

Article

Qualitative Research in the Digital Era: Obstacles and Opportunities

Ted Palys, PhD
Professor
School of Criminology
Simon Fraser University
Burnaby, British Columbia, Canada

Chris Atchison, MA
Research Associate
Department of Sociology
University of Victoria
Victoria, British Columbia, Canada

© 2012 Palys and Atchison.

Abstract

Although the many sites and opportunities available to researchers through the development and proliferation of the Internet are well known, little attention has been paid to what digital technologies and the world's developing digital infrastructure can offer qualitative researchers for the actual process of doing research. This article discusses opportunities that now exist that we have experimented with and implemented in our own research, such as viral sampling strategies, wireless interviewing, and voice recognition transcription, as well as impediments we have encountered that stand in their way. Included in the latter are research ethics boards who often lack expertise in issues that arise in computer-assisted research, hardware/software costs and technological expertise for researchers, and university administrations who have not embraced infrastructure for qualitative research to the same extent they have supported quantitative research. The article closes with a look at the implications of emerging issues, such as the trend to cloud computing, the proliferation of mobile devices, and the maturation of voice recognition software.

Keywords: Computer-assisted, qualitative, viral sampling, voice recognition, ethics, information management, CAQDA, NVivo

Author Note: An earlier version of this article was presented as a keynote address at the 10th Annual *Advances in Qualitative Methods* Conference of the International Institute for Qualitative Methodology held in Vancouver in October, 2009. The authors thank those in attendance for their comments and questions.

This article emerges from the intersection of two domains: (a) qualitative research, with all its diverse and illustrious history, and (b) the digital age, which is now most clearly upon us. While others have already considered the many research sites and opportunities that are now available to researchers through the development and proliferation of the Internet (e.g., Atchison, 1999; Bainbridge, 1999; Bauman, Airey, & Atak, 1998; Couper, Traugott, & Lamias, 2001; Palys & Atchison, 2008), little attention has been paid to what digital technologies and the world's developing digital infrastructure can offer qualitative researchers in the actual process of doing research. This article will address this issue by describing some of the ways that qualitative researchers can incorporate digital technologies into their work. Our focus is on the opportunities that the digital age offers as well as some of the impediments we have encountered that keep us from fully embracing those opportunities.

These observations bring together three elements: (a) digital technologies and the many opportunities they provide both to do new things that the qualitative research community has never done before, as well as to do better the things that it has always done; (b) the social dynamics that have been triggered, fostered, and fed by digital developments, most notably the colonization of the Internet, and the implications these have for sampling and solicitation; and (c) observations regarding some possible implications of these social and technological developments for qualitative research in the academy.

Opportunities and Efficiencies in Research

The Over-Arching Challenge: Information Management

At this point it would be hard to deny that many aspects of the research process that were formerly accomplished through more conventional approaches are becoming increasingly difficult to manage. The explosion of digital publication and e-holdings in many libraries, the growing popularity of interdisciplinary and multi-member research projects that can span institutions and even continents, and the noted impact of virtual colonization has had a huge impact on the way we manage information in the process of designing and conducting research.

Our approach has involved trying to move out of our comfort zone and to look for applications of technology that are useful at each stage of the research process. In doing this, we have found that many existing technologies are underutilized because of an overly narrow conceptualization of what they do. A classic example of this is NVivo, a well-known software program that is generally considered to be useful for the analysis of qualitative data. But its potential is far more than that. We construe NVivo as more of an information management tool that can be useful throughout the research process.

The planning stages of a project are often the time when information management is most important. The ubiquity of the portable document format (PDF) journal articles, the growth of online publications, and the dissolution of linguistic, cultural, and geographic borders that previously placed limitations on access to globally produced knowledge have made the management of information during the literature review, proposal writing, and research design process increasingly challenging. At the outset any literature one accesses can be imported into a NVivo project, and sophisticated and almost error-free optical character recognition (OCR) software programs such as Abbyy Fine Reader, OmniPage, or Readirus, can make any image-based document both searchable and codable. This is immediately useful for literature review and proposal writing, and continues through the research design, data gathering, and analysis stages—where NVivo's flexibility for many different sources of data (including audio, video, graphic, and text) are well known—of writing for publication. And one does not even need to be a qualitative

researcher or engaged in qualitative research to reap the benefits. In our own departments, for example, PhD students of all stripes are recognizing what a useful tool NVivo can be for managing the literature when they are preparing for comprehensive examinations and writing papers and proposals.

The Colonization of the Internet

We use the term “colonized” to convey that the Internet has seen the creation of places that individuals and groups have marked and settled for their own social purposes. Rather than refer merely to “settlement,” saying the Internet has been “colonized” acknowledges that settlement patterns have reflected the particular gendered, cultural, economic, and political interests of divergent hegemonic groups. Stated more simply, the Internet has fostered, more than any technology before it, the growth of social networks and the realization of an ever-growing number of “global villages” (see McLuhan, 1962, 1964) that represent a diversity of interests among members, and there is ample evidence that there is a growing will among some to control the shape and regulation of these emergent communities (Atchison, 2000; Atchison & Thomas, 2000). Although explaining the dynamics of this colonization process is beyond the scope of our article, recognizing that it is happening is a prerequisite to understanding some of the methodological implications that flow from it.

During the initial settlement of the Internet, virtual communities of like-minded people congregated using utilities such as Usenet (newsgroups), Internet Relay Chat (IRC), e-lists, and bulletin board systems (BBSs). Often global in composition, the members of those early virtual villages were typically weakly tied because of the realities of physical geography. Physically remote, single-interest bonds can be difficult to sustain, and few of us have access to the resources required to develop those relationships in person.

The first graphical web browser in 1993, and the accompanying wide-scale release of Netscape Navigator in 1994, along with the growth of commercial Internet Service Providers (ISPs), led to an explosion in the public use of the World Wide Web (WWW or the Web). This new graphical user interface (GUI) in turn led to the creation of Internet forums or message boards, which rapidly replaced Usenets, e-lists, and BBSs as the preferred medium for members of virtual villages to meet and communicate. While instant messaging technology had always existed in the form of IRC, the new GUI interface made this form of online communication even more appealing to a wider array of network technology users. The relative ease with which these forums could be set up, maintained, and used by anyone with storage space provided by their Internet Service Provider (ISP), combined with the emergence of meta-search engines such as Alta Vista, allowed growing numbers of people with access to the Web to migrate to and settle in virtual communities that were even more reflective of their particular social, political, economic, or geographic background, interests, and beliefs.

Two distinct processes were evident. First, the boundaries between “virtual” and “real” communities began to overlap as networks of individuals who were already part of “real world” friendship cliques or networks began to bond and interact with people of shared interests they met through their participation in these newly forming virtual communities. The local bridge club or Little League Baseball organization could share information not only among its local members, but also could network with similar groups elsewhere.

At the same time, some virtual communities that had been created on the basis of a more generic common interest grew to the point where the sheer size of the interest group made it both feasible and useful to organize on a more local basis. For example, while the Internet initially allowed

relatively smaller numbers of men who were the clients of prostitutes (“johns”) to connect nationally and internationally with others who shared their general interest, the growth of the Internet made it feasible for groups of johns to establish geographically-linked sites that allowed them to engage in conversation not only about broader issues involving prostitution, but also to exchange consumer-based information about local prices and services.

The demographic that appeared to be most receptive to this emergent form of communication was the younger generation of Internet users, many of whom were born only a few years before the introduction of the first GUI web browser. The more recent dawn of social networking and content sharing websites such as Facebook, Pinterest, and personal blogs, “notification” technologies such as Twitter, and the explosion in popularity of short message services (SMS), or text messaging, has resulted in yet another transformation in the ongoing process of virtual colonization. People are now able to connect, interact, and share resources with existing friends, as well as maintain ties with former friends and other contacts who are no longer a part of their current lives “offline.” Wireless network devices—such as the iPhone, Blackberry, and phones using Google’s Android OS—have made it even easier to connect with a wide variety of strongly and weakly tied members of an individual social network or community because of these devices’ portability, ease of use, and removal of barriers from information flow between people no matter the distance.

As a result, the Internet is now occupied by varying groups of people, including those in a highly dispersed physical space and others locally accessible, occasionally bridging online and offline relationships while at other times remaining within one or the other, and in some situations engaging in real time, while in others transcending it. Understanding this variation, and how different cultural and subcultural groups and networks use these virtual spaces to create and define their social identities, is a prerequisite for qualitative researchers hoping to benefit from the accessibility of these meeting spaces. Although much of this understanding has yet to be developed, we can in the interim at least start this dialogue by sharing some of our own research experiences.

Implications for Sampling

The possibility of using social network sites as sources of data is well known (e.g., Boyd & Ellison, 2008; De Bruyn & Lilien, 2004; Zimmer, 2010), but their potential utility for participant recruitment is not. In this regard, we have found the sorts of networks created by social networking sites can be extremely useful in acquiring samples, even among, and sometimes particularly among, marginalized, stigmatized, or otherwise socially isolated persons who have sought to connect with others in virtual space. Atchison, for example, has taken advantage of network communications technologies, such as text messaging, online discussion boards, and social networking sites (e.g., Facebook and Twitter), to develop recruitment strategies that build on the fact that modern network technology and communications are designed explicitly for the rapid transmission of information among and between members of distinct social networks (for a more detailed discussion see Kolar & Atchison, 2012).

In order to acquire a sample of clients of sex workers, a notoriously difficult population to locate and engage, Atchison employed a targeted-network recruitment strategy that involved approaching members of the community directly through advertisements placed within a variety of region- and topic-specific online newsgroups that sex buyers used to locate sexual service providers. The wording of the advertisements was phrased in such a way so as to highlight the importance of the research “giving voice” to this group of stigmatized individuals whose views had rarely, if ever, been sought. One particular advertisement read:

We appreciate that in an ideal world people who buy sex would be in a position to stand up and have their voices included in debates surrounding prostitution without fear of moral, social, legal or political persecution or prosecution; we recognize that in the world we currently live in this is rarely the case. We see this research as providing a chance for people who have paid for sex to be included in discussions that they have been excluded from for far too long—without having to fear that their safety, privacy, confidentiality or anonymity will be compromised in any way. If you are interested in finding out more about this important research project visit us on the web at www.johnsvoice.ca

In order to get the word out to potentially interested parties, Atchison posted a total of 450 advertisements using variations of the above wording over the duration of the project. Posting frequency was based on priority and traffic levels with the higher frequency areas of Vancouver, Toronto, and Montreal comprising 34.22% (n=154) of the posts made (an average of about 51 posts per area). The medium frequency areas of Calgary, Edmonton, and Ottawa accounted for 13.33% (n=60) of the posts (an average of about 20 posts per area). Finally, the 29 lower frequency areas (e.g., Victoria, Hamilton, Winnipeg, and Red Deer) accounted for the remaining 52.44% (n=236) of posts (an average of about 8 posts per area). This method of recruitment generated 72.6% (n=584) of questionnaire participants and produced 35 telephone or email inquiries about the research. Sixteen of these people ended up participating in more intensive follow-up interviews, accounting for 64% of interview participants.

Since at least “Doc’s” legendary introduction of William Foote Whyte (1943) to the street corner boys, network sampling techniques have, until now, relied on some form of face-to-face contact between the researcher and at least one participant or informant who agrees to vouch for or go forward on behalf of the researcher to recruit members of their social network for participation (e.g., see Lofland, Snow, Anderson, & Lofland, 2006). For his web-centered project on the clients of sex workers, Atchison developed a second network-based sampling strategy that he refers to as “viral recruitment.” Viral recruitment does not require the researcher to first identify and contact a member of the target population. Instead, this strategy requires that the researcher only know “how” to target the population and to construct a solicitation that appeals to them. Once a trusted member of the community deems the advertisement important or interesting, its transfer to further members of the community is exponential.

In Atchison’s case, the targeting of sex buyers came through identification of geographically-based networking sites catering to the interests of members of the sex-buying and -selling communities. There are currently seven such discussion boards in Canada with a collective membership of more than 220,000 people. While not all of these members are active, and some individuals are members of multiple boards (as was the case with the principal researcher for the purpose of this study), even if every individual member was a member of all seven forums, there would still be over 31,000 unique members.

Monitoring of the sites’ discussion boards over time revealed these virtual communities form around specific subcultural communities of buyers and sellers that are defined by their status (as a buyer or seller of sex), their venue experience or preference (e.g., street, escort, independent, and parlour), and their geographic location (major Canadian cities). Some members of the network have formed friendship cliques that engage in regular social interactions that sometimes extend beyond the boundaries of the virtual community to real world meetings over beer or attendance at annual community parties. Despite the strong ties that have developed among some members of the network, the majority of members appear to be peripherally related to these cliques through occasional participation in group conversations or chats, or through their interaction over the sharing of their sex-buying or -selling experiences.

In order to employ his viral recruitment strategy, Atchison informally analyzed the postings that community members had placed in the various publicly accessible discussion board threads in order to identify both conventional (e.g., newspapers, tabloids, or periodicals) and network-based (e.g., websites and newsgroups) media spaces that active online community members utilized. Once he identified these spaces, he arranged to post variations of the same advertisements that he placed in the commercial newsgroup spaces in the hopes that members would learn about the research and then pass this knowledge back to the community. The strategy proved to be quite successful, producing 21.6% of his participants for a web-based survey (n=174) and 16% (n=4) of his interview participants. Participants who came across information about the survey quickly posted messages about the research in the various discussion forums and these posts resulted in conversational “threads” where members of the community shared their thoughts about the “legitimacy” of the research and their desire to participate.

There is of course a potential downside to this more passive bait-and-wait technique, which is that while a single positive statement by a significant group member may be helpful in viral transmission of the researcher’s appeal, a single negative one can undermine the study in that forum. In this regard, monitoring of discussions at the site made clear that just as Atchison had used the Internet to his advantage in order to locate the group and engage them as the launching point for his appeal, so, too, did the participants employ their online savvy to find out more about him. Extensive searching led listserv members to the conclusion that Atchison was fair-minded and was providing a legitimate and unbiased opportunity. For example, one community member posted a message back to the community that stated,

I looked at the names of some of the people associated with the project (ain't Google wonderful!) and they have a number of legitimate research studies and analyses out there. As well, some have been involved in harm reduction strategies with the sex worker community and some have expertise in criminal justice issues. In a phrase, ‘looks legit to me.’

The intensity of prospective participants’ search should serve as a reminder to researchers that they should be careful about their online persona, and that they should expect their behaviour and reputation to come under critical scrutiny. The doors to the Internet open both ways.

Data Gathering

Intranet- and Internet-Based Techniques

Although perhaps of less interest to persons who engage solely in traditional qualitative research, those who engage in mixed-methods research also may be interested to know about some of the possibilities we have pursued in the realm of intranet- and Internet-based survey research. Although Internet-based surveys have become more common, we have yet to see anyone making creative use of the advantages that computer-based techniques bestow. Virtually every web-based survey we have seen has involved little more than taking what would have been on a white piece of paper and placing it on a screen.

In contrast, we start with a design process that ensures the look of the survey is appealing to the audience we seek. The computer also allows you to offer questions in virtually any form—we have experimented with having questions in text, audio, and video—and, depending on the hardware, to enable responses by text, audio, video, or touch, and thereby surveys can be adapted to virtually any category of population, including the illiterate and those who have any number of visual or motor impairments. In-person or purely computer-administered surveys also can be

prepared in any language, with language of administration changed at the click of a button, voice command, or touch, for use in multi-cultural and international research.

Computer-based surveys also simplify contingency questions by allowing internal links from particular responses to whichever option most appropriately follows, thereby negating any possibility for respondent or researcher/administrator error by programming contingency sequences seamlessly and invisibly into the administration process. We also have begun incorporating what might be termed “adaptive” questioning where key words from responses earlier in the survey are incorporated into the wording of relevant subsequent questions, thereby enhancing user friendliness. These can mirror the sorts of processes that effective interviewers would use (e.g., see Palys & Atchison, 2008). For example, a survey that included questions about sexual behaviour might ask early in the process what word or phrase the respondent would prefer to use when describing sexual intercourse, and from that point on any question that dealt with intercourse could incorporate the respondent’s preference into its wording.

Our examination of existing web-based survey programs and sites and every particular survey that has been brought to our attention also shows that such sites and surveys offer little or no opportunity for the more open-ended responses that would be of special interest to qualitative researchers, and that they appear to treat “data collection” and “data analysis” as two separate processes with no bridge between them. In contrast, as outlined elsewhere (see Palys & Atchison, 2008), the surveys we create typically include both structured/categorical and less structured open-ended questions and are designed using scripting that makes them easily downloadable to SPSS, NVivo, and/or Excel. This in itself creates savings and efficiencies by minimizing transcription time and the errors that would otherwise inevitably accompany that process.

Interactive Techniques

While the role that a capable interviewer or observer plays in the quality of data gathered can never be overstated, a tremendous advantage that digital technologies offer the data-gathering process is their ability to free up more research time by increasing comfort, maximizing efficiency, and minimizing error. For example, the tape recorder has always been a tool *sine qua non* for qualitative researchers, and the contemporary digital recorder is no less. But, instead of the tape recorder sitting in the middle of a table with a microphone pointing intrusively toward the participant, we have begun using two-way radio transmitter-receivers and digital recording devices where both interviewer and interviewee are equipped with a microphone and transmitter that attaches inconspicuously to a lapel or shirt collar.

Atchison had an opportunity to use this technique most recently in his in-person interviews with clients of sex workers, which allowed him to conduct interviews in true conversational style in pool halls, moderately busy lounges or cafes, or while walking about busy city streets. Not having the visual distraction of the recording device present, not having to worry about being restricted spatially, and not having any obvious hardware that would draw third-party attention as if an “Interview in Progress” sign were hanging about their necks, combined to put participants at ease, which made it much easier to establish a trust and rapport that resulted in the typical interview lasting well over 2 hours. Participants often commented at the end of the session how enjoyable the conversation was and how much they appreciated being able to talk so openly and freely.

The mobility associated with this approach also allows for interviews to be conducted in situ where research participants have visual cues they can use as they explain how particular events occurred and/or explain their behaviour or choices, which we speculate will enhance both the comprehensiveness of recall and its validity. For example, we recently advised a publishing

business in the health food industry that was undertaking in-house interview research with health food retailers regarding various aspects of the industry. Our advice included the suggestion that the interviews should be conducted in the actual store using wireless microphones while wandering about the venue. The feedback was highly positive. Retailers clearly felt more comfortable showing off and explaining “their” store, while the interviewer enjoyed having both concrete referents to anchor what was being spoken about and the ability to point to and ask about product, store features, and other issues that came to the interviewer’s or retailer’s attention.

As anyone knows who has seen a teenager speedily texting with both thumbs on their cell phone, there appears to be a preference among some members of the current generation of computer users to communicate via chat-direct or text messaging. Nevertheless, we do not recommend it in the research/interview context. The single computer-assisted on-line interview (CAOI) that was conducted in Atchison’s sex-client project was perhaps the most challenging and time-consuming of all the interviews. Unlike in person or telephone interviews, with the CAOI it was quite difficult to establish a clear sense of trust and rapport, and it was extremely difficult to obtain the same naturalness of conversation that produced such rich and thoughtful commentary in the other interview settings. In the end, while this interview did result in lengthy and well-articulated responses by the participant, and as a helpful by-product generated a comprehensive transcript of the interaction, it took nearly seven hours over the course of two evenings to complete. The equivalent amount of time spent in an in-person or telephone interview probably would have yielded a much more detailed and comprehensive account. Conversely, the same content could have been covered face-to-face in a quarter of the time. We also have concerns about confidentiality issues when sensitive information is being transmitted via text, and would feel more comfortable conducting the interview via Skype (using audio only or with video). The participant was informed of these options and their respective advantages and disadvantages for him; however, in the end he would not have consented to the interview at all if the CAOI was not an option.

Transcription

Once gathered, the next challenge awaiting the qualitative researcher is to get the material assembled into codable form. In the case of interviews, this traditionally has meant transcribing, which has often been described as the most labour-intensive part of the process, with one interview hour normally taking anywhere from 4 to 6 hours or more to transcribe. This becomes an extremely time-consuming and expensive task in any larger scale study.

In our own work we have begun experimenting with automated transcription processes using voice recognition software. The program we use is called Dragon Naturally Speaking (DNS), which is generally recognized as one of the top voice recognition programs available. Other popular commercial and open source voice recognition software applications include IBM ViaVoice, MacSpeech, Microsoft Windows Speech Recognition, Philips SpeechMagic, Sphinx, and VoxForge. For those unfamiliar with the program, DNS, once installed, requires a brief training process in which the user reads standard passages for approximately 15 minutes. The brief training process trains both the researcher and the program—the researcher to speak in a manner that is most recognizable to the program and the program to recognize the researcher’s unique speech—and is remarkably accurate once trained. This can be supplemented with other user-generated material such as documents, which allows the researcher to introduce vocabulary that is unique to his/her interests and research. For example, Palys does face-to-face and legal research with Indigenous peoples whose traditional tribal, place, and personal names are far beyond the program’s default capability, but can be produced unerringly by DNS with a few minutes of training. The same would be true of any technical vocabulary.

Our first experiment using DNS for transcription involved simply inputting a taped interview to the program for transcription. This yielded a highly error-filled copy that required almost as much time to edit as traditional transcription would have taken. Our next step was to exploit the fact that the program knew each of our voices by making us the medium through which the interview was taken from digitized audio file to transcript. This involved wearing headphones, slowing the tape down by 20-30% (using another program such as the freely available Express Scribe, the commercially available Adobe Soundbooth, or the playback provisions in NVivo), and then speaking out what we heard into DNS as the recording played. Because this process used our own voices, which had already been trained in the program, the error rate was trivial and easily edited with only a few interviews for practice. We found the total time taken dropped from 4-6 hours of transcription time for 1 hour of interview to about 1.5 hours of transcription time, which was a substantial time saving.

Another option would be to actually begin or end an interview by putting the respondent through the training process, the basic version of which only takes about 15 minutes. While such sessions would not be particularly useful or feasible for single-interview studies, even when each interview lasts for more than an hour, they may be highly useful for something like an oral history project where a given participant is sometimes interviewed for dozens of hours. Any digital recordings then could be input into DNS and auto-transcribed. Because some training of the program already will have occurred, accuracy will be high from the start. By correcting errors from the first interview, and thereby further training the program, subsequent accuracy will be even higher.

Yet another more promising option with which we have only begun to experiment is the batch conversion of digital recordings into a single voice file for the purposes of automated transcription. In this process, sophisticated audio software such as Adobe Soundbooth is used to apply a series of specially designed sound filters to digitally recorded files. During the first stage of this process, a filter is applied to remove all ambient noise from the recording that may have been present during the initial interview. In the second stage, a special filter is then applied in order to change the pitch and tone of all voices present in the recording into a single voice. Once the voice is converted, a third filter is applied to the new voice to slow down the rate of speech to 20-30% of the original rate so that the speech is clean and methodical. The final version of the file can then be auto-transcribed using the DNS batch transcription feature.

Data Analysis

This is another area where much has been written and where software such as NVivo and ATLAS.ti clearly market themselves (Auld et al., 2007; Crowley, Harre, & Tagg, 2002; Durian, 2002; Lewis, 2004; Smyth, 2006). While we will not comment on what programmes like NVivo and ATLAS.ti can do, we would like to observe that these programmes are in their relative infancy and companies like QSR (the parent of NVivo) are actively working with members of the qualitative and mixed methods research community (academic and commercial alike) to build features into the next generation of these programmes that reflect the needs and applications of those who could benefit from the use of this software.

Methodological Caveats

We have focussed on opportunities first, and hope they have opened readers' eyes to some possibilities that digital technologies can generate in their research. But there are other considerations that arise from exploiting the technologies and techniques we have outlined that represent a potential downside or disadvantage for incorporating these digital interventions. These

disadvantages fall into three main categories: (a) start-up issues, (b) ethical issues, and (c) academic and institutional support.

Start-up

There are two key aspects to start-up. The first is the hardware and software required to engage in the various practices we have described. The second is the expertise required to run that hardware and software.

In the grander scheme of things, hardware and software costs are not unreasonable, but neither are they inexpensive. We each own several desktop computers and laptops and prefer to work using multiple monitors, and then there are the microphones, radio transmitters, and web and data servers. Software is probably the most expensive element of the equation; we have mentioned several in passing—NVivo, Adobe Soundbooth, Dragon Naturally Speaking, Abbyy Fine Reader—which can total several thousand dollars, even with educational pricing. While the solo, largely self-funded researcher is likely to take some time to accumulate these tools, those who engage in larger projects should be able to write in such purchases fairly easily; they would be a relatively small line item on any half-million dollar budget. Indeed, we would hope that an article such as this one would help make it easier to justify these purchases by showing how much they can facilitate and gain efficiencies throughout the research process. It also should be noted that there are a wealth of open-source shareware and freeware software solutions available to those with the desire to implement technology into the research process, but who do not have the financial resources to invest immediately.

The bigger issue is likely whether the interested reader has the expertise, or has access to the expertise, to actually run all these programs and servers and equipment. We might be seen as an interesting case study in that regard, and probably represent the two major groups who might consider our suggestions. Atchison is the clear alpha technophile and the one who is typically the first to try out any new technology. Palys is also a technophile insofar as interest and educability go, but does not possess Atchison's level of expertise in being able to do the scripting and programming required to make them run. Not everything requires servers and scripting, and Palys has the main prerequisite of being willing and educable on the use of new software. Neither of us has a degree in computing science; both are self-taught.

Realistically, we presume and hope that some more adventurous and tech-savvy and/or financially independent readers of this document will start to experiment with the software and hardware they have, push their boundaries a bit further, and write follow-ups to this article that help create a burgeoning inventory of tricks and tips that will bring computer-assisted qualitative research into the 21st century. Although the situation at the moment would seem to favour those who are more advanced technologically—writing scripts and setting up and operating servers are not within most researchers' range of expertise, creating the demand for future development is crucial, and if prior development patterns are at all predictive, technologies once considered esoteric can become user friendly and mainstream very quickly.

Palys cut his academic teeth in the days of mainframes, and he is one of those who took his dissertation data to the computing centre in boxes full of 80-character IBM punch cards. These first versions of what was then an esoteric program called SPSS were themselves leagues ahead of the rooms full of mechanical calculators that he had been trained on as an undergraduate, and were at that time considered the province of the adventurous. Similarly, web pages once required designers to be proficient in scripting html. WYSIWYS (what you see is what you get) html editing programs such as CoffeeCup, FirstPage, Front Page, GoLive, HotDog, and Dreamweaver

soon came along to “do” html invisibly with a point-and-click interface that made every user a possible web designer. Qualitative researchers are now at that similar ground level of digital development and are thus in a position to promote and influence its development.

Ethical Issues

Because we have not addressed the various new venues for research that have been made possible by the Internet, we will not discuss the many ethical issues that arise, and have already been discussed, such as where the dividing line is between “public” and “private” on the Internet and what “netiquette” is involved when approaching and interacting with various social networking groups (but see Atchison, 1999; Palys & Atchison, 2008). Within the range of topics covered, the primary ethical issue we see is confidentiality and the impact that digital technologies have on the security of data. In this regard, two primary areas of concern can be identified.

The first concerns the security of digital data that is gathered interactively via the Internet, for example, through web surveys or interviews that in one way or another are computer-assisted (e.g., by using Skype or any of the various messaging or networking programs). Once the data are in our possession, we try to make them as invulnerable as possible by running our own data servers, anonymising as soon as possible, encrypting everything sensitive and/or taking it offline until needed, and never leaving any content information of any sort on our tablet laptops. Other researchers may need to take even greater care, depending on the sensitivity of the information they have and the potential consequences to their participants, and whatever unique legal constraints or enhancements exist for protecting research confidentiality. Researchers in the United States, for example, need to understand how the *USA PATRIOT Act* (2001) has undermined researchers’ ability to protect confidential, sensitive information unless they are in possession of a Confidentiality Certificate (through the National Institutes of Health) or Privacy Certificate (through the National Institute of Justice). The Simon Fraser University research ethics board now prohibits researchers from gathering sensitive data through any United States company because of the threat to confidentiality the *USA PATRIOT Act* represents, and it is not the only research ethics board in Canada to have done so. In Canada, no statute-based protections of any sort are available to researchers at this time other than those at Statistics Canada (via the *Statistics Act*), which means that those who do research involving sensitive information from identifiable sources that would be of interest to a powerful third party must be doubly careful, although such challenges occur much less frequently than in the United States (Lowman & Palys, 2001).

A second area that causes us concern, which is related to most researchers not having their own web and data servers, concerns the new tendency toward cloud computing, in which users store data on someone else’s server. Examples of such venues include picture-uploading sites, such as Picassa, as well as blogs, Facebook, G-mail, and so on. Owners of these types of sites emphasize the convenience and minimal storage requirements that this offers users who can now access their information from any mobile device anytime, anywhere. And while such sites are potential gold mines for their research potential, for researchers concerned about their participants’ confidentiality, this trend is fraught with danger because of the loss of control it involves; essentially it places the researcher and his or her data at the whim of another party’s privacy and use policies, very few of which are created with the rights and interests of researchers and research participants in mind. Many of these sites also are in the United States, which brings us back to the *USA PATRIOT Act*. Researchers doing computer-assisted research involving sensitive information outside North America should take the time to understand the legal context in which computing operates, and especially the extent to which the nation state in which their participants

reside is involved in monitoring computer traffic, as these practices and their potential consequences to participants vary widely.

We are especially distrustful of university-based web and data servers, which are often terribly vulnerable to hacking, and typically subject to use policies that give the university the right to inspect any information and to manage the information in the event that a challenge to research confidentiality arises. Although one would hope that, if such a situation arose, one's university would zealously engage the legal battle required to safeguard research confidences, and both the research participants and academic freedom their existence protects, thus far, the record of Universities doing so is less than perfect (e.g., see Lowman & Palys, 2000, 2001; Palys & Lowman, in press; Rusk, 2012).

Beyond that, however, we find more in the way of what we consider misunderstanding about the Internet. For example, Palys sat on a federal committee that was charged with advising the Presidents of Canada's major granting agencies on prospective future development of their federal ethics policy—the *Tri-Council Policy Statement* on ethics in research involving human participants—where one issue that arose was whether any specific ethical guidelines were necessary for Internet-based research. Palys was among those who balked at the idea, in part because he believed (a) any technological specifics the group might suggest would undoubtedly be obsolete by the time they were published and (b) that there was no more need for special policies on Internet-based research than there was a need for unique policies for telephone-based research, or kitchen-based research. In every case there are threats that theoretically can be envisioned: the telephone might be wiretapped, the kitchen might have a bug in it, but the likelihood of this threat being realized in most cases ranges from remote to non-existent.

The Academy

Perhaps most surprising for us in terms of obstacles to pursuing the digital vision we have outlined in our own research and with our students are the impediments that have arisen within the academy. Three of these are noteworthy.

The first would have to be Research Ethics Boards (or REBs as they are known in Canada) and Institutional Review Boards (or IRBs as they are known in the United States), who rarely have a member with significant computing expertise, and, it seems, equally rarely have anyone with qualitative research expertise (Social Sciences and Humanities Research Ethics Working Committee, 2004). Far too often we have heard the mantra that “nothing is private” on the Internet, without any serious consideration of what it would take for someone to intercept or steal a piece of identifiable information about someone in the context of an actual research project. The challenge is to avoid running away and suppressing sensitive research, while recognizing when a realistic threat exists and engaging a specific ethical and/or legal strategy to address it (e.g., Palys & Lowman, 2010).

A second obstacle comes from university administrations that, at least in our experience, seem willing to devote significant portions of their budgets to quantitative software, but are unwilling to do the same for qualitative software. At Simon Fraser University, for example, every single computer in every single lab on campus—hundreds and hundreds of computers—has SPSS on it, and faculty at the university are able to download the program for free as part of the university's lease arrangement. Meanwhile, there is not, to our knowledge, even one computer on campus that has a university-funded qualitative data analysis program like NVivo on it, even though we have located an extensive group of individuals across at least eight different disciplines who would be interested in seeing the university invest in leases, so that learning qualitative data analysis

software could be included in our undergraduate and graduate curriculum the same way that SPSS is. Tired of waiting for Computing Services, Palys's department decided to invest in six leases for its own graduate computer lab so that he can teach the program in his graduate qualitative methods class. Another dozen people in the department decided to invest in personal leases or outright purchases, but none of those funds came from the university administration or computing services.

And finally, we would be remiss if we did not also mention the resistance we have found among some members of the qualitative research community who use their computers for email and word processing but view the prospect of digitalizing data and engaging it with qualitative software as somehow sacrilegious. Although we understand that some persons are more technologically interested than others, in our view the kinds of tools and techniques we have discussed are simply part of an ever-diversifying qualitative research landscape, and an inevitable part of its future. The current generation of graduate students needs to know enough about what digital possibilities are available to them to make their own choices about the extent to which they want to pursue the options the digital world opens to them, and help develop the next generation of digital tools so that they are maximally beneficial for both researchers and participants alike.

Conclusion

We have attempted in this article to outline some of the benefits we have realized in our research by crossing the digital divide and asking what exists and can be constructed to facilitate qualitative and mixed methods field research. Sharing these efforts with our colleagues, both to assist their research and to encourage them to begin incorporating some of the possibilities we have described into their research, is the easy and enjoyable part. Trying to envision the possibilities and stay ahead of the curve in the realm of technology-assisted social and health research, where it seems every week brings a new development, is equally enjoyable, but also a considerable challenge. There are, however, two rising issues that will have considerable impact for researchers.

The first is the growing trend toward cloud computing. Although we understand its appeal for recreational computer users, we see it as a regressive trend that is good for those who run the data servers but bad for researchers and their participants because of the relinquishing of control it involves. Researchers need to be responsible for the care of their data and cannot do so when they are at the whim of their cloud service provider's use policies. To the extent this becomes inevitable, researchers should choose their web and data servers well, and ideally develop the capacity to run their own.

The second, which we have only begun to consider, involves the possibilities that are opening up because of the virtual revolution of portable computing power brought about by the iPhone, Blackberry, Google, and their respective apps. Because we are reaching the point where virtually everyone has a cell phone, wireless connectedness allows us to envision the possibility where everyone in our audience can download our survey or interview questions and respond right away by text, voice, or video. Voice recognition has the potential to make keyboards obsolete and open up data gathering possibilities even further. Beyond that, time will tell.

References

- Atchison, C. (1999). Navigating the virtual minefield: Using the Internet as a medium for conducting primary social research. In D. Currie, D. Hay, & B. MacLean (Eds.), *Exploring the social world: Social research in action* (pp. 145-158). Vancouver, BC: Collective Press.
- Atchison, C. (2000). Emerging styles of social control on the Internet: Justice denied. *Critical Criminology: An International Journal*, 9(1-2), 85-100.
- Atchison, C., Lowman, J., & Fraser, L. (1998). Men who buy sex: Preliminary findings of an exploratory study. In J. E. Elias, V. L. Bullough, V. Elias, & G. Brewer (Eds.), *Prostitution: On whores, hustlers, and johns* (pp. 172-203). Amherst, NY: Prometheus Books.
- Atchison, C., & Thomas, J. (2000). The internet and the state: Instrument of control or subversive technology? *Critical Criminology: An International Journal*, 9(1-2), 163-170.
- Auld, G., Diker, A., Bock, M., Boushey, C., Bruhn, C., Cluskey, C., . . . Zaghloul, S. (2007). Development of a decision tree to determine appropriateness of NVivo in analyzing qualitative data sets. *Journal of Nutrition Education and Behavior*, 39(1), 37-47.
- Bainbridge, W. (1999). Cyberspace: Sociology's natural domain. *Contemporary Sociology*, 28(6), 664-668.
- Bauman, S., Airey, J., & Atak, H. (1998, May). *Effective use of web-based technology: Using the Internet for data collection and communication applications*. Paper presented at the 53rd Conference of the American Association for Public Opinion Research, St. Louis, MO.
- Boyd, D., & Ellison, N. (2008). Social network sites: Definition, history, and scholarship. *Journal of Computer-Mediated Communication*, 13(1), 210-230.
- Couper, M., Traugott, M., & Lamias, M. (2001). Web survey design and administration. *Public Opinion Quarterly*, 65(2), 230-55.
- Crowley, C., Harre, R., & Tagg, C. (2002). Qualitative research and computing: Methodological issues and practices in using QSR NVivo and NUD*IST. *International Journal of Social Research Methodology*, 5(3), 193- 197.
- De Bruyn, A., & Lilien, G. (2004). A multi-stage model of word of mouth through electronic referrals. eBusiness Research Center Working Paper. Retrieved from http://www.smeal.psu.edu/cdt/ebrcpubs/res_papers/2004_02.pdf
- Durian, D. (2002). Corpus-based text analysis from a qualitative perspective: A closer look at NVivo. *Style*, 36(4), 738-742.
- Kolar, K., & Atchison, C. (2012). Recruitment of sex buyers: A comparison of the efficacy of conventional and computer network-based approaches. *Social Science Computing Review*. Advance online publication. doi:10.1177/0894439312453001.

- Lewis, R. B. (2004). NVivo 2.0 and ATLAS.ti 5.0: A comparative review of two popular qualitative data-analysis programs. *Field Methods*, 16(4), 439-469.
- Lofland, J., Snow, D., Anderson, L., & Lofland, L. H. (2006). *Analyzing social settings: A guide to qualitative observation and analysis* (4th ed.). Belmont, CA: Wadsworth.
- Lowman, J., & Palys, T. S. (2000). Ethics and institutional conflict of interest: The research confidentiality controversy at Simon Fraser University. *Sociological Practice: A Journal of Clinical and Applied Sociology*, 2(4), 245-55.
- Lowman, J., & Palys, T. S. (2001). The ethics and law of confidentiality in criminal justice research: A comparison of Canada and the United States. *International Criminal Justice Review*, 11(1), 1-33.
- McLuhan, M. (1962). *The Gutenberg galaxy: The making of typographic man*. Toronto, ON: University of Toronto Press.
- McLuhan, M. (1964). *Understanding media: The extensions of man*. New York, NY: McGraw Hill.
- Palys, T. S., & Atchison, C. (2008). *Research decisions: Quantitative and qualitative perspectives* (4th ed.). Toronto, ON: Thompson-Nelson Canada.
- Palys, T., & Lowman, J. (2010). Going boldly where no one has gone before? How confidentiality risk aversion is killing research on sensitive topics. *Journal of Academic Ethics*, 8(4), 265-284.
- Palys, T. S., & Lowman, J. (in press). Defending research confidentiality "to the extent the law allows": Lessons from the Boston College subpoenas. *Journal of Academic Ethics*.
- Rusk, D. (2012, January 2). Researcher hits out at Boston College in IRA interviews row. *Irish News*. Retrieved from <http://bostoncollegesubpoena.wordpress.com/2012/01/02/researcher-hits-out-at-boston-college-in-ira-interviews-row/>
- Smyth, R. (2006). Exploring congruence between Habermasian philosophy, mixed-method research, and managing data using NVivo. *International Journal of Qualitative Methods*, 5(2), 1-11.
- Social Sciences and Humanities Research Ethics Special Working Committee. (2004). *Giving voice to the spectrum: Report of the Social Sciences and Humanities Research Ethics Special Working Committee*. Retrieved from <http://pre.ethics.gc.ca/eng/archives/policy-politique/reports-rapports/reports-rapports/>
- Tapscott, D. (2008). *Grown up digital: How the net generation is changing your world*. New York, NY: McGraw-Hill.
- Uniting and Strengthening America by Providing Appropriate Tools Required to Intercept and Obstruct Terrorism (USA PATRIOT Act) Act of 2001, Pub. L. No. 107 –56. 115 Stat. 272 (2001). Retrieved from <http://www.gpo.gov/fdsys/pkg/PLAW-107publ56/pdf/PLAW-107publ56.pdf>

Whyte, W. F. (1943). *Street corner society: The social structure of an Italian slum*. Chicago, IL: University of Chicago Press.

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