Note: this is a Feb 2020 working variation of a paper forthcoming chapter for the book: *The Field of Qualitative Research* (Edited by Patricia Leavy, will be published by Oxford University Press, anticipated 2020). Please contact author amarkham@gmail.com for more information about the Handbook version.

DOING ETHNOGRAPHIC RESEARCH IN THE DIGITAL AGE

Annette N. Markham, PhD,
Co-Director, Digital Ethnography Research Centre, RMIT University

INTRODUCTION

As I write this piece, the world is undergoing a wave of what is being called “social distancing” to help slow down a pandemic. It involves staying at home, keeping a certain distance from other humans, and maintaining strict protocols of hygiene. My old job is in Denmark, my new job is in Australia, my home town is in the USA. I work strange hours. I socialize in ways that are alternately very familiar and very unfamiliar. As a long-time researcher of digital culture, I’m no stranger to online community. But unfamiliar to me are the number of chats, ambient hanging around, and conversations I’m having with people using video conferencing. Even my aging mother is getting on the video conferencing bandwagon, who struggles to figure out whether the icon on her screen is a speaker or a video camera. I tell her it doesn’t matter what it looks like, she should just push it and then we can see each other as we chat. This scene, these instructions, these galleries of faces on laptop windows, is played out in similar ways around the world in countries with good internet access.

There’s a lot of talk in my academic social media feeds about transforming traditional research into digital research. Researchers who are isolated need to either use digital tools to conduct their semester research projects or shift from physical to digital contexts to find and study their phenomenon.

What’s unique about digital research? In this era of constant connectivity and digital saturation in all spheres and moments of everyday life, very few research domains do not brush up against whatever we call “the digital” in some way. Immediately after this initial claim, we might back up to ask a more fundamental question: what is meant by “the digital” or “the digital age?” This opens up a floodgate of other questions, such as: How have nearly three decades of public access and use of the Internet changed the way qualitative and ethnographic research is conducted? What makes digital contexts unique? What sorts of questions do people ask if they call themselves digital researchers? What particular challenges do they face?

There are important differences between 1) research that is *digital by necessity* because of physical distance between researcher and participants; 2) research that uses *natively digital tools* to

study phenomena¹; and 3) research that is focused on digital culture. This chapter focuses on #3. It is situated in studies of people, societies, and sociotechnical relations that are digitally-saturated, or somehow impacted by transformations wrought by digitalization, widespread internet connectivity, and the largescale datafication of society.

This chapter offers an introduction to some key issues, challenges for ethnographic research of digitally saturated social environments, online social contexts, or digitally-mediated phenomena. It focuses on empirical approaches, meaning knowledge is derived from investigation, observation, experiment, or experience, in contrast to knowledge emerging from theoretical or logical philosophical writing, or close readings of literature. In the context of digital social research, this might then involve observing or collecting actual behaviors and actions in social networking platforms, or studying use and interactions with and around digital devices, technologies, and media in naturalistic environments. It might also involve recording and observing in contrived settings, like workshops, focus groups, experiments, or interviews. The target of one's study might include people in their physical forms or just data produced through human behaviors, movements, or flows of information. One's focus might seem small scale, whereby one is looking at a single case, instance, individual or small group, or it could seem largescale, such as exploring patterns in aggregated datasets, analyzing upswells or shifts of interest in events or crisis, examining how ideas flow or emerge through various groups, platforms, or networks. With such a broad range of topics, approaches, choices, there will obviously be different theories, concepts, methods, ethics, and best practices.

Digital research is not defined by a particular perspective, method, tool, or unit of analysis, but by the degree to which the digital is centrally relevant in the phenomenon, context, or focus of analysis. The research design and focus may look very different, as a result, but still qualify as "digital research." Take three researchers studying hate speech: Researcher 1 might study how hate speech manifests in digital platforms, but the way they access that knowledge is through conducting face-to-face interviews with journalists who have experienced hate speech. Researcher 2 might never talk with people at all, but analyze massive Twitter datasets to visualize and search for patterns in the memetic flow of hate speech memes, focusing on how these might begin in one way but change and morph as they move and are remixed or remediated through various social networks. Researcher 3 might explore how members define and frame hate speech in extremist Reddit communities. All of these qualify as digital research, so this chapter considers a very broad domain. *The key is that for the researcher, there is something about 'the digital' that is centrally relevant.*

Likewise, qualitative research design in the digital domain may not always look like it's 'qualitative' in the traditional sense. One researcher might draw heavily on tools of classic

¹ Many researchers use digital tools to capture, measure, visualize, or describe phenomena. This is not the study of digital society or digital culture, per se, but the use of digital media to automate some aspects of sensemaking that might otherwise (and traditionally) be conducted by hand. This is not the point of this chapter, since the use of these tools is a topic of its own. Visualizations, network maps, and largescale data analytics are useful in qualitative research for exploring large datasets, conducting pattern recognition, or immersing in various layers of the situation. One of the most well known books to summarize these approaches is *Digital Methods*, by Richard Rogers (2013).

anthropology but the focus is on how people use digital media. Another researcher's toolbox might resemble that of a computer scientist or statistician.

As preview for what this chapter does: Because terms like 'the digital' or 'digital research' do not have standardized meanings or operationalizations, the first part of this chapter offers some definitional clarification of what is meant by these terms, by tracing some key historical markers or trends since the internet became more publicly accessible in the early-mid 1990s. This is followed by a sketch of key characteristics of the internet and/or digital media that continue to influence the enactment and negotiation of identity, relationships, and sociality at granular levels of experience, where social researchers often focus their attention. Third, I discuss several issues (problems, dilemmas) around research design, fieldwork, data management, and analysis/interpretation in my own longstanding digital research practice. These are not all encompassing: To give examples or talk about best practices, I focus only on what I or my students have found useful over the years. Finally, at the end of this chapter, I discuss ethics and politics of engaging in research in the digital age.

The material presented in this chapter emerges from and reflects my own practice as someone who has been using digital methods and studying digitally saturated social contexts since the early 1990s. I study, broadly speaking, the impact of digitalization and datafication on society, looking at micro level interactions and practices in everyday life and also broad cultural and social structural implications. My empirical research is guided by interpretive ethnography, autoethnography, phenomenology, and symbolic interactionism. I also have a strong background in critical theory, so when I interpret generated materials, I often apply a more deductive set of critical cultural theories about how these micro practices are functioning at larger levels. I have used and experimented with almost every digital medium, platform, and tool possible.

DIGITAL RESEARCH: DEFINITION AND HISTORY

How do researchers delineate what is understood as "digital" and what difference does this make, in terms of a research stance and understanding of the unit of analysis or the subject of our researcher gaze? This chapter takes the stance that just using digital tools to conduct research is not digital research, *per se*.

At the most basic level, something can be considered digital when it "has been developed by, or can be reduced to, the binary—that is bits consisting of 0s and 1s" (Horst & Miller, 2012, p. 5). But the digital has come to represent much more than this. Even in 1995, Nicholas Negroponte described this shift from atoms to bits as not just a matter of making something digital, but the harbinger of existential changes. 'Being', in a digital era, emerges through --and because of-- our relationships with/in the sociotechnical, or drawing on Negroponte more directly, our "acquaintance over time," with machine agents who understand, remember, and respond to our individual uniqueness "with the same degree of subtlety (or more than) we can expect from other human beings" (1995, p. 164). We don't live life with media, as Mark Deuze notes, but in media, like fish in water (2012).

It is not an exaggeration to say that the internet has transformed cultural practices around the world. This influences what counts as a human subject in sociological studies, complicates the

legacies for how to gather or analyze material in qualitative research, and provides new sources for meaningful information about the situations we want to study.

If we trace the history of the digital era, we see certain milestones marking how social researchers from a range of related disciplines such as anthropology, sociology, human computer interaction, communication and media studies, or science and technology studies have defined and studied the digital over time and how the emphasis changes, along with salient issues. The important part of this history for purposes of understanding research design in the digital age is that subsequent shifts do not replace or negate previous issues/challenges, but simply add more considerations. Thus, researchers conducting analysis of how individuals communicate with each other through internet mediated systems still draw inspiration and methodologies from scholars who studied this in the early 1990s. Researchers in 2020 looking at how influencers developing Instagram Stories find value in looking at how previous researchers studied blogging practices in 2002 (e.g., Walker, 2008). Scholars exploring playful practices in Reddit might find value in early works on play in usenet groups (e.g., Danet, 2001). Flexible adaptation of methods is therefore balanced by being familiar with and drawing on the important legacies of earlier work.

Below, I depict five major waves that have influenced social research design. Other depictions or classifications may exist.

Wave 1, early-mid 1990s: Cyberspace

In the early and mid-1990s, when the internet was new for most people, researchers were focused on (dis)embodiment, identity, geographically dispersed community, and matters of virtuality. No wonder: all these ways of being in the world and with others were accomplished through the exchange of texts in various shared spaces online. The novelty of creating identity through text, or writing oneself into being (Sunden, 2003, Markham, 1998) was accompanied by the ability to be in synchronous interactions with people from all over the world. Anonymity was possible and people widely used pseudonyms or “nicks” (nicknames) instead of their given names (as memorialized in the famous 1993 New Yorker cartoon captioned “On the Internet, nobody knows you’re a dog”). Distance seemed irrelevant, so digital space and place were of great interest. Online communities emerged based on special interests rather than geography (cf the edited collections by Jones, 1995; 1999). The impact of communicating with people in Cyberspace was also quite visceral, often in ways that startled people and researchers into recognizing that (and studying how) physicality and sociality are not the same thing, that embodiment is performative, that textual action can harm (c.f., Dibble, 1993), and that identity and being are informational as much as they are embodied (e.g., Cherny & Weiss, 1996).

In the mid-late 1990s, the look and feel of computer-mediated communication grew more visual through the web, or the World Wide Web (WWW) as it was called at the time. There was a marked growth of commerce and information exchange online. The late 1990s was still considered somewhat a ‘wild west’ or ‘electronic frontier’ (in the terms of Barlow (1996) and Rheingold (1993), respectively) but an equally dominant metaphor was the information superhighway, popularized by US Vice President Al Gore. While online communities still thrived and researchers continued to focus

on digital culture in MUDS or MOOS (Multi User Dimensions), IRC (Internet Relay Chat), or BBS (bulletin board systems), many people only experienced the internet as a portal for information flow.

Wave 2, early 2000s: Web 2.0 and the age of sharing

The early 2000s were characterized by the return of the social web, but in markedly different ways than the sociality of the early 1990s. Here, rapid growth of software for interaction via the web created greater possibilities for commenting, writing back, and otherwise giving feedback to information that was being posted. This happened at the individual level of blogs, but also in news and commerce sites. Amazon reviewers or ebay sellers could be given stars, so it was interesting to study this type of interaction. Citizens could remark, correct, augment, or disagree with a news story. Thus citizen journalism became a relevant topic, Geolocation, remarkable for its absence in the early 1990s, became relevant again as phones shrank and the use of SMS grew rapidly. This confluence of technological development meant people could interact in ways that were not previously possible or easy. Using text messaging, activists could mobilize for social action (for example, the WTO protests in Seattle, cf Rheingold, 2002). Youth could engage in conversations with each other while in the classroom, or interact intimately in unsanctioned --because of religious restrictions or cultural norms-- relationships.

Wave 3, Mid-late 2000s: Platformization of social networking

Myspace, Facebook, Sina Weibo, Youtube, Twitter, Instagram. While not the first, these platforms represent how expression, interaction, networking, and sharing could be combined in a single online service. They're notable for me, not just because they were most common for me (with the exception of Sina Weibo, the Chinese social networking platform, all these are U.S. companies), but because they represent solid examples of the success of web 2.0 ideology added to the business models that provided so-called 'free' services, reliant on users generating traffic through uploads, sharing, and building networks of connections. They all emerged approximately in the order listed between 2002-2008. They are heralds of the rise of the 'platform,' or platformization of culture.

With platforms emerging that could withstand the test of time and the exponential growth of competition, researchers were interested in how sociality, norms of expression and interaction, and the flow of ideas were interlinked, organized. Networks become an interest area for ethnographic researchers. What we call networked culture or network society is meaningful not because networks are new, but because digitalization plus scale creates the possibility for dense interconnections and global flows. These interconnections at the very least² facilitate communities of interest rather than geographic location. Along with interconnectivity, there is swift dispersion of information. This made viral media possible, along with citizen journalism, the use of social media to spread news about

² This chapter presents only a small slice of the concept of network culture. For comprehensive texts, one could look at Manuel Castell's classic volumes on Network Culture. For an array of empirical studies and theoretical discussions, see the work of Zizi Papacharissi, as well as the many curated volumes she produces.

catastrophes, the memetic spread of disinformation or what we now call Fake News, and the rise of microcelebrities and influencers or celebrities (like Taylor Swift, whose breakout was sparked through Myspace).

There was another significant shift in the mid 2000s, to build one's identity with attention to authenticity. Likely, this was influenced less by the increasing viscosity of the interfaces and more by the rise of the idea that the internet, via these platforms, was now a place to nurture your real relationships and build your reputation. As mentioned earlier, none of these moments supplanted previous moments. Thus, many people were still engaging in immersive, anonymous environments. But the trend was to show your identity, and not just your everyday identity but your 'best of' identity, since it could potentially be marketable. Whereas in 1995, the cover of Time Magazine warned of anonymous strangers stalking children to peddle Cyberporn (July 3rd issue), the 2006 Time Magazine Person of the Year is "You" (December 25th issue) This wave of developing technologies for sharing is marked by the collapsing boundaries between entertaining and being entertained, looking and being seen, commenting on others and being commented on. Jenkins remarked on this already in 2006 with his book *Convergence Culture*. Marwick and boyd focused on "context collapse" as a specific outcome of convergence and networked sociality (2010). Bruns (2008) wrote about what he called "produsage," the specific blurring of boundaries between producers and users of user generated content.

Wave 4, early 2010s: The rise of big data

In 2011, several events had massive global impact: the earthquakes in Christchurch, New Zealand; the Japanese earthquake, tsunami, and nuclear disaster; and the Arab Spring, which was most broadly represented in the use of social media by journalists on the ground during the Tunisian and Egyptian uprisings. In these events, for different reasons of course, the social media responses were rapid but also widely and swiftly spread. Official and unofficial content being posted immediately from 'the front lines' was reposted, forwarded, liked, commented on. Tweets and replies were retweeted, content from one platform was cross posted to other platforms, texts and visuals were remixed into montages, creating more user generated content to spread. The general uptake of social media caused massive swells of information flows, which could map the qualities and characteristics of what Hermida called "ambient journalism," (2010), and also influenced how the viscosity of these events rippled through people's lived experiences on wider scales than the events themselves. The use of mobile phones and wireless networks combined with the dominance and reach of social media platforms like Youtube, Twitter, and Facebook. It's no surprise that researchers began to want to understand the scale at which people were functioning, responding, and interacting using these platforms. Twitter as a dominant medium for both effective crisis communication and false reports or rumour mills became a central focus for researchers in 2011 (c.f., Murthy, 2011; Vis, 2012; Papacharissi & Oliveira, 2012; Bruns & Highfield, 2012).

Partly in response to this, but also partly to create stronger personalization of news feeds, content generated by friends, and targeted web and in-app advertising, researchers were collecting larger and larger datasets. Scraping became a common term and with the beautiful renderings made

possible with off the shelf (easy to use) software like Gephi, the world was captivated by visualizations of global, large scale trends. “Big Data” became a common term to use to describe the computational analyses of datasets too large to be handled or understood by human cognition.

Between 2012-2014, there was a steady stream of critiques of this concept (e.g., boyd & Crawford, 2012; Grinter, 2013; Gitelman, 2013). Qualitative researchers insisted that ethnographic data was already always big (Boellstorff, 2013), that any metric for big data would yield only a partial and non-representative sample (Baym, 2013), that data was a vague concept to begin with (Markham, 2013a); that data was found, lost and made, certainly not given (e.g., Bruun-Jensen & xx (2013), and that small data, thick data, or deeper data were useful counterpoints to the big data surge (e.g., Brock, 2015). So a significant type of research in this moment was to both study how data could help augment qualitative research and to resist that very trend.

This focus on big data, or matters of concern at scale, was important for social researchers as well as computational analysts since there were global events that demanded detailed, field- and context-specific attention and comprehension. Using creative visualizations to show global patterns helped us trace the use of Twitter by journalists in global crises (for a good overview, cf Vis, 2012); the speed of responses to regional or country wide events (e.g., Bruns, Burgess, Crawford, & Shaw, 2012), trends in commenting, voting, and other event-based behaviors (c.f., Rossini, Hemsley, Tanapabrunsun, Zhang, & Stromer-Galley, 2018), and characteristics of what became known as “hashtag publics.” Using large teams for research projects, while still novel in many social research fields, has become a productive approach to retain interpretive, grounded, or ethnographic epistemologies and techniques while working on the issues at scale.³

Wave 5, mid-late 2010s: Algorithms, predictive analytics, and more than human relations.

Following the surge of interest in big data analytics, the use of automated data collection and analysis tools, social research of the digital has turned again, to consider what it means to live in a world that is data-driven and where machine learning creates automated analysis of who we are and what we (will) do. This includes the study of how social categorization is becoming an automated process influenced by algorithmic logics. The process of prediction, which drives everything from recommendation systems on Netflix, to auto complete on Google’s search bar, to predictive policing is accomplished not just through massive data collection but the more nuanced processes afterwards, when datasets are collected, aggregated and subjected to automated processing through machine learning algorithms. The scale and speed of computation is what makes targeted advertising seem so eerily accurate at times. In less obvious ways, algorithmically processed data influences many decision-making systems.

This wave also returns to the issue of identity that marked the first wave of internet studies. This time, beyond the performative aspects of identity made possible by digitalization and the capacities of the internet, which were the hallmark of earlier works, scholars are interested in how

³ Some of the best examples of this practice emerge from projects led by Daniel Miller, such as the 9-country ethnographic study of social media users, which is overviewed here: <https://www.ucl.ac.uk/why-we-post/>

people are identified, classified, and how our recommendation systems might not just reflect our identity but play a strong role in creating it. The number of recent book publications on topics of datafied or algorithmic identity demonstrates wide interest in this from a qualitative perspective (c.f., Cheney-Lippold, 2017; Lupton, 2019; Mau, 2019).

In some ways, this wave is heavily influenced by feminist technoscience, posthumanism, and new materialism. These schools of thoughts, represented by such scholars as Donna Haraway, Karen Barad, Rosi Braidotti, Jane Bennett, help digital researchers decenter the human and the individual, holding space for considering the role and agential characteristics of nonhuman, and more-than-human entities. While the scholarship in this arena is too broad to review here, I note significant work that focuses on the processes of becoming with technology (e.g., Svedmark, 2018); examines how the digital never stands alone but is always embedded in or entangled with the material (e.g., Pink, Ardevol, and Lanzeni, 2016; Thylstrup, 2019), critically interrogates how self-learning algorithms function independently as agents (Hayles, 2015). We need look only toward speaking personal assistants like Alexa or Siri to understand some of the ways in which agency can be broadened and balanced in research design in the digital age.

This most recent shift toward looking at the digital as a confluence of datafication, digitalization, predictive analytics, and machinic systems that think (or behave like they think) like humans loops us back to Negroponte's 1995 prescient insights that the world is fundamentally changed by the transformations from atoms to bits. For this reason, I have long resisted the use of 'digital' as an adjective, despite my situation as a researcher of what might be broadly labeled digital society. Because 'the digital' is too generic, I seek labels that are more specific. Thus, social contexts are *digitally saturated* rather than digital, to indicate they are heavily entwined or layered with digital features, regardless of whether the situation or context is actually online, offline, or a mix of both. Likewise, I have used phrases like *data-implicated* to specify contexts where data analytics are operating in a way that would be relevant, for example where newsfeeds are algorithmically sorted and presented, or where people experience numeric indicators in notifications on their mobile devices, or when people discuss something and seemingly, moments later, they receive advertisements related to that conversational topic.

Importantly, *one's adjectives to describe the scope or stance of one's research will likely be tied to what becomes salient or relevant in the situation, from both the researcher's perspective and from the characteristics of the context.*

Ethnographic research in the digital age has throughout these waves foregrounded the mediation of the internet on communication practice in specific contexts. Digital researchers in this domain label their research 'digital' not because they're using digital tools to collect or generate data, conducting interviews online rather than in person, or using automated software or digital tools to analyze data, although they might additionally do these things. Rather, they orient toward the interweaving, entanglement, or impact of the digital in interactions and relationships (e.g., Duguay, 2016; Tiidenberg, 2015; Tiidenberg, Markham & et al., 2017), sensemaking (Prieto and Schrieber, in press), everyday life and wellbeing (e.g., Pink, Postill, Hjorth & et al., 2015; Tania Lewis, 2019; Horst and Miller, 2012), sense of community (Orgad, 2006; Massanari, 2015); performance of identity

(Abidin, 2016; Boellstorff, 2008; Markham, 1998; Waskul, 2003); workplace (Hakim-Fernandez, in press), and labor (Baym, 2018; Gray & Suri, 2019).

PERSISTENT CHARACTERISTICS OF DIGITALITY

Physicality can be separated from sociality

The Internet is geographically dispersed, meaning people can congregate or have a sense of togetherness despite great distances. There is a simultaneous collapse of distance and an expansion of reach, something that Marshall McLuhan noted many decades ago when discussing how each new medium extends our senses, allowing us to see, listen, and reach well beyond our local sensory limits. More recently, Mirca Madianou develops the term “polymedia” (2016) as an all encompassing “concept of an environment of communication opportunities as an integrated structure of affordances” (Madianou & Miller, 2011, p. 172), whereby the possibilities for interpersonal communication are deeply interwoven with the affordances of digital media and their platforms as well as the overlaps and blurring of these communication environments. This is only one of many conceptual models attempting to adequately describe the multiple modes of being made possible through the simultaneity and ubiquity of digital media and technologies. Not only can we be experientially connected to situations far removed from our physical location, we can be engaged in multiple uniquely situated settings all at once.

What does it mean to ‘be with?’ Having a ‘sense of presence’ without actually being face-to-face is a hallmark of the Internet. Presence becomes a more complicated concept because it is determined by participation more than proximity. Meyrowitz discussed this as a separation of social from physical presence (1985). Waskul’s work (2005) takes this idea and flips our sensibility about what is passive or active in the creation of sociality. He writes that “places are transmitted from one locality to any and all users’ varied geographic ‘space’ (p. 55). Obviously, as the spring of 2020 is demonstrating for so many around the globe, sociality and a ‘sense of presence’ does not require physical presence.

Time/Temporality as a central and shifting factor

Internet-connected technologies for interaction can disrupt time, shifting it from a seemingly linear, forward moving, universal flow to a morphing and unstable variable in everyday interactions. Answering machines were once novelties whereby people could screen incoming calls, stopping the clock, so to speak, to delay an interaction. We now take for granted the ability to stop and start time in the midst of a digital conversation to consider and adjust our interactive choices. Most of us don’t notice that we are, in effect, manipulating time to suit our purposes. This is a relevant feature in contemporary social relations, facilitated by the fact that we are writing in digital formats that allow us to backspace and edit our writing but more to the point, send and receive in our own time zones, leaving messages to be picked up by others in their own time, picking up or slowing down the pace of

conversations for myriad reasons. Time and temporality may be relevant as an object of analysis, yet unnoticed because it is a seamless, hidden part of the infrastructures of digital life.

While sometimes we might feel we can control time or the tempo of conversations, time is not necessarily a controllable variable. It is simply a malleable construct. Time is often shifted, sped up, or disrupted in ways we cannot control and may not notice by the defaults or working protocols of the platforms we're using, the quality of our or others' internet or wifi connections, glitches in how information is being sent or received through networks. To what extent do these distortions or ruptures in time matter?

Data can develop a social life of its own

As we post, tweet, surf, search, and otherwise interact with others, we produce digital material that can be copied, archived, commented on, reposted. The qualities of persistence, replicability, scalability and reuse are salient characteristics of digital media. Remediation and reuse most obviously occurs in social networking sites, as Ellison and boyd note (2008), as an essential part of the ongoing flow of information exchange in such places as Facebook or Twitter or Wechat.

It also occurs in less obvious or automated ways. As individuals using phones, computers, or smart devices in our homes, wallets, cars, or workplaces, we leave massive data trails. Traces of our everyday actions and decisions are often swept up by nonhuman entities, who may archive, aggregate, sell, or pass along our data to other entities.

Once our words, uploads, mouse clicks, heartbeats, movements, or other human activities leave the body in ways that can be digitized or capturable in data form, they can develop a social life of their own. By this I mean they function on our behalf or represent us in other transactions. This can have mundane or profound impact.

The Internet is increasingly embedded and embodied⁴

For all intents and purposes, the internet has become embodied. We carry the internet with us and it is more fluidly attached to the accomplishment of everyday activities. Even if we don't have an obvious smart device in our hands, we use internet-connected cards on public transportation and at the cash register, connect to the internet through fitness or health trackers on our wrists. As the internet grows more mobile and embodied, it disappears as a visible infrastructure that governs how we make decisions, what we value, how we make sense of the world around us. Even if we're not online, we know that Google (Reddit, Safari, TikTok, Facebook, etc) is there, waiting in the background, just a tap or swipe away, ready to provide directions, maps, entertainment, and sociality.

The complication of the digital is also that it is embedded, embodied, and everyday (Hine, 2015), ubiquitous (Turkle, 2011; Deuze, 2012) and infrastructural (van Dijck, 2013; Bowker & Star, 1999). This creates an invisibility (Markham, 1998) whereby the digital actually disappears as an obvious intermediary and becomes simply a way of being. Normalized yet profoundly influential.

⁴ This notion is discussed in depth by Christine Hine (2015), who would add "everyday."

In many ways it can be said that to study the digital is to study society. This chapter exists as a subcategory of broader approaches because we still characterize this domain as something meaningfully separated from mainstream research. Perhaps the inextricable interweaving of the digital in how we make sense of and live everyday life in the midst of a pandemic signals the end of this separation once and for all. hereby scholars, policymakers, and citizens alike still struggle to make sense of technologies that are continuously becoming ever more central and essential to how we live everyday life. Like electricity studies 150 years ago, perhaps this phase of calling research 'digital' will pass and we will settle into the revised disciplinary boundaries for inquiry. But for now, we still need to consider how the internet has changed how we do social research.

CHALLENGES FOR THE PRACTICE OF QUALITATIVE RESEARCH

In what follows, I discuss some common problems for qualitative researchers in the digital age. Given the range of issues, the following represent only a small sample, common in my own practice of studying digitally-saturated or datafied, algorithmic social contexts. I focus on issues in four stages of inquiry: research design, fieldwork, data management, and interpretation/representation

Research design

Making, not finding, the boundaries of the field. How do ethnographic or so researchers make boundaries around what they study? Christine Hine emphasized in her early work that boundaries were matters of negotiation, constructed as the researcher moves and encounters the field (2000). The researcher plays no small role in creating the boundaries of the field through their own actions and attention. This may have always been the case, despite the seeming solid foundations of cultural borders for traditional anthropologists. Still, it can be daunting to realize that the boundaries of your field keep moving.

Pick any scenario of digitally-saturated social life. Any piece of digital material or data. If we begin to trace the origins or flow of digital media, we will swiftly realize that there are no actual or clear boundaries. I might notice a swell of interest in all things Japanese in the weeks following the earthquake in 2011. I start to take screenshots and collect posts, images, texts, comments on Facebook. Should Facebook be the boundary? Or reposts? Or a single post with hundreds of comments? These questions illustrate that the boundaries of the field are based on decisions I must make. Even asking a seemingly narrow research question may not be an adequate way to handle the complexity of most digital situations. I might ask: "How are Westerners depicting Japan in their sharing of visual imagery of the Japanese earthquake on Facebook?" Even with this narrowing of my RQ, the situation is impossibly large, with between 600-800 million users (at the time).

I can use the advice of Marcus (1998) or Latour (2005), both of whom advocate the method to "follow" something, whether it's the story or the data. By following a single video from a single post, I shift immediately from where it is posted (Facebook) to where it was originally uploaded on YouTube. Is the boundary the edges of the video? The edges of the post, the edges of the path it takes as it

travels through various networks? The edges of the screen of YouTube, where other 'recommended' videos are posted and an entirely different set of comments and interactions exist?

This example is intended to simply demonstrate that first, no matter the choice I make, I have to recognize that boundaries are not preformulated. They emerge as our questions are refined. And second, looking for a boundary may not be helpful, methodologically speaking, since this is likely not where the narrowing is most effective. This leads to the next challenge in research design.

Finding the (earliest) origin point of inquiry: Many research projects emerge from our experiences and interests. We live in digitally-saturated and internet-mediated societies, whether or not we think we are heavy users of digital media. Unless the research subject is completely foreign, the researcher has likely been immersed in or studied intensively the phenomenon for a long time before they started calling it 'research.' This means the emergent and inductive processes of inquiry have been ongoing. Recognizing this will help researchers situate themselves and the study itself. Likely, the study will benefit from some backtracking, memory mining, and reflexive consideration of previous unaccounted-for data gathering, analyses, interpretations, and otherwise building presumptuous boundaries around a field or research object. How much of these prior moments have been included as fieldwork and what might be necessary to recognize and then defamiliarize one's previous assumptions and knowledge about 'what is going on here?' This is particularly characteristic of digital research since digital media are embedded, ubiquitous, and nearly invisible infrastructures in our everyday lives, yet digital media still seems a foreign or novel thing to study.

Fieldwork. Or more specifically, the challenge of rethinking what counts as fieldwork.

Many research projects include what ethnographers call fieldwork, where the researcher immerses in the location of the situation to take an emic perspective, understand local meanings, explore what lies underneath or motivates cultural behaviors and norms and get closer to what Clifford Geertz (1973) would call "thick description," or find out, to put it in Marcus's (1998) deceptively simple terminology, "what is going on here." Immersion has long included some level of participation in common activities and key rituals of communities and cultures.

What counts as the field? This is not always obvious, or in any case, should not be taken for granted. What are we paying attention to when we scroll through Twitter or Instagram? What is relevant, salient, or meaningful differs from person to person, depending on numerous factors. Some might be paying attention to hashtags whereas others might find the person more important than the content. What does it even mean to conduct something like "fieldwork" or "participant observation" on Twitter? What prompts us to focus on certain individuals and seek them out for interviews? There are even more complications once we decide who to focus on, as we might want (or need) to translate a traditional interview to something more suited to the context, e.g., in digital platforms where people are anonymous or where communities are too vast to identify representatives who should be targeted for more in depth conversations. This issue can be addressed by going back to the origins of why people did interviews or observations in the first place. As I suggest elsewhere (Markham, 2013b), we might ask: What useful material did anthropologists obtain from certain interactions or behaviors in the field, which later were standardized as interviews, gaining informants, observation, taking fieldnotes,

photography, and participant observation? Significant power rests in the hands of researchers who make these decisions, since they are essentially deciding what “counts” as the field, and therefore what is ignored as “not the field.”

The second issue of rethinking fieldwork is to reconsider: *Who comprises the field?* If following Burrell (2009), we can see many challenges, or opportunities to creatively design research when the network is the field. This means that wherever you dip into the network, the field begins. This returns us to the matter of making boundaries. This also raises an often-neglected element of fieldwork: acknowledging and attending to the self as a participant. Researchers (especially if they’re new to the game) often fail to consider that they are an essential participant who should be observed and interviewed. To include oneself as a relevant party making meaning within the boundaries of the field is a hallmark of the interpretive sociological position⁵. The stance taken by many interpretive scholars --and, of course, autoethnographers-- to reflect on the role of the researcher is an important consideration, but at a more basic level, the point is that the researcher may hold valuable knowledge, insights, and experiences that should not be omitted from the study. This is a simple acknowledgement of the point made in the previous paragraphs, that one’s study of the phenomenon likely began long before it was called a study.

As a case in point, a colleague spent several years studying people whose physical exercise relied on self-quantified tracking tools and was frustrated that the vignettes of his participants were sometimes unsatisfactory, not rich enough. Once he included himself as a participant to be interviewed, he could explore and utilize the depth of experience he had with self tracking apps and exercise, which was the very reason he was interested in the topic to begin with. Acknowledging his expertise and long term personal immersion in this topic was a relief. He then interviewed himself, had someone else interview him to get a different perspective, and mined his research diary for evidence of his perceptions and insights.

Data management (of too much material)

We can collect an incredible amount of material in the field, which leads to datasets that are unwieldy, daunting, and impossible to fully analyze from a hands-on approach. To be fair, narrowing one’s focus is important in any qualitative research, since even the seemingly tiny projects produce innumerable possibilities for inquiry. This is what prompts UK sociologist David Silverman to insist: “if you think your project is too small, make it smaller.” Still, the past 20 years have proven that despite this good advice, ethnographers and sociologists are every bit as daunted as computational scientists by datasets that are enormously large.

In addition to collecting archives of what people post in written and image form, we could-- ethics aside for a moment-- collect a person’s eye movements, track where they go online, how long they spend on a site, how long they seem to linger on certain content, images, or posts appearing on

⁵ Too many scholars have written about this to include them all. I note only two: The conceptual features of this positionality are outlined by scholars in the classic edited collection *Writing Culture* (Clifford & Marcus, 1986); the difference in focus and outcomes of research is illustrated vividly by Margery Wolff in her book *A Thrice-Told Tale* (1991).

their screens, how long they watch or listen before moving on to something else. We can collect information about a person's typing and backspacing tendencies, witnessing how they construct, edit, delete messages. This list could be endless.

Even if we don't invade a single individual's data producing actions to that degree, it is tempting to gather much more than is needed, simply because one can. It is easy to scrape and archive hundreds of thousands of comments around a single event. But it is also an endless rabbit hole when we start to collect the materials generated by human and nonhuman actors that also influence the phenomenon, directly and indirectly. Of course, this can involve obvious stakeholders in ever wider spheres, which is a standard problem in collecting and managing relevant material in any qualitative study. But it can also include nonobvious, behind the scenes activities and entities, nonhuman agents with great deal of influence but not visible in what lies on the surface of our observing gaze. Here, I'm referring to metadata located in the code or tags; default settings that guide actions in browsers or apps; protocols on platforms that delimit how things look, appear, and disappear; lines of code that invoke an algorithm, which influences whether or not an advertisement will appear on a person's instagram feed; information on the corporate links of all the cookies planted when someone clicks on a supposed 'not for profit' website. There are endless factors, all part of the ecology of 'the digital', all potentially relevant from an ethnographic perspective, especially if we're seeking thick meaning versus surface level description. There are therefore endless potentially relevant data points that might become useful, meaningful, necessary. This is a trap even for senior scholars, but particularly plagues junior scholars, who often feel they must 'explain' the whole.

The answer to this conundrum is relatively straightforward to say but difficult to enact. Make a general question to guide the direction of your analytical eye. Pick any place or point to begin, explore and immerse yourself, and patterns will emerge, from which you can create more narrow research questions. Movement is a key to all research, whether we conceptualize this as entering and moving through a site to explore cultural meaning, being moved by the phenomenon, shifting from one perspective to another, or adopting a stance or positionality. Importantly, not only are we moving, the sites, phenomena, and meanings are also in continual motion.

What I have earlier called a "network sensibility" (as separate from a network analysis, cf, Markham & Lindgren, 2012) acknowledges that whatever we explore in a social realm in the digital age is likely to be composed of materials and media that are largely already remediated. Our own collection techniques are fraught with the challenges of temporality (we rarely will analyze a cultural phenomenon in the same way, pace, or time it was generated or experienced). Even if collected in so-called 'real time', the experiential stuff of cultural experience is translated into a format that can be analyzed. Since it is no longer in the flow, or networked, it is in some way necessarily removed from the context in which it was generated⁶. For researchers new to this domain, it is a key issue to highlight. These mediation processes comprise an important aspect of anything we discuss as digital.

Once we let go of the idea that we can capture it all, the process of collecting can be reframed as one of exploring widely to find some good questions, and continuing to narrow one's

⁶ There may be some minor exceptions to this, especially when one is immersed deeply in the field one is studying and can therefore experience the phenomenon as it happens.

focus until the question is clear enough to drive a more clear choice of material to work with, which then can be enacted through a narrow and purposive sampling of material to analyze, drawn from the larger collection of stuff one has collected.

Analysis / Interpretation

What counts as data? Once we've reached the stage of analysis, the researcher must make some choices about what should be analyzed and what can be set aside or disregarded as 'not data.' This series of choices is not simply logistic, but epistemological. In writing about this previously, I note that:

Obviously, we cannot pay attention to everything—our analytical lens is limited by what we are drawn to, what we are trained to attend to, and what we want to find. Borrowing from Goffman (1967), our understanding is determined as much by our own frames of reference as the frames supplied by the context. Our selection of data and rejection of non-data presents a critical juncture within which to interrogate the possible consequences of our choices on the representation of others through our research. (Markham 2005, p. 803).

To illustrate the implications of this I draw on an example from my early ethnography of self-described heavy users of the internet (Markham, 1998). To clarify what you'll see below, the interview occurred in a MOO⁷, an online text-only environment where one's writing often includes both content and actions. By inputting various commands or using particular punctuation, the computer would add information that would clarify that a person was "speaking", "exclaiming", "questioning", "whispering," "thinking," "falling," "rolling on the floor laughing," and so forth. The computer would automatically correct the written format, so it would appear on the others' screens as verbal statements in the form of, for example:

```
Annette says, "Hi."  
Annette exclaims, "Hi!"  
Annette asks "Hi?"
```

Alternately, one's actions could simply be a description of one's nonverbal behaviors or thoughts:

```
Annette scratches her head thoughtfully  
Annette . o O ( I wonder if the reader realizes this represents a thought bubble )
```

In my study, I decided to focus only on the main content of what people said, to cut down the size of the dataset, which included thousands of pages of interviews, transcripts of actions, conversation threads, descriptions of places in text-based open spaces online, ascii art inserted in descriptions or by users in conversations, and logs of group events. This seemed sensible, as my research question asked something like "How do people make sense of the internet through how they define or frame it in everyday life online?" One way to answer this question is to use their words to analyze their definitions or frames.

⁷ A close relative of MUDs, a MOO is a Multi User Dimension of the Object Oriented sort. There were also MUCKs and MUSHs.

The excerpt below exemplifies my “cleaned-up” dataset, where I have removed from Matthew’s interview some of the extraneous, repetitive, or what might be called metadata. This is what I used as the basis for analysis:

Markham: “okay here’s some official stuff for you Matthew.”
Markham: “I guarantee that I will not ever reveal your address/name/location.”
Matthew: “Fine about the secrecy stuff.”
Markham: “Matthew, I guarantee that I will delete any references that might give a reader clues about where you live, who you are, or where you work.”
Markham: “do you mind that I archive this interview?”
Matthew: “Log away,Annette” . . .
Markham: “what do you do mostly when you’re online? Where do you go?”
Matthew: “Mostly I’m doing one of two things. Firstly I do research. If I’m looking for academic research in software engineering, my specialty, a lot of it is on the Web . . .”
Matthew: “And a lot of tools to play with are there, too.”
Matthew: “Also, I use it for news and information, the way I used to use the radio.(I’m an unrepentant . . .”
Matthew: “real-lifer). For instance, if I’m going to go run (or bike or do something else outside) . . .”
Matthew: “I check the weather on the Web when in years past I would turn on the radio. Ditto for news” . . .
Markham: “how would you compare your sense of self as a person online to your sense of self offline?”
Matthew: “More confident online, because I’m a better editor than writer/speaker. I do well when I can backspace.”
Matthew: “But I’m the same me in both places. I guess I’ve been me too long to be anybody else without a lot more practice than I have time for.”
Markham: “hmmm . . . How would you describe your self?”
Markham: “i mean,what’s the ‘me’ you’re talking about?”
Matthew: “Kind of androgenous. Plenty of women for friends. But I was never good at dating or any of the romantic/sexual stuff.”
Matthew: “Also, somewhat intellectual.”
Matthew: has a delayed blushing reaction to the androgeny comment.
Matthew: “And a fitness nut.”
Markham: o O (I wonder why Matthew is blushing . . .)
Markham: “tell me about your most memorable online experience”
Matthew: “OK, it was a couple years ago and I was just getting on the Web and starting to realize all”

Although I had been intrigued by Matthew’s descriptions in this interview, I couldn’t regain that sense of wonder or curiosity as I analyzed this interview some months later. Why was it so uninteresting? I couldn’t figure out why but dismissed it as a faulty memory of the moment. Later, for some reason, I was rummaging around in the original data files and re-read the originally-archived interview. As you read the original below, notice the boldface items that I had deleted as ‘non relevant’ data:

Markham says, “okay here’s some official stuff for you Matthew.”
Matthew spills popcorncrumbs into his keyboard :- (
Markham says, “bummer,Matthew.”
Matthew says, “If you see me going away for a while, you know I went to make more popcorn ;-)”
Markham says, “okay here’s some official stuff for you Matthew.”
Markham says, “I guarantee that I will not ever reveal your address/name/location.”
Matthew says, “Fine about the secrecy stuff.”

Markham says, "Matthew, I guarantee that I will delete any references that might give a reader clues about where you live, who you are, or where you work."

Markham asks, "do you mind that I archive this interview?"

Matthew salutes and says "Yes'm

Matthew says, "Log away, Annette"

Markham says, "okay. i have a tendency to ask questions too quickly."

Matthew doesn't answer because he's too busy opening a box of rice cakes. . . .

Markham asks, "what do you do mostly when you're online? Where do you go?"

Matthew says, "Mostly I'm doing one of two things. Firstly I do research. If I'm looking for academic research in software engineering, my specialty, a lot of it is on the Web . . . "

Matthew says, "And a lot of tools to play with are there, too."

Matthew says, "Also, I use it for news and information, the way I used to use the radio. (I'm an unrepentant . . . "

Matthew says, "real-lifer). For instance, if I'm going to go run (or bike or do something else outside) . . . "

Matthew says, "I check the weather on the Web when in years past I would turn on the radio. Ditto for news" . . .

Markham asks, "how would you compare your sense of self as a person online to your sense of self offline?"

Matthew says, "More confident online, because I'm a better editor than writer/speaker. I do well when I can backspace."

Matthew says, "But I'm the same me in both places. I guess I've been me too long to be anybody else without a lot more practice than I have time for."

Markham asks, "hmmm . . . How would you describe your self?"

Markham asks, "i mean, what's the 'me' you're talking about?"

Matthew says, "Kind of androgenous. Plenty of women for friends. But I was never good at dating or any of the romantic/sexual stuff."

Matthew says, "Also, somewhat intellectual."

Matthew says, has a delayed blushing reaction to the androgeny comment.

Matthew says, "And a fitness nut."

Markham . o O (I wonder why Matthew is blushing . . .)

Matthew does pushups.

Markham stares

Markham . o O (should I be doing something too?)

Matthew says, "You should be asking me questions (the interviewee becomes the interviewer)"

Markham sighs and refocuses

Markham says, "tell me about your most memorable online experience"

Matthew gets very jealous of people who have sleep.

Matthew enters state of deep thought.

Matthew goes to raid the nearby refrigerator while composing reply in head

Matthew says, "OK, it was a couple years ago and I was just getting on the Web and starting to realize all"

A reader might simply say I missed the obvious. Or remark that while I noticed and commented on Matthew's embodied elements, I don't follow this trail in the original interview. Both are reasonable critiques of the process of elicitation. But here, I point to the later problem of cleaning up data, when a researcher applies some schema, or what we might now call algorithmic formulae, to decide what is relevant and irrelevant. This *cut* into the data is necessary for all analysis. And within these decisions to edit in and edit out there is the potential impact of ignoring essential elements of experience. Of course, this problem is somewhat the opposite of the challenge of managing too-large datasets and seems to give the opposite advice. However, it's more an illustration of how analysis is always a matter of choosing, editing and cutting what counts as data even before the analysis begins.

This poses the question of how much in a datafied era should researchers rely on only the scraped data, or secondary datasets to guide the analysis? Of course, empirical social science, interpretive or otherwise, is data driven, but our formulation of what counts as data is also guided by our current questions. And even if we understand conceptually that in qualitative research, we should not be method-driven, we often end up using the tools with which we have the greatest familiarity.

It is worth returning to one of the most fundamental aspects of the research process: Asking ourselves why we're doing research in the first place and how the material (data) gathered (generated) fits into that overall goal. Our research questions are attuned, whether or not we recognize it ourselves, to our research goals. Here, at the point of origin, or considering the reason we're doing the research in the first place, the researcher can rethink their analytical moves by returning to the intended outcome of inquiry. Is it to describe, understand, explain? Is this with the intention of generalizing? Or are we trying to spark social change, co-create meaning with collaborators, tell one of many possible stories about "what is going on here?"

Qualitative knowledge emerges from immersion and a somewhat open-ended sensibility, which enables findings --and methods, for that matter-- to be emergent. This is an ideal, more a goal than an actual possibility. It requires us to be reflexive about what we are actually paying attention to in our analysis of a situation. In the case of Matthew above, I was guided by method--that is, tools of analysis that focused on language and utterances. I therefore didn't pay as much attention to visual rhetoric, or how much of what was being 'said' was actually enacted, built into the environment and how I moved through it. Thus, no matter how much I thought I was immersed and the methods were emergent, my actions in cleaning the data were the first steps of analysis disguised as a matter of efficiency to facilitate the method. So even for a scholar well trained in interpretive ethnography, retaining a truly open-ended sensibility is difficult, if not impossible.

It would be tempting to throw up one's hands in frustration at this point, recognizing that no matter what you do, you'll miss, ignore, or forget something. However, it's useful to remember that this is always the case in all forms of inquiry, qualitative, academic, or otherwise.

Studying cultural phenomena in the digital age is at its best a matter of using multiple tools, perspectives, and sensibilities to find patterns, create pathways to meaning, and make sense of situations. This embraces the idea that the researcher is ultimately the primary lens, the body through which the phenomenon passes. Thick description may be closer when we know the field/data/material so well that we are able to let go of the idea that the data tells us everything, that everything that is important is in the data, or that the phenomenon is limited to what we can read in/through the data.

This sensibility, Baym notes (2017), can help counteract the current trend to remove the human from the process of analysis when dealing with large or complex datasets, or more insidiously, let machine learning conduct the equivalent of qualitative analysis for us.

The age of datafication and automated data analytics is also an age of erasure, even as it has produced amounts of data and eerily accurate predictive analytics. Today's technologies for automated capture often leave important things (and people) out of the datasets. The critical problems of automated machine learning systems are now well documented: People of color are missing from datasets used to build facial recognition technologies, basic data used to train machine learning AI is

misclassified, algorithms are biased, data modeling is flawed, and automated systems for curating online content make decisions on our behalf in invisible and immediate ways (for extensive elaboration, see Buolamwini, 2017; Eubanks, 2018; Gillespie, 2018; Noble, 2018; O'Neil, 2016).

To be clear, data have always been misrepresentative. Not only are they abstractions from actuality, as Latour pointed out in discussing dirt in the Amazon. They are partial samples of experience, never representing the entirety of what happens (Baym, 2013; Gitelman, 2012). They flatten and unify experience into comparable units labeled 'data', which stand in for reality. This re-figuring creates an illusion of completeness, that we have captured all there is, or that it is possible to capture what is there (Markham & Gammelby, 2017). This problem of representation is exacerbated by the sheer size and scope of datasets that give us innumerable possible things to see.

Indeed, the role of the interpretive qualitative researcher has never been more important. Each step in analyzing data is a choice about what remains, what is relevant, what is meaningful. We are entrusted to this role of interpreter. And while it's easy to make glaring mistakes, as my own research example demonstrates, the strength of the qualitative approach in a digital age is to understand what it is we're working with in our analytical stages, acknowledge the potential consequences of choices that will naturally take us down certain pathways to meaning and simultaneously cut off other possible paths, and continue to work toward depth of interpretation.

Conclusion: A note on Ethics

I conclude this chapter by addressing ethics as more than just a regulatory matter of making sure we protect participants or personal data. As I have written elsewhere, ethics is fundamentally an emergent quality of research, and the ethical principles and practices emerge as one engages in a series of choices at critical junctures, or as we make judgments in contexts of complexity and uncertainty. An ethic of care, for example, begins with a stance that anything we do in the name of research should be directly relevant to the accomplishment of care in a larger community, perhaps the same as the community we are studying. This may be an a priori stance, an attitude that guides our practice, but it is far better understood as the ethic not only emerges as we practice it, but is measured or assessed after the fact, as we make sense of what choices we made, or as we consider the outcomes of our research in retrospect. Knowing this, we might consider taking a stance that embraces this future-orientation. For example, examining our work through the lens of short- and long-term impact (Markham, 2018), rather than just whether or not it adheres to legal or regulatory policies, we can think about a study less as just harvesting other people's experiences for our goals to instead design and conduct engagements that enact social change with and in communities. This is an ethic that focuses somewhat less on predetermining what might cause harm--which is an important consideration of course, but only one of many--and more on future benefits and harms. This is a speculative ethic, since the impacts of our actions can only be understood in retrospect, after they are felt and known by people or societies in the future.

In other words, while adhering to extant guidelines, regional or disciplinary norms, or legal requirements is important, it is not enough to simply follow the rules. Because regulatory guidelines have tended to restrict experimentation and create only low risk research projects that won't be

flagged by institutional ethics review committees as problematic. This doesn't mean researchers should ignore regulations, but take additional proactive measures. This impact model of ethics interrogates how one's research feeds into the shape and characteristics of various possible futures. What sort of futures we want to work toward? And what sort of inquiry would these better futures perhaps require? Far from eliminating regulation, a proactive stance or 'beyond regulations' ethics can actually serve a regulatory function in that:

The anticipatory analysis modulates decision making. Rather than the regulation being asserted from the outside—by institutional or historical contexts, it is derived continuously and iteratively. This more internal form of regulation can be more actively attuned to the situation, as it is an ongoing activity. This not only transforms the regulatory function from the noun to verb form (regulation to regulate) but also distributes agency for ethics, or “doing the right thing,” in a more balanced way. (Markham, 2018, np).

Moving beyond self-regulation and avoidance of harm, an impact model for ethics encourages us to think not only about what we should *not* do, but what we *should and can* do. Especially in this era of global trauma. There is a powerful ethic in creating and influencing social change through research premises, practices, choice of unit of analysis, type of argument, form of writing, dissemination of findings, and politics of engagement. The shifts that we see over time in digital social research grow more critical, more daring. More inclined toward sparking social change, research in this domain often seeks to be critical and interventionist. This is because in a world that is increasingly mediated by the digital, data-driven, and automated, the stakes are high. For individuals, communities, and the planet itself. Today's most highly-regarded social researchers seek to directly confront and decolonize social, technological, and cultural systems and stakeholders that are associated with racism, surveillance capitalism, ecological destruction, and the like. We live in troubled times and now more than ever, the paramount ethic is to make a difference. And returning to the idea of an ethic of care, this would suggest that if we are not working collaboratively with people in the communities we study, we might wonder why not, and reflect critically on what is driving our inquiry practices.

The speed of technological development in robotics, automation, machine learning, nanotechnology, and biology are transforming humanity. We face these advancements in the midst of rapid climate change, which creates a situation whereby many of our human advancements are unsustainable, or creating unsustainable futures. This raises the bar for social scientists in this domain, and demands that we continue to ask tough questions: Why are we doing our research in the first place? How does our research genuinely contribute to solving the big problems we face in the world? These questions invite us to think more radically about the possibilities for what research means as well as how it is accomplished.

WORKS CITED:

- Abidin, C. (2016). “Aren't These Just Young, Rich Women Doing Vain Things Online?": Influencer Selfies as Subversive Frivolity. *Social Media + Society*, 1-17.
<https://doi.org/10.1177/2056305116641342>

- Barlow, J.P. (1996). *A Declaration of the Independence of Cyberspace*. Available from <https://www.eff.org/cyberspace-independence>
- Baym, N. (2013). Data not seen: The uses and shortcomings of social media metrics. *First Monday*, 18(10). doi:10.5210/ fm.v18i10.4873
- Baym, N. (2017). The Challenges and Benefits of Influencing Human Technological Futures. *Plenary presentation at the 13th annual International Congress of Qualitative Inquiry*, May 19, Urbana, Illinois.
- Baym, N. K. (2018). *Playing to the Crowd: Musicians, Audiences, and the Intimate Work of Connection*. NYU Press.
- Boellstorff, T. (2008). *Coming of age in second life: An anthropologist explores the virtually human*. Princeton University Press.
- Boellstorff, T. (2013). Making big data, in theory. *First Monday*, 18. doi:10. 5210/fm.v18i10.4869
- Bowker, G. C., & Star, S. L. (1999). *Sorting Things Out: Classification and Its Consequences*. MIT Press.
- boyd, d., & Ellison, N. (2008). Social Network Sites: Definition, History, and Scholarship. *Journal of Computer-Mediated Communication*, 13, 210-230. <http://dx.doi.org/10.1111/j.1083-6101.2007.00393.x>
- boyd, d., & Crawford, K. (2012). Critical questions for big data. *Information, Communication & Society*, 15, 662-679.
- Brock, A. (2015). Deeper data: a response to boyd and Crawford. *Media, Culture & Society*, 37(7), 1084–1088. <https://doi.org/10.1177/0163443715594105>
- Bruns, A. (2008) *Blogs, Wikipedia, Second Life, and Beyond: From Production to Produsage*. Peter Lang.
- Bruns, A., Burgess, J., Crawford, K., & Shaw, F. (2012). *#qldfloods and @QPSMedia: Crisis communication on Twitter in the 2011 South East Queensland floods - Research Report*. Brisbane, AU: ARC Centre of Excellence for Creative Industries and Innovation.
- Bruns, A., & Highfield, T. (2012). Blogs, Twitter, and breaking news: The produsage of citizen journalism. In R. A. Lind (Ed.), *Produsing Theory in a Digital World: The Intersection of Audiences and Production in Contemporary Theory* (pp. 15-32). Peter Lang Publishing.
- Buolamwini, J. A. (2017). *Gender shades: intersectional phenotypic and demographic evaluation of face datasets and gender classifiers* (Dissertation). MIT.
- Burrell, J. (2009). The Field Site as a Network: A Strategy for Locating Ethnographic Research. *Field Methods*, 21(2), 181-199. DOI: 10.1177/1525822X08329699
- Cheney-Lippold, J. (2017). *We Are Data: Algorithms and The Making of Our Digital Selves*. NYU Press.

- Cherny, L., & Weiss, E. (1996). *Wired women : Gender and new realities in cyberspace*. Seal Press.
- Clifford, J., & Marcus, G. E. (1986). *Writing Culture: The Poetics and Politics of Ethnography*. University of California Press.
- Danet, B. (2001). *Cyberpl@y: Communicating Online*. Bloomsbury Academic.
- Deuze, Mark (2012) *Media Life*. Polity.
- Dibbell, J. (1993). A rape in cyberspace: Or, how an evil clown, a Haitian trickster spirit, two wizards, and a cast of dozens turned a database into a society. *The Village Voice*, 21 (December), 36-42.
- Duguay, S. (2016). Dressing up Cinderella: Interrogating authenticity claims on the mobile dating app Tinder. *Information, Communication, & Society*, 20(3).
<https://doi.org/10.1080/1369118X.2016.1168471>
- Eubanks, V. (2018). *Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor*. St. Martin's Press.
- Geertz, C. (2013). *The interpretation of cultures*. Basic Books.
- Gergen, K. (2002). The challenge of absent presence. In J. Katz & M. Aakhus (Eds.), *Perpetual Contact: Mobile Communication, Private Talk, Public Performance* (pp. 227-241). Cambridge University Press.
- Gillespie, T. (2018). *Custodians of the Internet: Platforms, Content Moderation, and the Hidden Decisions That Shape Social Media*. Yale Press.
- Gitelman, L. (Ed.). (2013). *"Raw data" is an oxymoron*. MIT Press.
- Goffman, E. (1967). *Interaction ritual: Essays in face-to-face behavior*. Aldine
- Gray, M. L., & Suri, S. (2019). *Ghost Work: How to Stop Silicon Valley from Building a New Global Underclass*. Houghton Mifflin Harcourt.
- Grinter, B. (2013). A big data confession. *Interactions*, 20(4), 10–11.
- Hakim-Fernandez, N. (in press). Workplace-making among mobile freelancers. In A. Markham, & K. Tiidenberg (Eds.), *Metaphors of the internet: Ways of being in the age of ubiquity*. Peter Lang.
- Hayles, K. (2016). Cognitive assemblages: Technical agency and human interactions. *Critical Inquiry* 43(1), 32-55.
- Hermida, A. (2010). From TV to Twitter: How Ambient News Became Ambient Journalism. *Media-Culture Journal*, 13(2). <http://journal.media-culture.org.au/index.php/mcjournal/article/view/220>
- Hine, C. (2000). *Virtual ethnography*. Sage.
- Hine, C. (2015). *Ethnography for the Internet: Embedded, Embodied and Everyday*. Bloomsbury Academic.

- Horst, H. A., & Miller, D. (2012). The Digital and the Human: A prospectus for Digital Anthropology. In H. A. Horst, & D. Miller Digital (Eds.), *Digital Anthropology* (pp. 3-35). Berg.
- Jenkins, H. (2006). *Convergence Culture: Where Old and New Media Collide*. New York University Press.
- Jones, S. G. (Ed.) (1995). *Cybersociety: Computer-mediated communication and community*. Sage.
- Jones, S. G. (Ed.). (1999). *Doing Internet Research: Critical Issues and Methods for Examining the Net*. Sage.
- Lewis, T. (2020). *Digital Food: From Paddock to Platform*. Bloomsbury Academic.
- Madianou, M. (2016). Ambient co-presence: transnational family practices in polymedia environments. *Global Networks*, 16(2), 183-201.
- Madianou, M. & Miller, D. (2011). *Migration and New Media: Transnational Families and Polymedia*. Routledge.
- Marcus, George E. (1998). *Ethnography Through Thick and Thin*. Princeton University Press.
- Markham, A. (1998). *Life Online: Researching real experiences in virtual space*. AltaMira Press.
- Markham, A. (2005). The politics, ethics, and methods of representation in online ethnography. In N. Denzin, & Y. Lincoln (Eds.), *Handbook of Qualitative Research* (3rd ed., pp. 793-820). Sage.
- Markham, A. (2006). Method as ethic, ethic as method. *Journal of Information Ethics*, 15(2), 37-55.
- Markham, A. N. (2013a). Undermining 'data': A critical examination of a core term in scientific inquiry. *First Monday*, 18(10). DOI:10.5210/fm.v18i10.4868.
- Markham, A. N. (2013b). Fieldwork in social media: What would Malinowski do?. *Journal of Qualitative Communication Research*, 2(4), 434-446. DOI: 10.1525/qcr.2013.2.4.434.
- Markham, A. N., & Lindgren, S. (2014). From object to flow: Network sensibilities, symbolic interactionism, and social media. *Studies in Symbolic Interactionism*, 43, 7-41. Initially published in 2012 via SSRN.
- Markham, A. N., & Gammelby, A. K. (2017). Moving through digital flows: An epistemological and practical approach. In U. Flick (Ed.). *Handbook of Qualitative Data Collection* (pp. 451-465). Sage.
- Markham, A. N. (2018). Afterword: Ethics as Impact—Moving from error-avoidance and concept-driven models to a future-oriented approach. *Social Media + Society*, 4(3), 1-12. Available from: <http://journals.sagepub.com/doi/full/10.1177/2056305118784504>
- Marwick, A. E., & boyd, d. (2010). I Tweet Honestly, I Tweet Passionately: Twitter Users, Context Collapse, and the Imagined Audience. *New Media & Society*, 13(1), 114-133. <https://doi.org/10.1177/1461444810365313>

- Massanari, A. L. (2015). *Participatory Culture, Community, and Play: Learning from Reddit*. Peter Lang.
- Mau, S. (2019). *The Metric Society: On the Quantification of the Social*. Polity Press.
- Meyrowitz, J. (1985). *No Sense of Place: The Impact of Electronic Media on Social Behavior*. Oxford University Press.
- Miller, D. (2011). *Tales from Facebook*. Polity.
- Murthy, D. (2011). Twitter: Microphone for the Masses?. *Media, Culture and Society* 33(5), 779–89. doi:10.1177/0163443711404744.
- Negroponte, N. (1995). *Being Digital*. Alfred A. Knopf, Inc.
- Noble, S. U. (2018). *Algorithms of Oppression: How Search Engines Reinforce Racism*. NYU Press.
- Latour, B. (2005). *Reassembling the Social: An Introduction to Actor-Network-Theory*. Oxford University Press. (Follow the data?)
- Lupton, D. (2019). *Data Selves: More-than-Human Perspectives*. Polity Press.
- O'Neil, C. (2016). *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. Crown Books.
- Orgad, S. (2006). *Storytelling online: Talking breast cancer on the internet*. Peter Lang.
- Papacharissi, Z., & Oliveira, M. (2012). Affective News and Networked Publics: The Rhythms of News Storytelling on #Egypt. *Journal of Communication*, 62(2), 266–282. <https://doi.org/10.1111/j.1460-2466.2012.01630.x>
- Pink, S., Ardevol, E., & Lanzeni, D. (Eds.). (2016). *Digital Materialities: Design and Anthropology*. Bloomsbury Academic.
- Pink, S., Horst, H., Postill, J., Hjorth, L., Lewis, T., & Tacchi, J. (2015). *Digital Ethnography: Principles and Practices*. Sage.
- Rheingold, H. (1991). *The Virtual Community: Homesteading on the electronic frontier*. Addison-Wesley.
- Rheingold, H. (2002). *Smart Mobs: The Next Social Revolution*. Basic Books.
- Rogers, R. (2013). *Digital Method*. The MIT Press.
- Rossini, P., Hemsley, J., Tanupabrunsun, S., Zhang, F., & Stromer-Galley, J. (2018). Social Media, Opinion Polls, and the Use of Persuasive Messages During the 2016 US Election Primaries. *Social Media + Society*. <https://doi.org/10.1177/2056305118784774>
- Schreiber, M., & Prieto-Blanco, P. (forthcoming). Co-becoming hybrid entities through collaboration. In A. Markham, & K. Tiidenberg (Eds.), *Metaphors of the internet: Ways of being in the age of ubiquity*. Peter Lang.

- Sunden, J. (2003). *Material Virtualities: Approaching Online Textual Embodiment*. Peter Lang.
- Svedmark, E. (2018). *Becoming Together and Apart: Technoemotions and other posthuman entanglements*. Dissertation. Umea University.
- Tiidenberg, K. (2015). Great faith in surfaces – A visual narrative analysis of selfies. In A. Allaste, & K. Tiidenberg (Eds.). *"In Search of ..." New Methodological Approaches to Youth Research* (233-256). Cambridge Scholars Publishing
- Tiidenberg, K., Markham, A. N., Pereira, G., Rehder, M., Sommer, J., Dremljuga, R., & Dougherty, M. (2017). "I'm an Addict" and Other Sensemaking Devices: A Discourse Analysis of Self-Reflections on Lived Experience of Social Media. *Proceedings of the 8th International Conference on Social Media & Society, Article No. 21*. Retrieved from <http://dl.acm.org/citation.cfm?doid=3097286.3097307>
- Thystrup, N. B. (2019). *The Politics of Mass Digitization*. MIT Press.
- Turkle, S. (2011). *Alone Together: Why We Expect More from Technology and Less from Each Other*. Basic Books.
- van Dijck, J. (2013). *The Culture of Connectivity*. Oxford University Press.
- Vis, F. (2012). Twitter as a reporting tool for breaking news. *Digital Journalism*, 1(1), 27-47. <https://doi.org/10.1080/21670811.2012.741316>
- Walker, J. R. (2008). *Blogging*. Polity.
- Waskul, D. (2003). *Self-games and body-play: Personhood in online chat and cybersex*. Peter Lang.
- Waskul, D. D. (2005). Ekstasis and the internet: liminality and computer-mediated communication. *New Media & Society*, 7(1), 47-63. DOI: 10.1177/1461444805049144
- William, C. (2008). <https://ew.com/article/2008/02/05/taylor-swifts-road-fame/>
- Wolff, M. (1991). *A Thrice-Told Tale: Feminism, Postmodernism, and Ethnographic Responsibility*. Stanford University Press.