Qualitative Research at the Gates of the Digital Age: Obstacles and Opportunities

Ted Palys & Chris Atchison
Simon Fraser University

Keynote address for the 10th Annual Advances in Qualitative Methods Conference of the International Institute for Qualitative Methodology held in Vancouver, October 2009.

Introduction
Our paper emerges from the intersection of two domains: (1) Qualitative Research with all its diverse and illustrious history; and (2) the Digital Age, which is now most clearly upon us. Our intention is not to offer an extended or systematic analysis of either, but to draw together some observations of both that we have made and that have influenced the way we go about our work. We would be out there doing that work right now if it were not for one other observation that made us think it might be worthwhile to come and talk to you today: the observation that many of our colleagues, who presumably also see the world changing and also have come to the obvious conclusion that the Digital Revolution has arrived, have ignored what the digital world has to offer them and simply continued with business as usual. It is for that reason that the words “obstacles” and “opportunities” also appear in our title – because we intend to talk not only about some of the opportunities that the Digital Age offers us, but also to make note of some of the impediments that exist, both within the academy generally as well as among the community of qualitative researchers, that keep us from fully embracing those opportunities.

Our observations bring together three elements: (1) the social dynamics that have been triggered, fostered and fed by digital developments, most notably the colonization of the internet; (2) the digital technologies themselves 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 the things that we have always done better; and (3) some random observations on our part – random not in the statistical sense, but in the sense of numerous and sometimes disjoint, reflecting no more than our own interests and experience – regarding some possible implications of these social and technological developments for qualitative research in the academy. That’s a lot of turf, and we have less than an hour to till it, so let’s begin.

Colonizing the Internet
Let me [Ted] begin by saying that it is no accident that there are two of us up here talking with you today. Although I was the one originally invited to come and talk with you, it would have been dishonest and downright exploitative of me to come today and talk on this topic without sharing the stage with
Chris, who began some years ago in the role of student and from whom I have learned much. That in itself is one of the social realities of the digital age. Once upon a time, qualitative research was something that involved the mentoring of younger students by older, experienced, street-savvy researchers who would introduce their academy progeny to the contacts they had cultivated and the time-worn and time-tested tools of the trade – detailed field notes, the analogue tape-recorder, the hours required to create accurate transcripts from the recordings, the highlighter, the yellow sticky, and so on. Instead, as Don Tapscott (2008) observes in Grown Up Digital, the world is now comprised of folks like me, for whom computers and handholds and iPhones are “technology,” and a younger generation like my kids who have never known a world in which the personal computer and the Internet did not exist, and who would be as likely to call their iPhones “technology” as they would their refrigerator. For them, both just are.

An implication of this state of affairs is that we senior folk no longer have the corner on expertise. Although one hopes we still have something to offer on the basis of our experience, the fact of the matter is that it is our students (as a group) who are more likely than us to be the experts on digital technologies so that we have to be equally willing to learn from our students, and to share the credit with them.

A corollary of this, however, is that if we want to understand how digital technologies, and especially the internet, are being used, then the best people to watch are those youth whose lives revolve around it and who have known nothing else. In that regard, we understand the Internet as a place that has been colonized by at least two major groups of people: (1) those – especially, but not exclusively, youth -- who exploit it as a social medium; and (2) some entrepreneurs, who see it as a place to make money.\(^1\)

Although an examination of the overlapping and sometimes competing aspirations of those two groups are worthy of attention on their own, our interests are more focussed on the former than the latter because it is the social implications of the internet that provide qualitative researchers with new venues and new ways to do their research. We use the term “colonized” not to conjure up its more pejorative connotations of subjugation and confinement of people and spaces – although there are many who are trying to do exactly that -- but simply to convey that the internet has involved the creation of places that individuals and groups have marked and settled for their own social purposes. 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 “global village” or, more to the point, “global villages” (see McLuhan, 1962, 1964). Although explaining the dynamics of this colonization process is beyond the scope of our paper, accepting that it has happened is a prerequisite to our main objective, which is to understand some of the methodological implications that flow from it.

\(^1\) We say “some” entrepreneurs because while many do see the internet as a place offering new opportunities to make money, many others – often with vested interests in the “old economy” – see it as a threat that will displace their privileged position as a distributor of what in the past were conceived as hard items. Examples here include the music and film industries, who are among the most zealous in trying to defend a proprietary internet, in contrast to youth – ironically, their biggest market – who are seen as the enemy because of the letter’s proclivity to use the internet to network and share.
The Growth of Social Networks in the Global Village
An exciting feature of the internet – made possible in large part by its foundational tools, the web browser and the search engine -- is the way it has provided a medium for like-minded and previously often disparate and marginalized people to find each other and create virtual communities.

During the initial colonization of the Internet, virtual communities of like-minded people congregated using utilities such as Internet Relay Chat (IRC), bulletin board systems (BBS), and newsgroups. Often global in composition, the members of those virtual communities were typically weakly tied due to the realities of physical geography; while it is all well and good to connect with people who share your interests in other parts of Canada, England, India and Australia – single-interest bonds are not the most sustaining, and few of us have access to the resources required to develop those relationships in person.²

The advent of instant messaging brought about a new age that saw people using it as a way to connect with existing friends online. With the dawn of social networking websites such as Facebook and MySpace, personal blogs, and “notification” technologies such as Twitter, people are now able to connect and interact with existing friends as well as former friends and strangers who are not 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 their portability, ease of use and by removing barriers from information flow between people no matter the distance.

As a result, the internet is now colonized by varying groups of people, sometimes across physical space and sometimes within one, sometimes across online and offline relationships, and sometimes within real time and sometimes transcending it. Understanding this variation and how different cultural and sub-cultural 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 we do not purport to have a blanket expertise in such issues, I [Chris] will share my experiences in using my knowledge about one particular niche to benefit my research.

Bourgeoning Opportunities and Efficiencies in Research
We have only begun our explorations of how digital technologies can be used to enhance the qualitative research process, but are happy to share some of the opportunities and efficiencies we have realized thus far.

Research Venues
We have little to say about this today, in part because so much has been written about it already by us and others for more than a decade (e.g., see Atchison, 1999; Palys & Atchison, 2008). Certainly the digital world has created many more opportunities for data gathering because of the many new ways –

² However, it is interesting to note that it is during this period we started hearing about internet-generated romances where people who met online would develop relationships at a distance – presumably fostered by the immediacy of worldwide communication the internet made possible-- that would then be consummated in person.
many of which are highly public – that people have to engage each other directly (through real time interaction with their “real” or avatar or other proxy selves) or indirectly (e.g., by posting material for interested persons to find, consume and react to). However, you have probably heard much about this; our preference is to attend to issues you have perhaps heard less about and to consider other ways of making use of that knowledge.

The Challenge of Information Management
Our approach has involved looking for applications of technology that are useful throughout the research process and think many existing technologies are under-utilized because of an overly narrow conceptualization of what they do. A classic example of this is NVivo, a very well known software programme that is generally considered to be useful for data analysis. We construe NVivo as more of an information management tool that can be useful throughout the research process as you will see.

At the outset, particularly with the bourgeoning appearance of online journals, online access to journals, more frequent access to books, any literature one accesses can be saved as part of an NVivo project and, with sophisticated and almost error-free optical character recognition (OCR) software programs such as Abbyy Fine Reader, can make any document or text both searchable and codable. This is immediately useful for literature review and proposal writing and continues through data gathering and analysis – where NVivo’s flexibility for many different sources of data (including audio, video, graphic and text) are well known – to writing for publication. We encourage our students to utilize NVivo from the very start of their research the moment a project is anticipated. Indeed, one thing we have found is that the longer one leaves the identification of a project as a project, the less likely a student (or us) is likely to use the software because of the magnitude of catch-up to photocopy and/or scan (where applicable). The more that books and journals become available online, the less onerous this task will be.

Sampling
The possibility of using social network sites as sources of data is well known, 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 even especially among, marginalized, stigmatized or otherwise socially isolated persons who have sought to connect with others in virtual space. Chris, for example, has taken advantage of social networks reified using network communications technologies such as text messaging, online discussion boards and social networking sites (e.g., Facebook, MySpace, and Twitter) to develop what he has called a “viral recruitment strategy” that builds 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.

Network sampling techniques have, until now, relied on some form of direct 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. Viral recruitment doesn’t 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 the population. Once the advertisement is deemed important or interesting
by a trusted member of the community, its transfer to further members of the community is exponential.

In Chris’s case, the sample sought were the clients of sex workers, who have been notoriously difficult to locate and engage. The targeting 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 during this study), even if each 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 sub-cultural 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 strongly tied 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 between 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. Chris’s appeal came with a promise to “give voice” to this group of stigmatized individuals whose views are rarely if ever sought. In the end, this recruitment strategy produced almost a quarter of his participants for a web-based survey (21.6%; n=174) and 16% (n=4) of his interview participants, which was greater than any other single recruitment strategy. (?)

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 might mean the demise of the study in that forum. In Chris’s case it was clear from communications within the group – which Chris was able to monitor by lurking within the group – that he had been checked out via Google and come up as someone who would be fair and had no axe to grind.

**Managing Projects Large and Small**

We’ve already mentioned how a fundamental challenge to qualitative research for which digital technologies are ideally suited is the challenge of information management from literature to proposals to data to reports and publication. NVivo has shown itself to be well-suited for this task, particularly when its capabilities are supplemented with OCR software that allows those capabilities to be more fully exploited. This has proven enough for Ted, who is more the academic dinosaur – someone who prefers to be the lone researcher gathering archival or interpersonal data over time. For example, a current project looking at “qualitative research and the methods of everyday life” has me bringing together material involving criminal investigation and wrongful conviction, medical diagnostics and mis-diagnosis, and commissions of inquiry good and bad. The range of material this includes – text, web sites, videos,
songs, film and TV excerpts, photos – is considerable, but all manageable with NVivo8, whose capabilities now include such material.

Chris, however, exemplifies much more the emerging academic – someone who engages in his own research a good proportion of the time but who also is equally likely to become engaged in some large-scale study involving multi-disciplinary and/or multi-site teams. Keeping everyone onside in such an environment involves a whole new set of challenges, which Chris has dealt with by incorporating network technologies to the task.

An example is the Urban Aboriginal Peoples Study for which Chris was the Vancouver coordinator – one of seven sites across Canada. The challenge was … The solution was to … use message boards as an organizational vehicle.

Data gathering
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. The tape recorder has always been a tool sine qua non for qualitative researchers, and the contemporary digital recorder is no less. However, instead of the tape recorder sitting in the middle of a table with a microphone pointing obtrusively to 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. Chris utilized 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 and not having to worry about being restricted spatially 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.

Telephone interviews (n=15) were also done and they, too, were digitally recorded and conducted in the same conversational style as the in-person interviews. However, unlike the in-person interviews, the relative insecurity of the telephone setting made many participants sceptical of the research and weary of the legitimacy of the researcher conducting the interview. Concerns about legitimacy, privacy and confidentiality among participants meant that far more time had to be spent establishing trust and rapport before the interview could commence. Quite often this involved elaborating on who was funding and sponsoring the project as well as providing clarification about the precise objectives of the study and how the resulting data were going to be used. Additionally, several participants wanted to be assured that the interviewer was not a member of law enforcement and that they would not be identified in any way - despite the fact that many participants contacted the research team using a blocked telephone line and were explicitly instructed to refrain from saying anything that could be used to identify them during the course of the interview. Thankfully, once participants’ concerns were put to
rest they appeared to be quite willing to engage in unrestricted conversations that again often lasted longer than 2 hours.

There appears to be a preference among some members of the current generation of computer users to communicate via chat-direct or text messaging as they seem to find this mode of communication more natural and convenient. Despite this fact, the single computer-assisted on-line interview (CAOI) that was conducted for the sex-client project was perhaps the most challenging and time-consuming of all the interviews. Unlike with the 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, it took nearly 7 hours over the course of two evenings to complete. The equivalent amount of time spent in an in-person or telephone interview would have probably yielded a much more detailed and comprehensive account. Having said this, the participant 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 gathered 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 to transcribe. This becomes a monumental 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, which is generally recognized as one of the top voice recognition programs available. For those of you unfamiliar with the program, DNS once installed requires a brief training process in which you read standard passages, which can be supplemented with other user-generated material. This trains both you and the program – you to speak in a manner that is most recognizable to the program, and the program to recognize your unique speech – and is remarkably accurate once trained.

Our first experiment at 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 a traditional transcription would take. 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 (using another program such as Adobe Soundbooth) down by 20-30%, and then speaking out what we heard into DNS. Because this process used our own voices, which had already been trained in the program, the error rate was trivial and easily edited, while 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, i.e., a substantial time saving.

Another option which we envision in theory but have not yet tried in practice would be to actually begin or end an interview by putting the respondent through the training process, a 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, but may be highly useful for something like oral history research where a given participant is sometimes interviewed for dozens of hours.

Data Analysis
This is another area where much has been written and where software such as NVivo and Atlas clearly market themselves. Accordingly, we will not repeat that material here.

Methodological Implications:
We have depicted the opportunities first, and hope that you have gorged on the possibilities that digital technologies open in your lives. 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 the digital interventions we’ve described. These fall into three main categories: (1) startup issues; (2) ethical issues; and (3) academic and institutional support.

Startup
There are two key aspects to startup. The first is the hardware and software required to engage in the various practices we’ve described here today. 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 nor are they cheap. We each own several desktop computers and laptops, and then there are the microphones, radio transmitters, servers. Software is probably the most expensive element of the equation; we have mentioned several in passing here today – NVivo, Adobe Soundbooth, Dragon Naturally Speaking, Abbyy Fine Reader – which can total several thousand dollars. While the sole, 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.

The bigger issue is likely whether you have the expertise, or have access to the expertise, to actually run all these programs and servers and equipment. The two of us are actually an interesting case study in that regard, and probably represent the two major groupings who will take up our offer. Chris is the clear alpha technophile here and the one who is typically the first to try out any new technology. Ted is also a technophile insofar as interest and educability go, and is at the point of being able to build computers from scratch, but not yet at the point of building servers, and is nowhere near Chris’s level of expertise in being able to do the scripting and programming required to make them run. However, not everything requires servers and scripting, and I’m as willing and educable as the next person in terms of trying and using new software. And while I’m the first to laud and learn from Chris’s skills, it should also be noted that neither of us has a degree in computing science; both are self-taught.

Ethical Issues
Because we have not addressed the various new venues for research that have been made possible by the internet, we won’t go here into the many ethical issues that arise, and have already been discussed,
about such issues 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, for example, Atchison, 1999; Palys & Atchison, 2008).

Within the range of topics we have covered today, the primary ethical issue we see is confidentiality and the impact that digital technologies have on the security of data. In this regard, we see two primary areas of concern.

The first concerns the security of digital data that is gathered interactively over the internet – 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). In that regard, we would make a distinction between how we do things versus what is available to other researchers who do not have our infrastructure. Once the data are in our possession, we consider them bullet proof because we have our own servers, anonymize as soon as possible, encrypt everything sensitive and/or take it offline until needed, and never leave any content information of any sort on our tablet laptops.3 Other researchers may need to exhibit far greater care or simply understand the limits of what they have. Researchers in the United States, for example, have to understand the impact of the USA Patriots Act that has undermined the statute-based protections that were otherwise available through vehicles such as Confidentiality Certificates (through NIH) and Privacy Certificates (through NIJ). Our research ethics board will not let researchers gather sensitive data through any US company because of the threat to confidentiality the USA Patriots Act represents.

The second area that causes us concern, which is related to the same problem of most researchers not having their own servers, concerns the advent of cloud computing, in which users store data on someone else’s servers. Examples of such venues include picture uploading sites such as Picassa as well as blogs, Facebook, G-mail and so on. We do not support this trend, and think it is particularly fraught with danger for researchers because of the loss of control it involves, essentially placing you 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, and many of which are in the US, which brings us back to the USA Patriots Act problem.

Beyond that, however, we find more in the way of what we consider misunderstanding about the internet. For example, Ted sat on a federal committee that was charged with advising the granting agencies on prospective future development of their federal ethics policy – the Tri-Council Policy Statement on ethics in research involving human participants – where an issue that came up repeatedly was whether any specific ethical guidelines were necessary for internet-based research. Ted was among those who consistently balked at the idea, in part because he believed (a) any technological specifics they might suggest would undoubtedly be obsolete by the time they were published; and (b) that there

---

3 Stories of people leaving laptops with sensitive information on them in their cars and being stolen are far too numerous for our liking. We use external drives for identifiable and/or sensitive content whenever possible, and remove any sensitive information as soon as possible thereafter if not, and keep these memory drives separate from our computers.
was no more need for special policies on internet-based research any more than we would need unique policies for telephone-based research or kitchen-based research. In every case there are threats that can be envisioned – the telephone might be wiretapped; the kitchen might have a bug in it – but the odds of this threat being realized on some *a priori* basis is remote.

**The Academy**

Perhaps most surprising for us in terms of obstacles to being able to pursue these digital dreams we have outlined in our own research and with our students are the impediments to doing so that have arisen within the academy. Three of these are noteworthy.

The first would have to be research ethics boards, who rarely have anyone with significant computing expertise on them, and even more rarely have anyone with qualitative research expertise. We have found all too many of them threatened by the more open-ended and inductive/emergent nature of much qualitative research – which leads some of the more liability-oriented to run through endless scenarios of “what if?” – and astonishingly simplistic in their understanding of the internet. Far too often we have heard that “nothing is private” on the internet without any serious consideration of what it would take for someone to steal a piece of identifiable information about someone in the context of research.

A second obstacle for us has come from university administrations who seem entirely willing to devote extensive portions of their budgets to quantitative software, but are almost entirely unwilling to do the same for qualitative software. At SFU, 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 qualitative data analysis program like NVivo on it, even though Ted has found an extensive group of individuals across at least a dozen 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 courses the same way that SPSS is incorporated into our curriculum. Ted’s department last year decided to finally invest in six leases for its own graduate computer lab so that Ted can teach the program in his graduate qualitative methods class, and another dozen people in the department decided to invest in personal leases, but not a penny 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 thought 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 the qualitative research landscape, and an inevitable part of its future. We hope our introduction to it has tweaked your interest and whetted your appetite for how they might benefit you and your research.
**Conclusion: What the Future Might Hold**

Trying to envision the possibilities and stay ahead of the curve in the realm of technology-assisted social research is both enjoyable and a challenge. Some issues we see around the corner are:

More and more of a trend toward cloud computing. We see this as a regressive trend that is good for those who run the 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 server’s use policies ... and that includes their universities, who we have found to be far too willing to pander to legal power and authority and far too unwilling to take an aggressive stance toward protecting research participant rights and interests. To the extent this becomes inevitable, researchers should choose their servers well.

We have only begun to consider the possibilities that are opening up because of the virtual revolution of portable computing power brought about by the iPhone and Blackberry 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 then and there by text, voice or video. Voice recognition has the potential to make keyboards obsolete and open up data gathering possibilities even further.

**References**


