wider set of technologically mediated data collection methods collectively referred to as 'computer-assisted survey information collection' (CASIC). As Vehovar and Manfreda observe, online surveys provide some of the traditional benefits of self-completion methodologies, but with advantages over conventional paper and

pencil methods that include cost and error reduction, the possibility to increase respondents' motivation and understanding, as well the ability to use advanced design features not available within non-digital contexts. On the other hand, if researchers are to make effective use of online survey methods, they need to confront a range of issues and challenges. Among the considerations outlined by Vehovar and Manfreda are issues to do with recruitment, sampling and non-response, how design elements are used within a survey instrument and the use of post-survey adjustments. They then extend their discussion to the use of single and mixed-mode surveys as well as mixed-method approaches. Many of these topics are subsequently taken up in detail in the other chapters making up this section of the Handbook.

In his chapter on sampling methods for web and email surveys Ron Fricker swiftly but carefully rehearses the fundamentals of sampling before going on to review the applicability of a range of probability and non-probability sampling methods to online surveys. He profiles the various methods of sampling – including the use of pre-recruited panels – that might be used and looks at the issues and challenges associated with their use. As do other writers in this section, Fricker recognises that the difficulties involved in generating probability samples online encourages the use of mixed-mode surveys. Fricker concludes with a look to the future, suggesting that in the shorter term online survey sampling is likely to remain problematic, but noting that with online technologies still in their infancy it is unclear what the future might bring.

It is difficult to spend any time online without receiving a request to participate in an online survey. Low cost, ease of administration and apparent reach all combine to make survey delivery online attractive to marketers, bureaucratic administrators and academic researchers alike. As Vera Toepoel points out in her chapter on online survey design, intriguing possibilities emerge from

the move to online surveys, particularly in relation to mobile data collection, and the extension of survey materials beyond the merely textual. At every stage, however, researchers need to take on board the concomitant challenges to conventional survey practice thrown up by online surveys. Toepoel identifies these challenges and takes readers through the various stages of designing, collecting and administering an online survey.

Nowadays anyone wanting to mount a survey online can choose from a wide range of survey software products. Lars Kaczmirek makes the point that the market for such software is now very diverse indeed. Settling on a suitable product can be daunting. Kaczmirek's chapter clears a path through the complexities involved. In it he provides a conceptual schema that helps potential users of survey software to identify uses, needs and priorities in a systematic way, allowing them to focus on that which is likely to be best suited to their needs.

Email on its own is a rather imperfect mechanism for online survey recruitment. Researchers often need to combine it with other methods, such as mail or telephone, to obtain an adequate sample of survey participants. Mixing survey modes is not a simple matter, as Don Dillman, Feng Hao and Morgan Millar point out in their chapter on the topic. Dillman will be well-known to many survey researchers as the originator of the 'total design method' (1978) and later the 'tailored designed method' (2000). Rather in the spirit of that work, he and his colleagues offer a holistic, comprehensive and practical account of mixed-mode work, setting out a series of detailed recommendations dealing with the timing and staging of contacts, the use of incentives and the possible ramifications of using different question formats across modes. As do other writers in this section, Dillman and his colleagues draw attention to the possibly problematic implications of the 'smartphone revolution' for survey practice.

## DIGITAL QUANTITATIVE ANALYSIS

Much material available online lends itself to quantitative analysis. The cost of ready availability, however, has often been analytic complexity. While for some this might constitute a barrier, the opportunity afforded by online methods to study dynamic and inter-linked aspects of social life in ways that are often absent from more traditional approaches has also brought newer tools and approaches to the fore.

It might be a truism to say not one of us is an island, but that social life is inherently relational – with each one of us linked to others through a web of strong and weak ties – is one of the fundamental insights of the social sciences. As its very name doubly implies the Internet is inherently relational. It is not surprising, therefore, that researchers quickly turned to the study of online phenomena such as email, web linkages and social networking sites. Often, as Bernie Hogan points out in his chapter, such studies utilise network analysis, a thriving area of research that emerged from the convergence of work on the mathematics of graphs with empirical studies of social relations by anthropologists and others. Hogan provides a useful primer on network analysis. He looks at the analytic choices one might make in studying an online network. Should one decide, for example, to focus on the relationships within a particular bounded population, the networks associated with particular individuals or the relational paths one can follow from a particular starting point? He points to the practicalities involved in extracting and managing data from online sites and gives a useful outline of techniques involved.

Javier Borge-Holthoefer and Sandra González-Bailón take up and extend the discussion of network methods by focusing on advanced analytic techniques. Noting the importance that now attaches to social media data in studies of social interaction and the potential thus created to revitalise long-standing debates in areas related to

interpersonal communication, they argue that analytic techniques suitable to data generated by traditional methods such as surveys need to be revamped. Specifically, they point to the need to define rules for aggregating and filtering data available from online social networks. In their chapter, Borge-Holthoefer and González-Bailón describe a range of newer methods, including approaches borrowed from studies of physical or biological systems that have recently come to the fore.

Introducing her chapter on simulation methods Corinna Elsenbroich observes that the social world is inherently dynamic. There is also a duality to it that social scientists have often encapsulated in distinctions between the micro and the macro, agency and structure and the like, and yet our methods seem best fitted to capture the static elements of social life and only one side or other of its polarity. For Elsenbroich, simulation overcomes these deficiencies. Although there are a number of different kinds of simulation, Elsenbroich focuses on agent-based modelling, a computer-based method in which interactions between micro-units called agents are used to generate macro-level patterns. For example, from simple assumptions about preferences for neighbourhood composition it is possible to examine how patterns of residential segregation might emerge. Heretofore, social simulators have had to rely on sources of data not necessarily well-suited to their purpose. Elsenbroich sees considerable potential for synergy between simulation methods and online research. The availability and dynamic character of much online data makes it amenable to analysis using agent-based modelling, which in turn allows often hard-to-study processes such as diffusion to be analysed.

Gaming was early on an important aspect of online culture. Harko Verhagen, Magnus Johansson and Wander Jager address the issues involved in researching games. One can study how games are played online or look at the social worlds that surround gaming and how they manifest themselves online. Since gaming is typically an immersive

activity, the study of games poses a number of methodological challenges, as well as a range of ethical issues. Games shade over into simulations, making them a research method in their own right. Thus, playing a game in which the participants must engage with a difficult problem through processes of interaction and negotiation allows one to gain insight into how such a problem might be dealt with in the real world.

Hans Rosling, doctor, statistician and anti-poverty campaigner, once reputedly said 'Most of us need to listen to the music to understand how beautiful it is. But often that's how we present statistics; we just show the notes we don't play the music'. In their chapter Helen Kennedy and William Allen aim to help online researchers go beyond simply showing the notes by using visualisation techniques to represent data in clear and, more often than not, beautiful ways. Of course, form can sometimes overwhelm content and after defining data visualisation and discussing both the possibilities and the limits of what visualisation can achieve, Kennedy and Allen emphasise the need for a strongly reflexive approach to the use of visualisation. Beyond this, they characterise the state of the art through an examination of the tools and techniques available for creating visualisations and give examples from their own work.

## DIGITAL TEXT ANALYSIS

The metaphor of the 'field' comes fairly readily to social scientists, a comfortingly agricultural metaphor for a place where one goes to 'gather' data. For online researchers though, the notion has begun to seem like an anachronism. Rather it is as if one is standing in a river with data flowing, cascading even, from a variety of data providers – individuals, social media sites, companies and so on – and in need of capture. The necessity to deal with volume and flow has encouraged social

scientists to think about ways of automating the analytic process.

Roel Popping looks at the use of content analysis as an analytic strategy. Content analysis is understood here as a systematic, quantitative approach that provides a basis for an understanding of a text or set of texts of interest to a researcher. Popping provides a clear overview to the field. He identifies the major theoretical approaches involved, discussing in each case both manual and machine coding methods. In particular, Popping explores the use of 'modality' analysis, an approach useful for the analysis of opinion statements of the kind often found in newspaper editorials that proclaim the need for some action or promote the desirability of a particular state of affairs. He concludes by providing information about appropriate software and emphasises the need to train coders and to ensure intercoder-reliability.

One approach that has come to the fore especially with the advent of social media is 'opinion mining' or more broadly 'sentiment analysis', terms used to refer to the automated identification and extraction of opinions and information about affective states from (often voluminous) online texts. Observing that such methods have become increasingly effective, Mike Thelwall discusses the main features of sentiment analysis and the various forms it takes, which might include the detection of subjective statements, the strength of a sentiment, its polarity, emotional tone and so on. The possible applications of sentiment analysis, which include both academic and commercial uses, are now quite extensive and, as Thelwall shows, hold considerable potential for studying patterns of affective communication hitherto not always well-studied by traditional methods.

Edward Brent suggests that the need to deal with large-scale digitised data flows might best be met by means of automated processes, specifically the use of 'intelligent agents' that leverage natural language processing and other artificial intelligence techniques to develop ways of coding data

as it flows towards capture by the researcher. Brent sees the vision he sets out as one that will become increasingly important in the future. Inevitably, he observes, concerns arise about privacy, intellectual property, and about the possible deskilling of researchers. Nevertheless, the possibilities are intriguing.

Around the turn of the millennium, the weblog or blog, a relatively new form of online communication, began to become popular. Through the medium of a blog one could produce online content relatively easily and link readily to the work of others similarly engaged. For Nicholas Hookway and Helene Snee part of the interest in blogs lies in the ways in which they make the personal public. In their chapter Hookway and Snee see blogs as 'documents of life' (Plummer, 2001), narratives produced spontaneously that give us insight into how people live their lives, more like traditional forms such as diaries or journals. Using case studies from their research, Hookway and Snee give a clear and detailed account of the processes involved in researching blogs. They look at the practical and technical aspects of doing blog research, as well as issues to do with selecting blogs for analysis and extracting data from them. Analytic issues are also addressed, for example the important issue of authenticity, and Hookway and Snee conclude with discussion of ethical and legal matters.

## VIRTUAL ETHNOGRAPHY

Peter Steiner's celebrated 1993 *New Yorker* cartoon in which one dog tells another 'On the Internet, nobody knows you're a dog' hints at some of the attraction online worlds had early on for ethnographers. The online was a space that was novel and exciting and – because or in spite of its technological carapace – perhaps even a bit mysterious. Within that space it might be possible to learn interesting things about identity, culture and the presentation of self. Thus, the online became grist to the ethnographer's mill.

Today, Christine Hine argues in her article on virtual ethnography that there exists 'an internally diverse array of approaches oriented to ethnography in and of online space' rather than a single dominant approach. Hine draws some of the strands together by identifying key methodological issues that surround participation and observation within online research settings and by addressing complexities in the definition of field sites. She offers a typology of ethnographic approaches depending on the degree to which the activities studied are interconnected and how these relate to the goals the researcher brings to the study. Looking forward, Hine addresses the potential for autoethnographic approaches while seeing challenges ahead related to the growing commercialisation of the Internet and the difficulties involved in studying the consumption of online material.

Henrietta O'Connor and Clare Madge point out that despite the proliferation of online methods, online synchronous interviewing where interviewer and interviewee interact remotely but in real time remains, for the present at least, relatively underused. In their detailed chapter, O'Connor and Madge look at the advantages and disadvantages of interviewing online and contrast online interviewing with interviewing face-to-face. As well as discussing the ethical issues involved, they address the practicalities of interviewing online and give advice on available software. In their conclusion, O'Connor and Madge emphasise the need to weigh carefully the strengths and weaknesses of online interviews and look forward to the ways in which newer technological developments might expand the scope for online interviewing.

The opportunities offered by online focus groups, as well as the issues involved in their use, are discussed in the chapter by Katie M. Abrams with Ted J. Gaiser. In this chapter readers will find a discussion of the methodological and technical considerations they will need to bear in mind when selecting a medium for conducting an online focus group. Various approaches and tools are discussed,

the medium to be used, whether communication is synchronous or asynchronous, recruitment and the demanding task of moderating an online group. The factors that need to be considered in choosing a particular technology to be used in data collection are also outlined. As with the online interview, ongoing technological developments are likely to open up a space for greater and probably more innovative focus group practice online. The chapter closes with a look at some of the possibilities.

Drawing on the work of a research project devoted to developing tools to support remote working with video data, Jon Hindmarsh's chapter looks to the needs of qualitative researchers who analyse digital video, an area of growing importance in the social sciences. Although software tools for qualitative analysis have become increasingly sophisticated, Hindmarsh notes that they do not always meet the needs of video analysts working in a research tradition associated with ethnomethodology and conversation analysis who focus on small slices of locally situated and occasioned interaction and who prefer to use video data because it allows recurrent viewing and inspection of the data with a high degree of granularity. The analytic needs of such users intersect with an institutional form within the field, the 'data session' in which researchers collectively and collaboratively view video materials for the purpose of analysis. Such sessions require both the physical co-presence of participants and a means of interacting with the video in immediate and complex ways. Hindmarsh describes recent technological developments that provide tools for allowing colleagues who are physically remote from each other to collaborate in a highly interactive and responsive manner in the analysis of visual data.

Beginning in the late 1980s, researchers began to use software tools for the analysis of data from qualitative research studies. Originally somewhat controversial, such tools eventually moved to the mainstream and became what some would regard as an

essential feature of contemporary qualitative research practice. Now, as Christina Silver and Sarah L. Bulloch discuss in their chapter, the field of Computer Assisted Qualitative Data Analysis (CAQDAS) is being shaped by its relationship to online research methods. As they point out, a number of key trends have become apparent in the past few years. CAQDAS packages are now capable of handling a wider range of data formats, moving beyond textual data to incorporate material from visual, audio, bibliographic and online sources. There is a trend to technologically mediated collaborative working and a move to make software available on a wider range of platforms and in mobile versions, all of which have interesting implications for ethnographic styles of work. Citizen research, collaborative work, as well as use in commercial environments are all facilitated in various ways by recent developments. Silver and Bulloch chart these trends and their ramifications based on a detailed familiarity with available software and the changing nature of the field.

## **ONLINE SECONDARY ANALYSIS: RESOURCES AND METHODS**

Probably most of us today make a fairly serious attempt to reuse and recycle what we produce and consume; however, 'waste not, want not' makes not just environmental sense. The benefits to researchers of using previously collected data as a resource for further study are now well understood and well documented. In addition, many of the tools and resources for doing so are now available online.

Some of the uses to which secondary data can be put are rehearsed by Louise Corti and Jo Wathan in their chapter on online access to quantitative data resources. These include the contextualisation of existing studies, comparative research, replicating existing studies, the asking of new questions of old

data, methodological research and so on. Focusing on the United Kingdom's Data Archive at the University of Essex and the Interuniversity Consortium for Political and Social Research (ICPSR) in the United States, Corti and Wathan point to the role of data services in ensuring the availability for reuse of high quality research data. Most users interact with data services via online portals that make it relatively easy to find and access data, but the availability of data in this way depends on a great deal of background work to produce data files and documentation in serviceable and durable form. Now data archives have to deal with new and emerging forms of online data available, for example as the result of open government initiatives, data from online transactions, social media and crowd-sourced data. Corti and Wathan explore how data services assess the provenance and quality of these newer forms of data, look at some existing examples and point to future developments.

The issue of how far qualitative data might lend themselves to secondary analysis has been a somewhat contentious one in the field. Although he recognises the sensibilities some qualitative researchers have in relation to the issue of secondary analysis, Patrick Carmichael underlines the diversity of form, purpose and content that can be found in existing collections of qualitative data and makes a pragmatic case for reuse, not least in relation to research training. Using as a case study a project designed to develop a digital archive of the data emerging from a series of educational evaluation studies, Carmichael addresses issues of various kinds that arise from the secondary analysis of qualitative data. He discusses in a relatively non-technical way strategies for data description and their relationship to existing and emerging network technologies, all of which opens up, in his view, a range of interesting possibilities for the provision of data that can be utilised in highly complex and novel ways. Carmichael concludes by discussing a range of new developments such as 'linked'

and 'open' data, the possibilities that exist for methodological innovation and the ways in which the role of researchers might change in terms of research impact, for example.

Taking a bus to work used to involve turning up at the bus stop and hoping that the service was running to schedule. Now, a smartphone app tells you where the bus is and when it is going to arrive. This is just one example of the role geographical data now plays in everyday life. As David Martin, Samantha Cockings and Samuel Leung point out in their chapter on finding and investigating geographical data online, although much social science data is analysed without reference to its spatial location, almost all the objects of study that social scientists are interested in have a spatial location. In their chapter they examine a range of online sources of geographical data before going on to identify online tools for data linkage, various forms of mapping and spatial analysis. Martin and colleagues are enthusiastic about the potential for greater use of geo-referenced data by social scientists, although they draw attention to the rapid pace of change in the field and caution that there are issues to do with scale, projection, accuracy and precision that might not be apparent to non-geographers.

As Matthew Zook, Ate Poorthuis and Rich Donohue point out, for most of us a map describes locations; it shows us where things are. Social scientists, however, are generally interested in thematic maps that show how social attributes or variables are spatially distributed. Zook and colleagues walk non-specialists through the various stages involved in producing such maps, paying attention to issues of measurement, generalisation and graphic design and detailing some of the software tools available. They then illustrate the issues involved using as the basis for a case study a sample of geotagged tweets using the term 'pizza' sent in the United States between 2012 and 2015. Spatially mapped, these tweets give insight into regional and cultural variations in food consumption, the analysis of which allows

Zook and colleagues to describe the methodological complexities associated with spatial analysis.

## THE FUTURE OF ONLINE SOCIAL RESEARCH

New methods throw up unexpected challenges and opportunities and place old problems in a new light. Technological change often makes previously intractable problems and bottlenecks resolvable. The prospect is to know the world in ways not previously possible with tools still to be envisaged. That prospect is an exciting and compelling one, and one that will have widespread methodological implications for social research. At the same time, it should not be forgotten that new technologies also shift the social relations of intellectual production. A case in point is the extent to which access to online data is increasingly constrained and controlled by commercial entities and proprietary interests. The balance of power between researcher and researched has also shifted. Interesting possibilities for citizen research, action research and the use of participatory approaches have opened up as a result. The wider implications of all of this are not entirely clear at present but require careful attention nevertheless.

A critique emerging in recent years associates dominant research traditions in the social sciences with Western colonialism and imperialism and emphasises by contrast the importance of using research to advance the needs, aspirations and cultural integrity of colonised peoples, as defined and articulated by those peoples themselves. Using as a case study their work with First Nations communities in Northern Canada, Brian Beaton, David Perley, Chris George and Susan O'Donnell point to ways in which new technologies coupled to participatory research styles can aid the empowerment of marginalised groups. They describe how

the availability of broadband networks and the use of video-conferencing tools enabled collaborative and participative working with small, widely scattered, remote First Nations communities with some history of suspicion towards research conducted by metropolitan academics.

The advent of mobile communication technologies opens up many possibilities for continuous and mobile data collection; however, as William Revelle, David M. Condon, Joshua Wilt, Jason A. French, Ashley Brown and Lorien G. Elleman suggest, the ability thus provided to collect data online from a large and diverse pool of participants is somewhat constrained by design considerations that limit their ability or willingness to respond to large numbers of items. Using an approach for dealing with the problem, described as 'Synthetic Aperture Personality Assessment', Revelle and colleagues suggest a strategy in which participants are given a small set of items of interest which are then analysed through the use of synthetic covariance matrices using software tools that are freely available.

What Harrison Smith, Michael Hardey, Mariann Hardey and Roger Burrows refer to as the 'Geoweb' or 'geo-spatial web 2.0' is based on what they call a 'new social cartography' that harnesses new technologies to allow ordinary citizens to create and use maps through practices such as crowdsourcing. The contrast here is with 'cartographies of knowing capitalism' in which the power of Geographic Information Systems is harnessed to produce knowledge that aids processes of capital accumulation. Smith and colleagues explore the epistemological dynamics of the Geoweb and the implications that developments such as knowledge production by non-experts and wider use of open data sources have for the social relations of data production. They examine a number of Geoweb tools applications that have potential for future research and praxis.

Michael Fischer, Stephen Lyon and David Zeitlyn look to the future of social science

research under the impact of what they call 'Internet and related communications technologies' (IRCT). Fischer and colleagues suggest that short-term trends at least are probably foreseeable from an inspection of what is happening now at the cutting edge (much of which is represented in this Handbook). Extrapolation into the medium term and long term, however, remains problematic. Certainly, online research will become more important as time goes on, although as they argue, current distinctions between online and offline might largely disappear as the two worlds increasingly interpenetrate. Continuing developments in IRCT will have implications right across the research process from the collection of data, through its handling, manipulation and analysis to the means by which findings are disseminated. Moreover, beyond the execution of research, new possibilities will open up for the design, conceptualisation and theorisation of research, while the emergence of formidable ethical challenges is also a possibility. Social scientists will need to respond to developments such as the advent of 'smart' technological assistants and come to terms with the research implications of the Internet of Things. At the very least, the possibilities and options open to coming generations of social scientists will be very different from those faced today.

In the concluding chapter of the Handbook, Grant Blank reminds us that the complex relationship between theory, method and the technologies for recording and analysing data has stood at the heart of disciplined inquiry since the dawn of the Scientific Age. Now, the advent of online methods casts that relationship anew. The promise of new information and communication technologies seems to be that we will have so much data available so readily, in such volumes and in such detail that there will be little need for theory. Usefully revisiting many of the topics discussed in individual chapters of the Handbook, Blank argues by contrast that theory is deeply and continually embedded in the choices we make to deploy online research methods.

## CONCLUSION

Information and communication technologies have affected research capacities in all fields of scientific endeavour but, arguably, they are of particular importance to the social sciences, offering means to address some hitherto intractable methodological problems of social science methods while providing a view onto the overall terrain of contemporary human knowledge, albeit one that is very large, very unruly and constantly changing. It is clear that online technologies have had, are having and will have transformative effects on what it is that social researchers do. In the meantime, the emergence of even newer technologies, some of which we can only now imagine, will engage the attention of social researchers. It is with this in mind that we have brought together a range of contributions relating to online research methods. Drawing on authors well known in their field from the United Kingdom, North America, Continental Europe and Australasia, we deliberately sought broad topic coverage in compiling the Handbook. Although all committed to the importance of empirical research, the authors of the preceding articles come from a range of epistemological traditions and embody a variety of methodological styles, substantive commitments and disciplinary affiliations. Many are early adopters who have contributed to the substantive literature in their own particular field and have demonstrated how the often previously unrecognised affordances associated with online methods were capable of extending and enhancing the doing of social science research. Authors who contributed to the first edition of the Handbook have brought their contributions up to date to ensure that readers have the clearest sense of the current state of the art. (Regrettably, we were unable to include updated versions of two chapters from the first edition because authors had competing claims on their time.) In addition, we have added or expanded coverage of some areas – for example big data, gaming and participatory research – where

there appears to be new and promising developments. Although some of the areas covered in the Handbook are technically complex, we have encouraged authors to address issues in a clear accessible way so that newcomers have a clear introduction to a particular field while those already familiar with it can be appraised of the newest developments.

New methods throw up unexpected challenges and opportunities and place old problems in a new light. Technological change often makes resolvable previously intractable problems and bottlenecks. In thinking about technological innovation in social research, it seems important to steer a path between a number of different positions. Quite obviously one of these is the kind of naive enthusiasm that is largely a matter of being in thrall to the latest fads and foibles. The newness of a method can lead to unthinking application and a distancing of users from the craft aspects of a particular methodological approach. For any given innovation someone has to be an early adopter. However, just as in artistic experimentation, where what seems outrageous to established taste might be, from the artist's point of view, a subtle exploration of where the boundaries of possibilities lie, so too it is important methodologically to assess what we gain and what we lose with any new way of doing things. This suggests that any assessment of online research methods needs to be sober enough to undermine exaggerated claims but open-minded enough to spot potentiality where it exists. Self-evidently the contributors to this Handbook are enthusiasts for the methods they describe. What they share in addition, however, is a commitment to the critical understanding of those methods. That is, they recognise that the very considerable opportunities opened up by online methods must also be assessed and evaluated. The implications of those methods need to be teased out and the contexts and consequences of their use analysed and theorised. Neither unthinking advocacy of the new or its curmudgeonly rejection serve well the cause of methodological innovation.

There are indications in the early decades of the twenty-first century that the boundaries of social research itself face possible reconfiguration. Although individuals, organisations and governments have always controlled access to data, the extent to which data sources and the methods for extracting data are now controlled by commercial entities represents a new challenge to social scientists. Indeed, it is the political economy of online methods, not always apparent at the level of day-to-day practice, that remains perhaps the most opaque and complex aspect of future methodological development. The ongoing dance of competition and cooperation, accommodation and antagonism between corporations and governments that has been shaped differentially by culture, history and self-interest in North America, Europe and elsewhere will no doubt continue to affect the balance of power between knowledge producers and consumers, including social researchers. Against this, the increasing availability and tractability of online tools and sources makes for a more research-literate and research-inclined orientation amongst non-academic users (Savage and Burrows, 2007). Indeed, it can be argued that the availability of online tools has facilitated a trend to research by 'ordinary' citizens. Citizen research looks like a trend that it would be futile to try to brake, which can presumably be seen as desirable at a time when disengagement from established political institutions is widely remarked. It could also lead to some improvement in the accessibility and design of online information resources on the grounds that lay people will not put up with the more forbidding kinds of information resource that the technically proficient may presently tolerate. This development, however, might also conceivably lead to a degree of competition between amateur and professional researchers, a circumstance that has implications for resources, such as this Handbook, which might have a role in educating or even regulating an expanded user base.

No longer a large rather foreboding machine, its console full of blinking lights, humming away in an air-conditioned room, the computer is now in your pocket. It is used to make calls, send messages, take pictures, check the time of the next train and what is showing at the local multiplex. The quotidian character of computing nowadays as well its massive interconnectedness draws researchers to online environments, just as their traditional tools are themselves being transformed by technology. Soon, everything will be 'smarter', more embedded and more interconnected. Interesting times ahead!

## REFERENCES

- Dillman, D. A. (1978). *Mail and telephone surveys: The total design method*. New York, NY: Wiley.
- Dillman, D. A. (2000). *Mail and internet surveys: The tailored design method (Vol. 2)*. New York, NY: Wiley.
- International Telecommunication Union. (2015). *Measuring the Information Society Report 2015*. Geneva: International Telecommunication Union.
- Lee, R. M. (2004). Recording technologies and the interview in sociology, 1920–2000. *Sociology: Journal of the British Sociological Association*, 38(5), 869–89.
- Plummer, K. (2001). *Documents of life 2: An invitation to a critical humanism*. London: Sage Publications.
- Savage, M. and Burrows, R. (2007). The coming crisis of empirical sociology. *Sociology*, 41(5), 885–99.
- Star, S. L. (1999). The ethnography of infrastructure. *American Behavioral Scientist*, 43(3), 377–91.
- Webb, E. J., Campbell, D. T., Schwartz, R. D. and Sechrest, L. (1966). *Unobtrusive measures: Nonreactive research in the social sciences*. Chicago, IL: Rand McNally.