¹Chapter 28

Mixed Methods and Multimodal Research and Internet Technologies

Christine Hine

Abstract

This chapter explores the varied ways in which qualitative and quantitative, large-scale and small-scale research designs have been combined in the study of the Internet. The chapter also considers multi-modal and online/offline studies, which explore a research object by combining diverse forms of data drawn by crossing the offline/online boundary or combining digital and analogue data. The focus of the chapter is on exploring the motivations for these combinations and outlining the distinctive issues that heterogeneous research designs in these domains encounter. Internet research has occasioned many innovative research designs, as the sheer quantity of data available online often prompts even those overtly committed to a qualitative paradigm to explore techniques for visualizing data and exploring patterns on a larger scale. The spatial

---

¹ The Oxford Handbook of Multimethod and Mixed Methods Research Inquiry, Edited by Sharlene Nagy Hesse-Biber and R. Burke Johnson. Oxford University Press (2015).

complexities entailed in networked communication often, however, introduce a challenge for researchers seeking to combine perspectives offered by different methodological approaches according to straightforward triangulation or integration.

Keywords: Internet, online, ethnography, visualization; hyperlink, social network analysis, social networking site

Chapter objectives

This chapter will:

- Introduce Internet research as a field rich in examples of mixed methods research
- Explore the motivations for mixing of methods in Internet research, particularly within ethnographic studies of the Internet
- Give examples of large scale approaches including network analysis and visualization made possible by “born digital” data and explore how these are complemented by smaller scale qualitative approaches
- Describe the methodological advantages and challenges of mixed mode approaches in Internet research

Introduction

The Internet is still a relatively young phenomenon, but in a short and lively history it has stirred up debate and questioned assumptions in many fields of endeavour, including social science research. The Internet has occasioned the development of new social science research methods and

the re-articulation of traditional methods, and has also stimulated researchers to come up with many examples of novel combinations of method. The history of Internet research contains numerous studies that, whether overtly claiming to use a mixed methods approach or not, are heterogeneous in their designs, combining approaches derived from different methodological traditions. Within the scope of this chapter are included many different forms of mixed method and mixed mode study, including combination of qualitative and quantitative techniques or an amalgamation of data derived from online and offline settings. Some such studies aim for integration at the outset or adopt contrasting approaches synchronously, whilst others implement a sequential approach or leave integration to a final stage where the contribution of different approaches to understanding an overall research question is compared or synthesised. The studies discussed in this chapter fall within the broad definition offered by Johnson et al. (2007, 129) of method combination as an “intellectual and practical synthesis,” in which the mixing of methods is a choice in its own right made according to the demands of a particular research question.

As Creswell (2011) notes, the field of possible ways in which to combine methods is both baffling and complex, and the variety of practice defies attempts to produce a straightforward typology. This is certainly true in the case of Internet research. When I appealed to the mailing list maintained by the Association of Internet Researchers (<http://aoir.org/email-list/>) for suggestions of current mixed methods studies in Internet research, I was overwhelmed by responses describing published studies, ongoing dissertation research and major funded projects (full list of responses available at <http://christinehine.wordpress.com/>). It became clear that the combining of methods, and of methodological traditions, was proving extremely fruitful within Internet research. Many Internet researchers were able to identify themselves with the practice of mixing

methods even if they had not overtly set out to do a mixed methods study, or to describe it in terms of triangulation, sequencing of methods and the like. It was also apparent that it would be challenging to provide a taxonomy of different approaches according to a systematic typology such as that mapped out by Creswell and Plano Clark (2007), and indeed, as Creswell (2011) acknowledged, it often proves to be the case that such a typology turns out to be insufficient to capture the complexity of actual practice. This chapter does not, therefore, review current practice in the mixing of modes and methods in Internet research in a systematic way intended to map differing approaches to design, although it aims to capture some of the diversity of approaches.

The field of Internet research does indeed provide examples of many of the forms of a mixed method study identified elsewhere across the social sciences by Creswell and Plano Clark (2007), including studies where concurrent methods are intended to provide triangulation, those where one method succeeds another in explanatory mode and studies where one method is embedded in another, such as a small quantitative component within a largely qualitative ethnographic study of an online environment. Many of these designs have arisen in somewhat ad hoc fashion within Internet studies rather than as a concerted program of methodological development, as researchers explore the affordances of the data available to them and progressively formulate research questions which morph over time, with research projects sometimes employing unanticipated methods over the lifetime of the project in the face of the development of the Internet itself. To enforce a structure on this emergent diversity of experience would possibly not be the most helpful thing to do. Instead, I have chosen to focus on what the experience of Internet research reveals about mixed and multimodal approaches to research more broadly, focusing on the opportunities and challenges that Internet researchers have

experienced in their attempts to build research designs appropriate for the job at hand. The chapter therefore focuses on why Internet researchers combine methods and explores the distinctive twist that a focus on the Internet confers on attempts to mix research modes and methods.

Instead of systematically assigning studies to a typology of mixed methods approaches, this chapter therefore instead begins with a focus on the motivations for mixed method studies in Internet research, attempting to understand why these approaches seem to flourish particularly well in this field. The first section of the chapter explores these motivations, by considering the emergence of Internet research as a relatively recent phenomenon not colonized from the outset by any particular methodological stance but also characterized by a certain anxiety about the capacity of any one method to capture the situation. The following section then focuses in on Internet ethnography, which has been a particularly rich source of mixed research designs. The next sections then move on to consider some of the distinctive qualities of Internet data which have encouraged mixed methods research, dwelling first on the notion of “big data” and considering the opportunities and challenges offered by the advent of large amounts of digital data emanating from social interactions online. This luxurious abundance of data has notably stimulated the methodological imagination and prompted a wide array of studies that explore both the small scale qualitative understanding of experience and the larger scale notions of pattern and emergence. The chapter then moves on to explore the research opportunities offered by combining studies of different sites in the research of Internet-mediated phenomena, and it concludes by looking specifically at the benefits and challenges of combining data derived from both online and offline sites within a single study.

Methodological paradigms of Internet research?

I have argued previously (Hine, 2005) that Internet research was characterized, almost from the outset, by a mood of methodological innovation: simply because the Internet seemed so new, and because it appeared to occasion new ways of being present in the world and interacting with one another, it appeared almost self-evident that new approaches would be required in order to grasp these new phenomena. Discussions of the challenges of researching Internet phenomena at this time were thus not focused simply on the development of some new techniques or the emergence of new methods, but touched on a methodological level of debate too. This methodological level of debate involved considering what the conditions for production of robust knowledge about Internet phenomena might be, given the apparently fundamental uncertainties about such issues as the identity and location of research subjects and the appropriate units of analysis to use for the phenomena that researchers were encountering. In fact, as social science research progressively came to grips with the Internet this mood of self-evident novelty gradually dissipated, and many more continuities with previous modes of existence than had been anticipated came to light. However, a legacy of interest in the Internet as ripe field for discussions about methods and methodology remains.

One characteristic of Internet research which has endured from the earlier days is a consciousness of the limitations of relying on single approaches and a willingness to explore various combinations of methods appropriate to the job at hand. It may be that the relatively recent emergence of the Internet as a mainstream phenomenon of interest to social science means that it was not, from the outset, locked into one particular methodological paradigm (where a methodological paradigm implies a particular set of views on the nature of the phenomenon to be

studied, how questions should be formulated and what methods should be used for collecting and analysing data). In contrast to some other fields, where distinct methodological paradigms have come into sharp contrast and even competition with one another, in Internet research the distinctions have been less starkly drawn. Internet research is young enough to have largely avoided the “paradigm wars” experienced in other substantive domains (Bryman, 2006). As a field, Internet research has also tended to be multidisciplinary rather than constituting itself as a single discipline with an agreed set of central theoretical premises or common discourse (Sterne 2005). That there is a common meeting ground for scholars interested in the Internet at all is due in no small part to The Association of Internet Researchers, which held its inaugural meeting in 2000 and has met annually since, continuing throughout to welcome researchers from diverse disciplines. These Internet researchers come from departments of sociology, anthropology, information science, library studies, communications, psychology, geography, politics and many more (Baym 2005), and this disciplinary diversity may have contributed to inter-visibility and mutual respect among researchers espousing different methodological approaches to understanding the Internet. At early conferences of the Association of Internet Researchers there was, for example, no clear distinction made between Internet researchers who espoused qualitative approaches and those who sought quantitative understandings of online phenomena.

This methodological diversity also means, however, that different communities of scholars come into contact around studies of the Internet but are not necessarily talking about precisely the same object. There have been some fundamental ontological complexities in determining what the Internet is, which affect how we define it as an object of research. These complexities arise, for example, when deciding whether to separate out

the online sphere from other aspects of everyday life for research purposes, and whether it is appropriate to focus on “the Internet” as an object as opposed to a specific site such as Facebook, or even a particular user’s experience of Facebook. In addition, because the Internet has been identified as the territory of many different disciplines at once, there is a corresponding complexity in methodological approaches. One researcher might see the Internet as a site for ethnographic exploration, while another might view it not as a location in its own right, but as a source of documents for textual analysis (Hine 2000).

The diversity in definitions of the Internet as a research object has not however necessarily led researchers to see different methodological approaches as incommensurable paradigms. Schneider and Foot (2004), for example, pointed out that the World Wide Web as an object of research had been constituted in multiple ways. Among researchers who shared a belief that there are phenomena of social interest to study on the web, they distinguished: those who took a discursive or rhetorical approach to understanding website content; those who undertook analysis of the structures and features of individual websites, or mapped patterns of linkage between sites; and those who undertook what they described as sociocultural analysis of the web, exploring the circumstances of production and maintenance of websites as a context for understanding website content and linking. Having identified these three quite different formulations of the web as a research object, Schneider and Foot (2004) positioned their own approach as an overtly multi-method approach which encompassed aspects of all three. Their “web sphere analysis” involved saving a collection of websites related to a chosen topic or theme, capturing not only the sites themselves but also the patterns of linkage between the sites. Changes in sites over time might also be captured in this archive. The analysis then deploys a combination of

computer-assisted and manual annotation of features of the web sphere, which is triangulated with accounts given by producers and users of the sites.

Scolari (2009) describes a field of cyberculturalⁱ studies that has, over time, encompassed empirical research, philosophical speculations, journalistic analysis, apocalyptic visions, optimistic forecasts, literary criticism, and cyberpunk literature. As Schneider and Foot (2004) exemplify, Internet research represents an area in which researchers from diverse disciplines are exposed to one another's theoretical and methodological resources, and become conscious of different ways of constituting the Internet as a research object. It seems little surprise then that the field has been methodologically diverse and disposed towards mixed method research designs. As Peng et al. (2012, 15) describe it: "Internet studies are a melting pot that attracts researchers from different disciplines to transcend their disciplinary boundaries to develop new theoretical, methodological, and practical concerns." Peng et al. (2012) based their claims about the nature of the field of Internet research on analysis of a corpus of journal articles published in English between 2000 and 2009, retrieved for their use of a set of indicative keywords: Internet, web, cyberspace, cyber-space, online and on-line. The resulting corpus was subjected to cluster analysis and text-mining for keywords. The cluster analysis demonstrated that the field spanned the social sciences and clustered most readily into substantive themes such as e-health, e-business, e-society and human-technology interactions. The methods used in these areas of research were explored by manually coding a sample according to the type of method. Peng et al. (2012) found that Internet research in the broadest sense had been dominated by quantitative methods in the time period that they studied. They did not specifically look for mixed methods, but they found that the research cluster of e-

society was particularly methodologically heterogeneous, suggesting a mutual visibility of different methodological approaches at the very least. Internet studies appear to be an area within which commonality of substantive concerns coupled with an inherent multidisciplinary may promote awareness of, and respect for different methodological traditions.

This mutual visibility of different disciplinary perspectives and different methodological traditions, clustered through their mutual concern with substantive areas focused upon the Internet, has produced fields of literature, which form a methodological mosaic of studies looking at related research objects through different methodological traditions. Each method is seen as having its own strengths and weaknesses, and focuses the object of research in a different way: Table 1 outlines some of the key methods which have become established in Internet studies, outlining the insights which they offer and the potential limitations associated with each. The awareness of overlapping and related, if not identical research objects does not in itself, of course, guarantee that individual studies will contain a mixing of methods, even though it potentially encourages learning about methods across disciplinary divides and might position researchers to experiment with approaches not usually within the methodological remit of their home disciplines. Combining of methods within individual studies was, however, also an early phenomenon, as researchers struggled to find effective ways of grasping the Internet as a social phenomenon. These mixed methods approaches particularly flourished within Internet ethnographies. The next section looks specifically at the mixed methods tradition in Internet ethnography, exploring the factors which drew ethnographers towards combinations of qualitative and quantitative analysis.

[Insert Hine-Table 1 here]

Internet ethnography as a mixed method approach

The claim that the Internet provided a field site for the study of culture using ethnographic techniques was made relatively early on in the mainstreaming of the Internet (Hine 2005). Baym's (1995; 2000) work on a rec.arts.tv.soaps newsgroupⁱⁱ was highly significant for the field in demonstrating that the Internet could be a place where rich social formations emerged, and as a corollary, that it was appropriate to set out to understand these social formations through the combination of participation, observation and systematic exploration that characterizes ethnography. Subsequently, the notion of Internet ethnography has expanded and diversified (for discussions see Markham 1998; Hine 2000; Miller and Slater 2000; Boellstorff 2008; Hine 2008a; Murthy 2008; Garcia et al. 2009; Kozinets 2009; Nardi 2010). Authors vary on the extent to which they see the online domain as a discrete field amenable to ethnographic study in its own right (Boellstorff 2010), with Miller and Slater (2000) having early on questioned the extent to which the "virtual" should be taken to be an objectively existing discrete domain of experience. Hine (2000) stressed that the Internet could be construed as both culture and cultural artefact, and noted the significance of offline circuits of interpretation for shaping online experiences. Markham (1998) noted the multiplicity of the Internet experience, specifically arguing that it could be encountered as a tool, place, or a way of being, and that each entailed a different orientation from the ethnographer. Both across and within ethnographic studies of the Internet from the outset there has been recognition of the ontological multiplicity (Mol 2002) of the Internet, such that it is experienced and defined quite differently from different perspectives. Since ethnographers tend to strive for a rich and multi-faceted

understanding of the phenomenon they study it is therefore not surprising that multiple methods should have been employed to develop an ethnographic understanding of the Internet, and that in many instances researchers have combined these multiple approaches within a single study.

Baym's pioneering Internet ethnography (1995; 2000), which has acted as a model for many subsequent studies, employed a combined methodological approach. Baym (2000) recognised three senses in which the group that she studied was a community: as a discrete Internet newsgroup; as an audience community oriented to a specific soap opera; and as a community of practice. Her understanding of these multiple community formations was itself developed through multiple forms of engagement with the group. She was a long-standing member of the group (and viewer of the television program that they discussed), and much of her initial insight was developed through a deep understanding of how participation was structured and how it felt to participate. She deployed a variety of data collection techniques, including archiving of messages themselves, interviewing participants, attending offline events when they occurred, and surveying members to collect background demographic information. She conducted a variety of modes of analysis on samples of messages, including structured coding of message genres and a content analysis of agreements and disagreements. Baym (2000) was thus able to combine an impressionistic, rich, and situated account of participation with a structured account of patterns emergent across the group. Just as many qualitative researchers recognize the necessity of providing quantification as a mode of description (Maxwell 2010), Internet ethnographers have found that they can quantify, and that developing structured, content analytic descriptions can complement more impressionistic approaches.

The urge to quantify, for an Internet ethnographer, can arise because so much data presents itself quite readily in a form that the researcher can search, copy and manipulate. Conversely, if one does not carry out some form of quantitative analysis or data visualization as an online ethnographer, the sheer amount of data that is available can come to seem overwhelming. For example, Boellstorff (2008) describes his ethnography of Second Lifeⁱⁱⁱ as producing over ten thousand pages of field notes based on participant observation, interviews and focus groups together with another ten thousand pages of documents such as websites, blogs and newsletters. He explains that for an ethnographer in an offline setting producing field notes and collecting data are often hard work, and a lot of filtering therefore happens at the stage of deciding what to write down, and which data to focus on acquiring. Online data can be much more ready-to-hand. Traditionally an ethnographer in an offline setting might have to laboriously conduct a house-to-house survey in order to be able to describe a population and to formulate quantitative statements about the nature of relationships there to complement qualitative descriptions based on participant observation. In an online world surveys can still be conducted (by sending messages rather than trudging from house to house), but the ethnographer can also develop an overview of the participants and their activities by analysing ready-to-hand data in archives of messages, membership lists or log files recording details of interaction.

Many forms of online interaction are already archived for the use of participants themselves quite independently of the activities of the ethnographer, so an ethnographer can not only experience real-time interactions as a participant observer, but also explore the precedents and

antecedents of the activities they are experiencing now through the archives that record the history of the group. For example, whilst studying an online discussion group of biologists I found that it seemed a quite natural step to combine insights based on participant observation and interviews with participants with a quantitative analysis of the group membership using membership records given to me by the list owner and an analysis of the temporal patterning of participation using data taken directly from the archive of messages (Hine, 2008). These descriptive statistics gave me a context in which to situate my own impressions of the dynamics of the group and the accounts of other participants. I also found that looking at the archives gave me the opportunity to rapidly develop a sense of the history of the group, and to situate activities I was observing in the present against the past of the group. It was possible to travel back and forth in time through the archive, finding that activities currently taken for granted as normal, such as asking for identification of a plant by sharing an image of it with other participants in the group, were once deemed remarkable, even risible. The archive makes certain forms of temporal mobility possible for ethnographers, as well as encouraging use of descriptive statistics to complement insights from participant observation.

Ethnography is often an inherently mixed method in its attempts to capture multiple dimensions of its substantive focus, and Internet ethnography has continued and developed this tradition, with extra fervour thanks to the particular qualities of the phenomenon being studied. In particular, so much of the data that Internet ethnographers encounter is “born digital” and thus relatively readily rendered as a corpus for structured analysis, that quantitative forms of enquiry become accessible even for single researchers or small teams. Ethnography online thus builds on the existing mixed methods tradition of ethnography and raises it to a new level:

As a form of ethnography, netnography encompasses multiple methods, approaches and analytic techniques. Experienced ethnographers are inevitably bricoleurs, and I would fully expect netnographers to deploy any of a number of techniques as they approach the rich lived worlds of cultural experience that people share and experience online. These techniques would include projective techniques, historical analysis, semiotic analysis, visual analysis, musical analysis, survey work, content analysis, kinesics and any of a world of specialities, as well as the more obvious observational, participative and interview techniques. (Kozinets 2007, 132)

Ethnography has been used here as an example of an approach which is already pre-disposed to being adaptive in its choice of methods, and where mixed methods research has particularly flourished in Internet research. However, the mixing of modes and methods is of course not confined to Internet ethnography. The experimental tradition has also thrived in Internet contexts, and here again combining of approaches has been found beneficial. Reips and Buffardi (2012), for example, discuss various web-based approaches for studying migrants using psychological techniques, and explore the purchase offered by combining different approaches. These approaches include Internet –based interviews and surveys, nonreactive data collection of existing documents and traces of interaction such as from social networking sites, Internet-based tests and online experiments. Each approach carries its own challenges, affordances, and limitations, but each is seen to offer a form of knowledge that remains unavailable from other perspectives.

Ess and Dutton (2013) talk about mixed methods designs in Internet research as a response to the weakness of particular methods, and to some extent this is apparent in the motivation of Internet researchers towards mixed methods designs, but this formulation possibly underplays the extent to which the methodological imagination has been stimulated in a positive sense by the particular complexities of the Internet as a

research object, the interdisciplinary nature of Internet research as a field and the kind of data brought into being by use of the Internet. It is not entirely accurate to describe the mixing of methods in Internet research as wholly a response to the perceived inadequacies of singular approaches. Mixing of methods in Internet research has been as much occasioned by a mood of exploring new possibilities and accepting the opportunities offered by the wealth of data the Internet makes available, as it has by a sense of weakness or lack in any particular method. The next section illustrates this sense of the opportunities of mixed methods designs in Internet research by focusing in further on some of the unusual qualities of the Internet as a source of data for doing social science. This section looks particularly at the affordances of “born digital” data for large scale analysis and explores in more detail some of the methodological responses this kind of data has prompted.

Combining large scale and small scale approaches to Internet research

Whilst discussing mixed methods research designs within Internet studies, it is crucial to recognize that this is not simply a question of mixing together pre-existing methods and approaches. The distinctive qualities of the Internet, as a vast, interconnected searchable territory, generated through the activities of myriad uniquely situated participants offers up possibilities for study not anticipated within previous methods. Rogers (2013) argues that the advent of this “born digital” data requires not only the transfer of traditional methods to the new domain, but also the development of digital methods, which offer a new way of conceptualizing what society is, and are particularly attuned to the qualities of this new domain. For example, analysis of the links between websites offers new possibilities for exploring the emergence of issues and the development of political debates (see for example Rogers and Marres 2000). The patterns of hyperlinking^{iv}, in this approach, are taken to

constitute a form of novel social connection in their own right, the issue network, rather than simply being artefacts reflective of an underlying social structure that could be uncovered by conventional techniques. The study of the issue network is thus construed as offering a distinctive window on contemporary social existence. Given its claims to offer such a distinctive perspective, it is perhaps no surprise that this kind of technique has been employed within mixed methods designs. For example, this technique was used in a study of political party organization and activism, which explored political blogs through a combination of interviews, content analysis, surveys of users and data on usage, and hyperlink analysis to explore the emergent networks of connection (Gibson et al. 2013).

Whilst the Internet has occasioned some altogether new methods, such as issue network analysis, it has also re-invigorated some pre-existing methods. Social network analysis, and related techniques for visualizing social structure on a large scale have thrived in the face of the visibility of social connections, which the Internet offers (Garton et al. 1997; Hogan 2008). A social network analysis aims to understand the social world as a network of connected entities, for example treating individual people as nodes in a network and various forms of relationship, such as friendship, kinship or working relations as the ties between the nodes. This form of analysis long pre-dates the Internet, but data collection was often very laborious because people had to be asked to recount in exhaustive detail the relevant relationships in which they were engaged. The advent of the Internet, with persistent data on many forms of relationship and interaction, has made large scale approaches such as social network analysis a much easier prospect. Provided that the relationship one wishes to study is reflected in some way in online interactions,

archived data from those interactions can be used as a proxy to portray the nature and strength of those relationships and enable the social network to be visualized.

One form of large scale analysis fostered by the Internet focuses on depicting relationships by using the patterns of connection between websites or the amount of online interactions between individuals as a proxy for the strength of the relationship. Initially web sites were often produced by organizations, and researchers were able to study the various forms of connection between them depicted by the patterns of linking in cyberspace^v. Subsequently, the advent of large-scale user-generated content (the participatory web, or web 2.0 (Blank and Reisdorf 2012)) offered up for analysis a dramatic increased array of traces of the daily activities and connections of large swathes of the population. Blogging, social networking sites such as Facebook and micro-blogging (for example Twitter) allow people to express their thoughts on a public stage, and as a by-product make those thoughts available for social researchers to mine for patterns. The visibility and persistence of digital data makes it possible to map out networks and explore structure on a grand scale, removing the previously often insurmountable barrier to large-scale analysis presented by the lack of resources to collect the necessary data.

In Internet studies it is clear that many novel research approaches have been occasioned by the distinctive qualities of the data, particularly its persistence, searchability and ready accessibility. The possibility of generating large data sets for relatively little cost, whether through collection of publicly available traces such as hyperlinks or social networking site activities, or via negotiation of access to institutionally held log files and

message corpora, makes large scale quantitative analysis more feasible than it has been before. Small teams of researchers, with relatively little resource, can now access large data sets and, with appropriate computational support, make sense of them. The advent of publicly available tools specifically designed to make sense of these large corpora, such as the Issue Crawler (<http://www.govcom.org/>)^{vi}, SocSciBot (<http://socscibot.wlv.ac.uk/>)^{vii} and SentiStrength (<http://sentistrength.wlv.ac.uk/>)^{viii}, means that each researcher does not necessarily have to design their own analytic tools.

Whilst this form of large scale analysis is undoubtedly powerful in its sheer scale and ability to see a “big picture”, it has also been subject to criticism for the lack of fine detail and ability to interpret findings for what they mean to participants. Large scale analysis of “found data” from the Internet has often, therefore, featured in mixed methods research designs, that exploit the ready availability of large scale data but also seek other means to assist in interpreting the resulting patterns. In the rest of this section I outline some different examples of mixed method designs which combine a large scale analysis of publicly available data with other forms of contextualising data. First, I outline the conditions under which digital data becomes available for research purposes and offer some diverse examples of large and small scale approaches to analysis of found data from social networking sites, blogs and Twitter, before turning to a more in-depth exploration of the motivations for combining methods and the methodological assumptions behind different forms of combination.

The most obvious motivation for conducting large scale analysis based on “born digital” data is prompted by the simple fact that this kind of data is readily available. Almost every conceivable facet of human activity is reflected in some way on the Internet (albeit often in a somewhat selective or partial way). Hyperlinks on the web are generally publicly available data, as are blogs by their nature, and as such can be accessible for use by researchers without the need to negotiate additional access. Social networking sites offer up a rich array of “found data” for quantitative analysis, provided the individual users’ privacy settings render the data visible to researchers. For example, Thelwall (2008) inventively deployed public MySpace comments to conduct a study of contemporary swearing and cursing practices. The study utilized the tendency of many users not to restrict access to their profiles only to their known friends. Some forms of data, by contrast, are amenable to large scale analysis but can be harder to come by. For example, access to institutional intranets, or log files of internal communications such as used by Dirksen et al. (2010) requires negotiation with stakeholders. Whilst the Internet does promote the use of “big data” approaches as a complement to qualitative studies, then, it does depend on the individual research object whether this big data is readily accessible or not. However in many cases, the research imagination is piqued by the actual or potential availability of a set of data that before the Internet could not have existed on such a scale.

Provided that the data is publicly available, the same kind of data can often in fact be viewed both as fodder for a large scale analysis and as a potential site for a detailed qualitative enquiry. Blogs offer up data which can be subjected to both large-scale qualitative and more in-depth quantitative approaches: Herring (2010) discusses the challenge that blogging poses to conventional forms of content analysis and suggests that

researchers can usefully combine different approaches in order to attend to the various qualities of blogs as meaningful contexts in their own right and as part of structured and inter-connected fields of discourse. For example, Herring et al. (2005) carried out a study of blogging, which combined a qualitative analysis of the comments that were made on inter-connected blogs with a social network analysis and visualization that allowed the researchers to explore the emergent patterns in the blogosphere. Efimova (2009) carried out a qualitative study of blogging practices, in which she complemented her ethnographic observations with a systematic visualization of blog conversations as they emerged over time. The temporal dimension can be particularly significant in analysis of blogs: Park and Kluver (2009), for example, conducted a longitudinal analysis of the blogs of politicians in South Korea, involving analysis of hyperlinking over time together with interviews.

Twitter also offers up rich data for both large-scale approaches and more targeted studies. Murthy (2013) presents a commentary on Twitter, which combines both qualitative and quantitative approaches to understanding its role in mediating disasters, promoting activism, and enacting health concerns. Bruns (see for example Bruns 2011; Bruns and Burgess 2011; Bruns and Liang 2012; Bruns and Stieglitz 2012) has developed quantitative approaches to characterizing and comparing the emergence of debates on Twitter, as a complement to case study and anecdotal approaches. Bruns and Liang (2012) discuss the methodological diversity within quantitative Twitter studies, which can deploy a mixture of re-purposing of existing tools and custom-designed analytic approaches. Proctor et al. (2013) discuss the mixed method design, which they developed to analyze the corpus of tweets relating to the riots, which took place in England in 2011, provided directly to the research team by the Twitter organization. Here the large-scale computational analysis of the substantial corpus of tweets (2.6 million tweets from 700,000 distinct

user accounts) was used to direct the researchers' attention to potentially interesting sections of the corpus for an in-depth qualitative exploration. This combination of methods allowed the researchers to handle the corpus on a range of different scales: as a corpus, as information flows, and as individual communicative events.

Various forms of network analysis and visualization described here have been deployed by researchers interested in online phenomena to orient their studies and add different perspectives, conscious of the affordances that different methods offer. The research designs within which these methods are combined can however differ widely, as can the understanding of their affordances. Designs can be built on a very specific notion of how methods complement one another. For example, like Proctor et al. (2013), Howard (2002), used a sequential design, deploying social network analysis of the patterns of communication across a distributed organization to orient his choice of case study sites for an in-depth exploration. As Howard describes it, social network analysis can act as a guide to conducting the kind of purposive sampling that an ethnographer will often deploy to home in on a potentially interesting site to study. It is, however, overly limited as the sole portrayal of a field, particularly with regard to temporal dimensions:

On its own, social network analysis misses much of the rich information that the researcher can obtain by participating in the hypermedia organization and observing small-group interaction. As a method it can bring perspective to complex-layered social networks, sometimes artificially making employment, peer, and personal networks congruent. More importantly, social network analysis has limited use in revealing stories of mobility within communities. Narratives about how people enter and leave a network, or about how people move from periphery to core and back, are difficult to reduce to comparative values. (Howard 2002, 560)

As he points out, most social network analysis involves some form of on-the-ground verification. However, the combination of a formal social network analysis followed by a full ethnographic exploration of selected field sites informed by the social network analysis that Howard (2002) uses is distinct in the more equal weight it gives to both methods as different modes of understanding. As he portrays it, the network analysis gave him a comprehensive overview of the territory, but in a rather static format, whilst the subsequent participant observation offered a more processual understanding.

In similar vein, De Maeyer (2012) conducts a critical analysis of hyperlink studies and finds that many authors combine an automated or large scale link analysis with some form of contextualization or qualitative exploration to allow the meaning of the links being analyzed to be inferred. In itself, a hyperlink, or a Twitter hashtag, or a “like” on a social networking site does not tell us what meaningful social connection is being forged. De Maeyer stresses the importance of supplementing link analysis with other methods to allow for both patterns and meanings to be explored. Such combinations of link analysis to study structure and qualitative methods to explore process, depth and meaning have been powerful approaches in Internet studies. Vasileiadou and van den Besselaar (2006) discuss a mixed method design for studying collaborations between organizations in the field of communication motivated by these concerns, combining an analysis of web links with qualitative interviews. Brügger (2013) discusses similar issues in the context of historical analysis of web archives. He argues that “a hyperlink is always a semantic, a formal, and a physically performative entity”, and whilst it might be informative, a network analysis of hyperlinks only explores the physically performative entity (Brügger 2013, 308). He argues that alternative approaches based on textual analysis are needed in order to

explore other aspects of the hyperlink. For a historian, Brügger explains, the challenge of working out just what a hyperlink might signify is compounded by the imperfect and selective rendering of the web in archived form, and the multiple ways of rendering the web as an object of study, such that a singular methodological approach is often insufficient.

To some extent, then, researchers combining large scale quantitative analysis with smaller scale qualitative approaches to online phenomena have done so because one or other approach on its own seems to lack explanatory power, or, to put it more positively, because the different methods offer complementary and distinctive ways of understanding the phenomenon. These researchers combining large scale analysis with qualitative methods tend to see the qualitative method in general as adding depth or meaning, but not all necessarily see the combination as a process of triangulation on a singular research object. Howard's (2002) sequential understanding of the mixing of methods in online ethnographic work with social network analysis as the preliminary step is only one of the possible models for understanding how the different methods combine. Dirksen et al. (2010) describe a somewhat different model in their study, which used social network analysis of log files from a company Internet to enable them both to interpret qualitative data and to target their detailed attention to various online and offline sites. Dirksen et al. describe their approach as "piling on layers of understanding" (2010, 1045), positioning the different methodological approaches as layers which contribute to development of a picture of the overall phenomenon of the dynamics of online social practices within the organization. Dirksen et al. describe their approach as a connective one, moving between sites to explore meaning-making as linked by various discourses and forms of technologically-mediated practice. Their approach to combining methods in this way is inspired by Strathern's (2002,

303) comment that different methods add “layers of understanding”, not necessarily triangulating with one another or providing a straightforward contextualization, since each method offers up a new viewpoint and a new version of what a context might be, effectively creating a new research object. A researcher might therefore make use of network analysis or visualizations alongside qualitative analysis in order to provide distinctive perspectives on a field without necessarily accepting that one straightforwardly contextualized the other.

Researchers even vary in the extent to which they see different forms of analysis as separable and distinct. Markham (2012) folds the role of network analysis into ethnographic research, describing network analysis as a means of developing “network sensibilities,” explaining “the reflexive, ethical power of using network analysis to help shift our perspective constantly and radically, by attending to different senses we might use to locate, interpret, and represent that which we might call ‘data’” (Markham 2012, 48). She describes network analysis as an enhancement to ethnographic work, neither necessarily preceding qualitative and participatory work nor depending upon it, but both become part of the ways in which the researcher senses and moves through the field. Geiger and Ribes (2011) take a related stance in their discussion of “trace ethnography”, which takes seriously the role of various forms of document and trace in rendering participants present to one another, as well as to the researcher. Their aim is to integrate techniques, taking in both the richness of participant observation and the decoding of various documentary traces: an approach which they demonstrate using the example of the log file traces of edits to Wikipedia pages. They stress that the analysis of traces does not precede or inform the ethnography, but to a large extent is the ethnography, since this is the means through which participants are visible to one another.

As these latter examples from Dirksen et al. (2010), Markham (2012) and Geiger and Ribes (2011) show, some Internet researchers from largely qualitative traditions may thus have adopted larger scale approaches, but they have done so because these are seen as appropriate adaptations to the circumstances they find rather than because these large scale approaches are seen as somehow inherently better than small scale qualitative research. boyd and Crawford (2011) argue that there are concerns that “big data” approaches may come to dominate the research landscape as appearing inherently more informative, and this “scaling up” may instead entail loss as well as gain in terms of our understanding of social phenomena. They stress that different forms of data are not necessarily interchangeable; for example, a social network analysis based upon people’s use of social networking sites does not necessarily tell us about their social networks as they would be represented by other forms of data. They formulate a distinction between the articulated networks that people specify through their use of social networking sites, the behavioural networks of the connections people forge through the various forms of communication available to them, and the personal networks which entail a more subtle understanding of the value people place on their relationships with other people. In their caveats about the use of big data, boyd and Crawford (2011) remind us that data that tells us about one of these forms of network may well not tell us about the others. They offer a principled argument for a form of mixing which remains attentive to the different claims that different forms of data make about the significant phenomena for our research attention. Thus stance appears consistent with the line taken by ethnographers such as Markham (2012), Geiger and Ribes (2011) and Dirksen et al. (2010).

Thus, even while many researchers have combined various forms of quantitative and network analysis with qualitative approaches to the Internet, then, their conceptions of the necessary link between the components of the study afforded by different methods can differ dramatically. This section has included designs which deploy methods in sequential fashion as one orients and directs the next, designs which use different methods to offer different perspectives on a phenomenon without claiming that either explains or contextualizes the other, and designs which use multiple methods precisely in order to provide contextualization. Internet researchers are very conscious of being overwhelmed by the sheer quantity of data available, and in this context the possibility of new ways of exploring and visualizing patterns holds considerable promise. These new approaches do not necessarily displace other forms of analysis, nor do large-scale “big data” approaches make small-scale qualitative research redundant. Instead, many researchers are discovering the explanatory power of combined approaches, which look both for patterns and for meanings. Rather than triangulating on a singular object, however, many researchers in this domain are highly conscious of the complex and multiple nature of the Internet as a research object, and see different methodological approaches as complementary but not necessarily cumulative in the understanding that they offer.

Multi-modal research on and off the Internet and the challenges of triangulation

A further aspect of the complex and multiple nature of the Internet as a research object which has troubled researchers and influenced research designs concerns the extent to which the Internet is seen to be understandable through data collected solely online. Because the Internet is as much a component of everyday life offline as it is a discrete online space, many studies of Internet phenomena do not collect their data solely

through the medium of the Internet itself. These studies often mix both methods and modes in their attempts to gain a rounded perspective on a phenomenon, which both understands the Internet for its own qualities and explores its embedding in diverse circumstances offline (Wellman and Haythornthwaite 2002). Baym (2009) argues for multimodal approaches to understanding new media, stressing that these new technologies are not experienced in isolation, and that our practices connect them in complex and unpredictable fashion. This has serious implications for the design of adequate research approaches for understanding new media:

We cannot bank our research future on the technological forms. Instead we need to interrogate the underlying dynamics through which technology use is patterned across media, relationships, and communicative purposes and with what effects for how we understand and conduct our relations, our communities, and ourselves. (Baym 2009)

Where our concern is to find out what the Internet means to people, it is often difficult to render a study that focuses only online as adequate to describe the phenomenon. At the same time, a study of what the Internet means to people, which focuses only on face-to-face interviews and ignores what they actually do when they are online, despite the multiple traces that these activities leave, would seem somewhat perverse. Studies of the embedding of the Internet in everyday life thus find themselves almost inevitably deploying a mixture of methods and modes to explore their key concern with understanding the way that online and offline are entwined. There are, indeed, many studies that could be cited here, so the following represent only a very selective sample for illustrative purposes, focusing on studies in which the combination of methods notably allows for an expansion in the scope of the explanatory or descriptive possibility of the study.

One such example of a mixed mode, mixed method study is boyd's (2007; 2008) exploration of teenage use of social networking sites. Her work combined online fieldwork in social networking sites themselves, offline interviews, and observation in less structured and less predictable environments where teenagers hang out. Through this combination of modes and methods boyd was able to position her study as an exploration of how teenagers lived their lives in and through social networking sites and at the same time how this activity was rendered salient in various other aspects of their lives. In another example of research aimed at finding out what people make of the Internet, Hargittai (2002) has conducted studies of the ways in which people find and interpret online materials, deploying mixed modes and methods, including structured observation sessions based upon participants completing an online task together with more open interviews and larger scale surveys, often conducted on a longitudinal basis. Livingstone et al.'s study of The Class (<http://clrn.dmlhub.net/content/class>) focuses on exploring the networks of one school class of 13-14 year olds, combining observations within and between lessons; interviews with children, parents, and teachers; and mapping of engagements with digital technologies. Hampton has conducted various studies of the embedding of the Internet in everyday life, focusing on multi-method and multi-mode data collection to explore such issues as the role of Internet communications in neighbourhood community relations (for example Hampton 2007). One notable study combined a national telephone survey of Facebook users with analysis of their Facebook profiles and log files of their activity on the site, to allow for a triangulation between what people said about the site and what they did (Hampton et al. 2012). These mixed mode, mixed method studies have in common that they seek both to understand in detail what people do online but also to explore what those online activities mean within people's everyday lives. This kind of study must, perforce, move around and adopt different perspectives in order to capture these different aspects of Internet activity. Similarly, connective and multi-sited approaches to

ethnography have developed through the ethnographer's urge to "follow the thing" as it is manifested in diverse locations (Marcus 1995), and increasingly this includes both online and offline locations (Hine 2007; Burrell 2009).

A different kind of mixed mode study is a rather more pragmatic response to the affordances of the Internet for conducting research interactions with participants. For example, the mixing of modes in interview-based studies is now relatively commonplace, in part due to pressure from research participants, some of whom will feel comfortable with one or other form of online interview, and others of whom will strongly prefer face-to-face interviews. The qualitative researcher's interest in finding the medium for interviews that will put participants most at ease provides an impetus towards the combination of different modes of interview in the same study, as does the urge to maximize participation (Kazmer and Xie 2008). This set of concerns provided the impetus for a methodologically plural study of the role of the Internet in gay men's sexual practice, which recruited from online and offline sites and used both face-to-face and online interviews (Elford et al. 2004). Similarly, Malta (2012) studied the online dating experiences of older adults, and deployed interviews conducted face-to-face, by email, by instant messaging and by telephone, according to the choice of the participant.

Another set of pragmatic motivations for mixing modes in Internet research occurs among survey-based studies. According to Dillman (2011) survey approaches mix modes for a variety of reasons, including: collecting the same data from different members of a sample; collecting panel data from the same respondent at a different time; collecting different data from a respondent at one time; collecting comparison data from

different populations; and using one mode to prompt a respondent to complete a survey in another mode. Each occurs for different pragmatic motivations, and introduces distinctive concerns for the interpretation of measurement differences across modes. There are therefore quite different motivations for mixing modes in Internet research studies, some of which are based in quite fundamental ways on a particular conception of the research object, and others of which are rather more pragmatic in their orientation to the most effective or efficient ways to collect interview or survey data.

In some methodologically plural and mixed mode studies, it appears relatively straightforward to view different approaches as triangulating on the same research object. One might study an online community for example, through a mixture of approaches including observation, interviews online or offline, and online surveys (as in the studies by Williams et al. (2006) and Nip (2004b; 2004a) described by Hesse-Biber and Griffin (2013)) and still be relatively straightforwardly able to claim that one was studying the same community. Temporal fluctuations in the membership of the community and sampling issues can however mean that the research object is not completely the same as viewed through each approach. Even in the case of a single context study, it is not necessarily straightforward to determine a single research object, since participants may come and go over the course of a study, or different approaches draw in different sub-sections of the membership, such as active posters or lurkers. In some mixed mode studies which combine online and offline but do not follow a distinct sample from one site to another, it is markedly less straightforward to identify a single research object on which different methods triangulate. For example, boyd's (2007; 2008) study of teenage use of social networking sites used a variety of methods, but unlike the study by Hampton et al. (2012) described

above, did not follow the same teenagers between online and offline because of ethical concerns. Thus boyd could be said to have studied teenage use of social networking sites but not as experienced by a specific sample of teenagers across online and offline. Mixed mode studies, which combine interviews online and offline can also have difficulties in aligning the two forms of data. Orgad's (2005b; 2005a) study of the use of online forums by women experiencing breast cancer moved between online and offline but found that the two modes of interviewing did not necessarily straightforwardly align: the two modes provided two different views of the phenomenon, even where the same woman was encountered both online and offline. Orgad found that online and offline interviews were not straightforwardly windows on the same phenomenon, nor was either inherently more informative.

Conclusion

It appears that mixed methods research designs are occurring relatively easily to the methodological imagination of Internet researchers. These designs have flourished in part due to the prevailing interdisciplinary tradition within Internet research, and the relatively recent emergence of the Internet itself, within an era more accepting of mixed methods research than may be the case for more established disciplines. Adaptive approaches such as ethnography have readily embraced the possibilities for combining qualitative and quantitative designs on the Internet. Even where there is a pre-existing commitment to qualitative research and its related philosophical paradigms, the sheer amount of data which comes readily to hand thanks to Internet search tools, log files and the like often prompts the adoption of a quantitative stance, or at least some means of

aggregating, visualizing, and exploring patterns. In adopting this stance, indeed, Internet researchers are aligned with developments in the world around them, which is increasingly oriented to the potential offered by “big data” (Savage and Burrows 2007).

Internet researchers are often in the position described by Johnson et al. “when the nexus of contingencies in a situation, in relation to one’s research question(s), suggests that mixed methods research is likely to provide superior research findings and outcomes” (2007, 129).

Particularly, the complex, fragmented and embedded nature of the contemporary Internet means that it often feels inadequate to rely on a single source of data, and instead methods are combined and data from different sources examined in order to shed alternative sources of light on facets of a problem (Mason 2011) if not to triangulate on a singular answer. Different methods and modes intersect in the illumination of a phenomenon, but not necessarily in an additive sense in which each simply contributes “more” to understanding. Sometimes different forms of data and different modes of enquiry will align to add to certainty, but such is the complexity, interpretive flexibility and multiple embedding of the Internet that this is rarely the case. More often the combination of modes and methods in the study of the Internet adds another form of understanding to our grasp of a complex phenomenon without triangulating on a singular solution.

Working against the methodological innovation and mixing promoted by a culture accepting of methodological innovation, a ready availability of data and an emergent set of tools, however, is the volatility of the Internet itself. The rapid development of Internet technologies and the constant emergence of new forms of data that has been experienced in recent years make it a considerable challenge for researchers to maintain

the necessary skills to access and make sense of online phenomena and for developers of tools to maintain the currency of their technologies for the contemporary situation. Mixing of methods in Internet studies brings with it challenges as well as promises, and not least of these is accessing the range of skills and tools needed to combine qualitative and quantitative perspectives on an ever-shifting array of phenomena.

Discussion questions

How is a large scale social network analysis of an online community compatible with an ethnographic approach? How, and when, might the ethnographer best make use of the network analysis?

Why might a researcher choose to conduct both online and face-to-face interviews? What problems of interpretation might arise when the researcher tries to analyse the two sets of interviews side-by-side?

Is it appropriate to talk about mixing methods in Internet research as a form of triangulation? What other motivations do Internet researchers have for mixing methods?

Websites for further reading

Association of Internet Researchers <http://aoir.org/> The Association of Internet Researchers is an interdisciplinary scholarly network and membership organization which organizes regular conferences, hosts a lively discussion list and publishes ethical guidelines for Internet research.

Digital Methods Initiative <https://www.digitalmethods.net/Digitalmethods/WebHome> This site originating from Richard Roger's research group at University of Amsterdam contains training materials on a range of approaches to social analysis of digital data.

Introduction to Webometrics <http://webometrics.wlv.ac.uk/> This site by Mike Thelwall contains material related to his work on methods for quantitative analysis of the web and social media.

Mapping Online Publics <http://mappingonlinepublics.net/> This site describes an inter-linked series of research projects conducted by a research team led by Axel Bruns, focusing on the use and impact of social networking tools in Australia and beyond.

References

- Baym, N. (1995) The emergence of community in computer-mediated communication. In *Cybersociety*. Ed. S. Jones. Thousand Oaks, CA: Sage pp. 138 - 163.
- Baym, N. K. (2000) *Tune In, Log On: soaps, fandom and online community*. Thousand Oaks, CA: Sage.
- Baym, N. K. (2005) Introduction: Internet Research as It Isn't, Is, Could Be, and Should Be. *The Information Society* 21(4): 229-232.
<http://dx.doi.org/10.1080/01972240591007535>
- Baym, N. K. (2009) A Call for Grounding in the Face of Blurred Boundaries. *Journal of Computer-Mediated Communication* 14(3): 720-723.
<http://dx.doi.org/10.1111/j.1083-6101.2009.01461.x>
- Blank, G. and B. C. Reisdorf (2012) The participatory web. *Information, Communication & Society* 15(4): 537-554.
<http://dx.doi.org/10.1080/1369118X.2012.665935>
- Boellstorff, T. (2008) *Coming of Age in Second Life: an anthropologist explores the virtually human*. Princeton, NJ: Princeton University Press.
- Boellstorff, T. (2010) A typology of ethnographic scales for virtual worlds. In *Online Worlds: convergence of the real and the virtual*. Ed. W. S. Bainbridge. London: Springer pp. 123-133.
- boyd, d. (2007) Why Youth ♥ Social Network Sites: The Role of Networked Publics in Teenage Social Life. In *The John D. and Catherine T. MacArthur Foundation Series on Digital Media and Learning - Youth, Identity and Digital Media volume*. Ed. D. Buckingham. Cambridge, MA: MIT Press. - pp. 119-142.

boyd, d. and K. Crawford (2011) Six Provocations for Big Data (September 21, 2011). , September 2011. *A Decade in Internet Time:*

Symposium on the Dynamics of the Internet and Society, September 2011. Oxford Internet Institute.

boyd, d. m. (2008) Taken Out of Context: American teen sociality in networked publics Doctoral thesis submitted to Graduate Division,

University of California, Berkeley. <http://www.danah.org/papers/TakenOutOfContext.pdf>

Brügger, N. (2013) Historical Network Analysis of the Web. *Social Science Computer Review* 31(3): 306-321.

<http://ssc.sagepub.com/content/31/3/306.abstract>

Bruns, A. (2011) How long is a tweet? Mapping dynamic conversation networks on Twitter using gawk and gephi. *Information, Communication*

& Society 15(9): 1323-1351. <http://dx.doi.org/10.1080/1369118X.2011.635214>

Bruns, A. and J. Burgess (2011) #Ausvotes: How twitter covered the 2010 Australian federal election *Communication, Politics & Culture* 44(2):

37-56.

Bruns, A. and Y. E. Liang (2012) Tools and methods for capturing Twitter data during natural disasters *First Monday* 17(4).

<http://pear.accc.uic.edu/ojs/index.php/fm/article/view/3937/3193>

Bruns, A. and S. Stieglitz (2012) Quantitative Approaches to Comparing Communication Patterns on Twitter. *Journal of Technology in Human*

Services 30(3-4): 160-185. <http://dx.doi.org/10.1080/15228835.2012.744249>

Bryman, A. (2006) Paradigm peace and the implications for quality. *International Journal of Social Research Methodology* 9(2): 111-126.

Burrell, J. (2009) The field site as a network: a strategy for locating ethnographic research. *Field Methods* 21(2): 181-199.

- Creswell, J. W. (2011) Controversies in mixed methods research. In *The SAGE Handbook of Qualitative Research*. Ed. N. K. Denzin and Y. S. Lincoln. Thousand Oaks, CA: Sage pp. 269-283.
- Creswell, J. W. and V. L. Plano Clark (2007) *Designing and conducting mixed methods research*. Thousand Oaks, CA: Sage.
- De Maeyer, J. (2012) Towards a hyperlinked society: A critical review of link studies. *New Media & Society*.
<http://nms.sagepub.com/content/early/2012/12/04/1461444812462851.abstract>
- Dillman, D. A. (2011) *Mail and Internet Surveys: The Tailored Design Method -- 2007 Update with New Internet, Visual, and Mixed-Mode Guide*: Wiley.
- Dirksen, V., A. Huizing and B. Smit (2010) 'Piling on layers of understanding': the use of connective ethnography for the study of (online) work practices. *New Media & Society* 12(7): 1045-1063.
- Efimova, L. A. (2009) Passion at work: blogging practices of knowledge workers Doctoral thesis submitted to Novay PhD Research Series Vol 24 (2009), <http://igitur-archive.library.uu.nl/dissertations/2009-0626-200434/UUindex.html>
- Elford, J., G. Bolding, M. Davis, L. Sherr and G. Hart (2004) The Internet and HIV study: design and methods. *BMC Public Health* 4(1): 39.
<http://www.biomedcentral.com/1471-2458/4/39>
- Ess, C. M. and W. H. Dutton (2013) Internet Studies: Perspectives on a rapidly developing field. *New Media & Society*.
<http://nms.sagepub.com/content/early/2013/04/24/1461444812462845.short>

- Garcia, A. C., A. I. Standlee, J. Bechkoff and Yan Cui (2009) Ethnographic Approaches to the Internet and Computer-Mediated Communication. *Journal of Contemporary Ethnography* 38(1): 52-84. <http://jce.sagepub.com/content/38/1/52.abstract>
- Garton, L., C. Haythornthwaite and B. Wellman (1997) Studying online social networks. *Journal of Computer Mediated Communication* 3(1). <http://jcmc.indiana.edu/vol3/issue1/garton.html>
- Geiger, R. S. and D. Ribes (2011) Trace Ethnography: Following Coordination through Documentary Practices. *Proceedings of the 2011 44th Hawaii International Conference on System Sciences*, IEEE Computer Society: 1-10.
- Gibson, R. K., K. Gillan, F. Greffet, B. J. Lee and S. Ward (2013) Party organizational change and ICTs: The growth of a virtual grassroots? *New Media & Society* 15(1): 31-51. <http://nms.sagepub.com/content/15/1/31.abstract>
- Hampton, K. (2007) Neighborhoods in the Network Society: The e-Neighbors Study. *Information, Communication & Society* 10(5): 714-748.
- Hampton, K. N., L. S. Goulet, C. Marlow and L. Rainie (2012) Why most Facebook users get more than they give. The effect of Facebook ‘power users’ on everybody else Washington, DC: Pew Research Center’s Internet & American Life Project. <http://pewinternet.org/Reports/2012/Facebook-users.aspx>
- Hargittai, E. (2002) Beyond Logs and Surveys: In-Depth Measures of People's Online Skills. *Journal of the American Society for Information Science and Technology* 53(14): 1239-1244.
- Herring, S. C. (2010) Web Content Analysis: Expanding the Paradigm. In *International Handbook of Internet Research*. Ed. J. Hunsinger, L. Klastrup and M. Allen. Springer Netherlands pp. 233-249.

- Herring, S. C., I. Kouper, J. C. Paolillo, L. A. Scheidt, M. Tyworth, P. Welsch, E. Wright and N. Yu (2005) *Conversations in the Blogosphere: An Analysis "From the Bottom Up"*, Big Island, HI, USA. <http://dx.doi.org/10.1109/HICSS.2005.167>
- Hesse-Biber, S. and A. J. Griffin (2013) Internet-Mediated Technologies and Mixed Methods Research: Problems and Prospects. *Journal of Mixed Methods Research* 7(1): 43-61. <http://mmr.sagepub.com/content/7/1/43.abstract>
- Hine, C. (2000) *Virtual Ethnography*. London: Sage.
- Hine, C. (2005) Internet research and the sociology of cyber-social-scientific knowledge. *The Information Society* 21(4): 239-248.
- Hine, C. (2007) Connective ethnography for the exploration of e-science. *Journal of Computer Mediated Communication* 12(2).
<http://jcmc.indiana.edu/vol12/issue2/hine.html>
- Hine, C. (2008a) Virtual ethnography: modes, varieties, affordances. In *The Sage Handbook of Online Research Methods*. Ed. R. M. Lee, N. Fielding and G. Blank. London: Sage pp. 257-270.
- Hine, C. (2008b) *Systematics as Cyberscience: computers, change and continuity in science*. MIT Press.
- Hogan, B. (2008) Analyzing social networks via the Internet. In *The SAGE Handbook of Online Research Methods*. Ed. N. Fielding, R. M. Lee and G. Blank. London: Sage. 141-160.
- Howard, P. N. (2002) Network Ethnography and the Hypermedia Organization: New Media, New Organizations, New Methods. *New Media & Society* 4(4): 550-574. <http://nms.sagepub.com/content/4/4/550.abstract>

- Johnson, R. B., A. J. Onwuegbuzie and L. A. Turner (2007) Toward a Definition of Mixed Methods Research. *Journal of Mixed Methods Research* 1(2): 112-133. <http://mmr.sagepub.com/content/1/2/112.abstract>
- Kazmer, M. M. and B. Xie (2008) Qualitative Interviewing in Internet Studies: Playing with the media, playing with the method. *Information, Communication & Society* 11(2): 257 - 278. <http://www.informaworld.com/10.1080/13691180801946333>
- Kozinets, R. (2007) Netnography 2.0. In *Handbook of Qualitative Research Methods in Marketing*. Ed. R. W. Belk. Cheltenham: Edward Elgar Publishing pp. 129-142.
- Kozinets, R. V. (2009) *Netnography: Doing Ethnographic Research Online*: SAGE Publications.
- Malta, S. (2012) Using Online Methods to Interview Older Adults about their Romantic and Sexual Relationships. In *Researching Later Life and Ageing - Expanding Qualitative Research Horizons*. Ed. M. Leontowisch. Basingstoke: Palgrave Macmillan.
- Marcus, G. (1995) Ethnography in/of the world system: the emergence of multi-sited ethnography. *Annual Review of Anthropology* 24: 95-117.
- Markham, A. N. (1998) *Life online: researching real experience in virtual space*: Altamira Press.
- Markham, A. N. (2012) Moving into the flow: Using a network perspective to explore complexity in Internet contexts. In *Network Analysis: Methodological Challenges*. Ed. S. Lonborg. Aarhus, Denmark: University of Aarhus Center for Internet Research Monograph Series pp. 47-58.
- Mason, J. (2011) Facet Methodology: the case for an inventive research orientation. *Methodological Innovations Online* 6(3): 75-92. <http://www.pbs.plym.ac.uk/mi/pdf/8-02-12/MIO63Paper31.pdf>

Maxwell, J. A. (2010) Using Numbers in Qualitative Research. *Qualitative Inquiry* 16(6): 475-482.

<http://qix.sagepub.com/content/16/6/475.abstract>

Miller, D. and D. Slater (2000) *The Internet: an ethnographic approach*. Oxford: Berg.

Mol (2002) *The Body Multiple: ontology in medical practice*. Durham, NC: Duke University Press.

Murthy, D. (2008) Digital Ethnography. *Sociology* 42(5): 837-855. <http://soc.sagepub.com/content/42/5/837.abstract>

Murthy, D. (2013) *Twitter: social communication in the Twitter age*. Cambridge: Polity.

Nardi, B. A. (2010) *My Life as a Night Elf Priest: An Anthropological Account of World of Warcraft*. University of Michigan Press.

Nip, J. Y. M. (2004a) The Queer Sisters and its Electronic Bulletin Board. *Information, Communication & Society* 7(1): 23-49.

<http://dx.doi.org/10.1080/1369118042000208889>

Nip, J. Y. M. (2004b) The Relationship between Online and Offline Communities: The Case of the Queer Sisters. *Media, Culture & Society*

26(3): 409-428. <http://mcs.sagepub.com/content/26/3/409.abstract>

Orgad, S. S. (2005a) From online to offline and back: moving from online to offline relationships with research informants. In *Virtual methods:*

issues in social research on the Internet Ed. C. Hine. Oxford: Berg pp. 51-65.

Orgad, S. S. (2005b) *Storytelling Online: talking breast cancer on the Internet*. New York: Peter Lang.

Park, H. W. and R. Kluver (2009) Trends in online networking among South Korean politicians — A mixed-method approach. *Government*

Information Quarterly 26(3): 505-515.

- Peng, T.-Q., L. Zhang, Z.-J. Zhong and J. J. Zhu (2012) Mapping the landscape of Internet studies: Text mining of social science journal articles 2000–2009. *New Media & Society*. <http://nms.sagepub.com/content/early/2012/11/22/1461444812462846.abstract>
- Procter, R., F. Vis and A. Voss (2013) Reading the riots on Twitter: methodological innovation for the analysis of big data. *International Journal of Social Research Methodology* 16(3): 197-214. <http://dx.doi.org/10.1080/13645579.2013.774172>
- Reips, U.-D. and L. E. Buffardi (2012) Studying Migrants with the Help of the Internet: Methods from Psychology. *Journal of Ethnic and Migration Studies* 38(9): 1405-1424. <http://dx.doi.org/10.1080/1369183X.2012.698208>
- Rogers, R. (2013) *Digital Methods*. Cambridge, MA: MIT
- Rogers, R. and N. Marres (2000) Landscaping climate change: a mapping technique for understanding science and technology debates on the World Wide Web. *Public Understanding of Science* 9(2): 141-163. <http://pus.sagepub.com/content/9/2/141.abstract>
- Savage, M. and R. Burrows (2007) The Coming Crisis of Empirical Sociology. *Sociology* 41(5): 885-899.
<http://soc.sagepub.com/cgi/content/abstract/41/5/885>
- Schneider, S. M. and K. A. Foot (2004) The web as an object of study. *New Media & Society* 6(1): 114-122.
- Scolari, C. A. (2009) Mapping conversations about new media: the theoretical field of digital communication. *New Media & Society* 11(6): 943-964. <http://nms.sagepub.com/content/11/6/943.abstract>
- Sterne, J. (2005) Digital Media and Disciplinarity. *The Information Society* 21(4): 249-256. <http://dx.doi.org/10.1080/01972240591007562>

Strathern, M. (2002) Abstraction and Decontextualization: An Anthropological Comment. In *Virtual Society? Technology, Cyberbole, Reality*.

Ed S. Woolgar. Oxford: Oxford University Press pp. 303–13.

Thelwall, M. (2004) *Link Analysis: an information science approach*. Amsterdam: Elsevier.

Thelwall, M. (2008) Fk yea I swear: cursing and gender in MySpace. *Corpora* 3(1): 83-107.

<http://www.eupublishing.com/doi/abs/10.3366/E1749503208000087>

Thelwall, M. (2009) *Introduction to Webometrics: Quantitative Web Research for the Social Sciences*. San Rafael, CA: Morgan and Claypool.

Vasileiadou, E. and P. van den Besselaar (2006) Linking shallow, linking deep. How scientific intermediaries use the Web for their network of collaborators. *Cybermetrics: International Journal of Scientometrics, Informetrics and Bibliometrics* 10(1).

<http://cybermetrics.cindoc.csic.es/articles/v10i1p4.html>

Wellman, B. and C. Haythornthwaite, Eds. (2002) *The Internet in Everyday Life*. Oxford, Blackwell.

Williams, D., N. Ducheneaut, L. Xiong, Y. Zhang, N. Yee and E. Nickell (2006) From Tree House to Barracks: The Social Life of Guilds in World of Warcraft. *Games and Culture* 1(4): 338-361. <http://gac.sagepub.com/content/1/4/338.abstract>

Method	Data source	Typical research questions/insights	Potential limitations
Participant observation in an online setting	Field notes based on direct observation and participation by the researcher in group activities. In an online ethnography, field notes based on participant observation are often supplemented by interviews, surveys etc. to provide descriptive statistics and to corroborate the researcher's emerging interpretations.	What are the group norms and behaviours? What is the social hierarchy? How is social order maintained within the group? Participant observation allows the researcher to develop an in-depth experiential understanding.	Focusing on the online group as a bounded social space may limit understanding of how people move between and combine different online and offline spaces in their everyday lives.
Online interviews	Online interviews can be conducted in real time, using chat and messaging services, or asynchronously, for example by email. Interviews may be very structured, or may involve use of a more flexible interview guide adapted for each individual encounter. Data will be in the form of text files recording the messages exchanged during the interview.	How do interviewees experience and understand my topic of interest? Online interviews allow the researcher to focus in on areas of specific concern and can allow for uninhibited discussion of sensitive topics.	Interviewers can find that reliance on online interviews alone limits their ability to understand interviewees' offline worlds. Interviews provide an artificial interactional context. There is little possibility of cross-checking between what interviewees say they do and what they actually do.
Online surveys	Online surveys involve collecting answers to a structured set of questions, often presented on a web site. Invitations to take part in the survey may be posted to a known population, for example via a mailing list, or may be publicised more widely, for example via social networking sites, to an undetermined population of potential participants	How do my population of interest experience and understand my topic of interest? What socio-demographic factors account for differences? Surveys allow a structured approach to exploring participants' experiences and understandings and lend themselves to quantitative analysis.	Generalizability can be problematic for online surveys where the underlying population is unknown: it can be unclear who saw the survey invitation, and response biases are thus hard to determine. Surveys provide an artificial context for expression of attitudes.
Face-to-face interviews	Interviews with Internet users can be conducted face-to-face, in either structured, semi-structured or open formats, resulting in audio- or video-recordings for transcription into text files, accompanied by field notes.	How do interviewees experience and understand my topic of interest? A face-to-face interview can be a powerful way to build rapport with an interviewee.	Face-to-face interviews can be limited as a tool for understanding people's Internet activities if not cross-checked with what people actually do online.
Structured activity sessions	Internet users are asked to complete a set task whilst being observed by the researcher.	How do people retrieve information from the Internet, and how do they interpret what they find? By asking people to focus on a task of interest the researcher generates a data set directly targeted on their topic of interest and avoid the distortions of	These task-based sessions can be somewhat artificial, in that they do not represent the way that people ordinarily go about using the Internet in their everyday lives.

		retrospective accounts.	
Surveys using offline media	Surveys can be completed on paper, by telephone or various forms of computer-assisted interviewing. A structured set of questions is answered by a population which may be systematically sampled.	How do my population of interest experience and understand my topic of interest? What socio-demographic factors account for differences? A survey using offline media to explore aspects of people's Internet use can use representative sampling techniques, and can avoid response bias in favour of intensive Internet users.	Surveys about what people do online can fail to pick up on the nuances of the meaning they place on these behaviours. Recall about online activities in survey responses can be inaccurate.
Content analysis on a corpus of data found online	Structured analyses can be conducted on the content of a corpus of data such as blogs, or messages sent to an online forum	What different kinds of contribution can be identified in this content? How do different online settings differ? A topic of interest can be explored in naturally occurring data, avoiding the need to ask people questions.	Analysis may be restricted to taking what people say at face value, and the ability to understand nuances can be limited. Content analysis may neglect the very different meanings which people may be giving to overtly similar online interactions.
Online experiments	Experiments can be conducted online via websites, allowing for people who are not geographically co-located to undertake an experimental task aimed at exploring some aspect of their behaviour, attitudes, cognitive function etc.	Online experiments can provide large pools of participants, and allow for a structured analysis of the effects of varying the conditions in relation to the focus of the research.	Some data collected may be unusable since the commitment of participants to sensible completion of the task cannot be guaranteed. The experimental conditions may not mirror more complex real-life circumstances.
Social network analysis	Network analysis can be carried out on any data set which can be used as a proxy for different forms of relationship. For example, log files of interactions, comments on blog posts or hyperlinks between websites can all be used to conduct network analysis.	What is the pattern of connections between nodes (e.g. people, organizations) in this network? How do patterns of connection change over time? Which entities are more central to the network, and which are peripheral?	The assumptions made by the network analysis about the relationships depicted by the dataset may not be valid – numerous hyperlinks between two websites may not always, for example, mean that the two organizations producing them are ideologically close in some way.
Sentiment analysis	Sentiment analysis uses machine learning, deploying techniques from natural language processing to determine the emotional mood of communications. It is often latterly used on large data sets from social networking sites and Twitter.	What are the large scale patterns of positive and negative sentiments expressed by users in relation to my topic? How does opinion change over time?	Sentiment analysis relies on the ability to develop stable means of determining the mood of a message. This is inevitably error prone and may miss out on specific nuances. It is particularly difficult to reliably identify sentiment in very short messages.
Semiotic	Semiotic analysis looks at the composition of	Who are the audience addressed by this	Semiotic analysis depends on the cultural

analysis of individual online artefacts	an artefact, such as a website, as a set of signifying elements chosen by the creator. Semiotic analysis looks at the consequences of the choices of signs and their arrangement, looking beyond overt meanings to the connotations of the various choices.	text, and what meanings does it convey to them?	competencies of the analyst. The analyst must be able confidently to identify the repertoire of signs available to the author and the denotations of the different choices. This may differ from the way that diverse actual readers may interpret the text in their own local circumstances.

Table 1. Commonly used Internet research methods which may feature in mixed methods research designs.

ⁱ Cyberculture is a term used to denote cultural formations occasioned by use of computer networks.

ⁱⁱ A newsgroup was a publicly available asynchronous discussion group, focused on a specific topic. The rec.arts.tv.soaps newsgroup was focused on discussion of television soap operas – Baym (1995; 2000) focused in particular on threads related to one US soap opera, studying the social norms, hierarchies and practices of the group of contributors to threads of discussion.

ⁱⁱⁱ Second Life is a graphical virtual world in which the user is visually represented by an avatar. Users can socialize, conduct business and create objects within the virtual world.

^{iv} A hyperlink on the World Wide Web is way of connecting two websites, or two parts of a site, so that the user can navigate between them. Patterns of hyperlinking on the web can therefore potentially tell us something about patterns of affiliation or affinity between websites, as conceptualised by the creators of the links.

^v Thelwall (2004; 2009) provides an overview of techniques for subjecting the patterns of hyperlinks between websites to quantitative analysis, which can, he argues, inform social science questions and can be combined with qualitative techniques.

^{vi} The Issue Crawler collects data from a set of websites specified by the researcher and visualizes the links between them, based on the pages that they link to. It allows the researcher to explore the patterns of affiliation and connection between sites.

^{vii} SocSciBot collects data on the links between a specified set of websites and also allows researchers to carry out textual analysis on word frequencies and concordance for basic corpus linguistics approaches.

^{viii} SentiStrength analyses positive and negative sentiment in short texts such as Twitter data, allowing researchers to explore emotional content and trends.