

**Using NVivo 9 to Analyse Academic Texts concerning
Justifying, Negotiating and Designing Ethical Approaches
to Ethnographic Research Online**

Andrea Nevitt – May 2012

Word count: 4,856

Introduction

This paper comprises a literature review of the ethics of ethnographic research on the Internet. The literature reviewed has been handled and coded using the qualitative data analysis software NVivo 9. First will be a narration of the coding and analysis process, followed by a critique of that process and the software. Second will be a review of academic literature exploring the necessity of researching on the Internet in contemporary culture, the obligation of the research community to develop ethical guidelines for that research, and the unique aspects of the Internet that complicate the process of ratifying those guidelines.

NVivo in practice

Sources

The literature analysed is interdisciplinary. Sociology and media studies are augmented by the fields of communication, culture, marketing and law. Spanning the last decade, the literature is academic discourse on what is, due to the constantly-changing nature of the Internet, an always fluctuating subject area.

Despite the capabilities of the software to code books, chapters and journal articles line by line, I imported my notes on the literature rather than importing the literature itself. My preliminary reading process is already a coding of sorts where irrelevant data is screened out. The notes comprise

summaries and direct quotes, with annotations about the text and its relevance to the subject in question.

Coding

I began the NVivo project by importing Word documents into the existing 'Internals' folder and labelling them with author name, year of publication and title. Case nodes were created from each source and the cases then classified using node classifications by year of publication, source type and field (Fig 1).

Fig 1: Classification sheet

Node Classifications					
Name	Created On	Created By	Modified On	Modified By	
Literature source	14/04/2012 20:32	JAN	21/04/2012 12:26	JAN	
Literature source					
	A: Year of pu...	B: Type of pu...	C: Field of...	D: Field of stu...	E: Field of stu...
1: James, M and Busher, H (2009) Online Interviewing	2009	Book	Sociology	Education	Not Applicable
2: Dicks, B et al (eds) (2005) Qualitative Research and Hypermedia - Ethnography for the Di...	2005	Book	Sociology	Not Applicable	Not Applicable
3: Kendall, L (2004) Participants and Observers in Online Ethnography - Five Stories About I...	2004	Chapter	Sociology	Not Applicable	Not Applicable
4: Markham, A N (2004) Representations in Online Ethnographies - A Matter of Context Sen...	2004	Chapter	Sociology	Media	Not Applicable
5: Bakardjians, M and Feenberg, A (2000) Involving the Virtual Subject	2000	Article	Sociology	Technology	Communication
6: Hine, C (2000) Virtual Ethnography	2000	Book	Sociology	Not Applicable	Not Applicable
7: Rheingold, H (2000) The Virtual Community - Homesteading on the Electronic Frontier	2000	Book	Sociology	Technology	Not Applicable
8: Hadas, L (2009) The Web planet - How the changing Internet divided Doctor Who fan fic...	2009	Article	Media	Culture	Not Applicable
9: Mitell, J (2009) Sites of participation - Wiki fandom and the case of Lostpedia	2009	Article	Media	Not Applicable	Not Applicable
10: Gray, J and Mitell, J (2007) Speculation on Spoilers - Lost Fandom, Narrative Consump...	2007	Article	Media	Culture	Not Applicable
11: Christians, C G and Chen, S-L S (2004) Introduction - Technological Environments and t...	2004	Chapter	Media	Communication	Technology
12: Kleinman, S S (2004) Researching DURNET - A Case Study of a Multiple Methods Appr...	2004	Chapter	Media	Not Applicable	Not Applicable
13: Bassett, E and O'Riordan, K (2002) Ethics of Internet Research - Contesting the Human...	2002	Article	Media	Technology	Not Applicable
14: Kozminski, R V (2010) Netnography - Doing Ethnographic Research Online	2010	Book	Marketing	Not Applicable	Not Applicable
15: Hozer, B and Nitschke, T (2010) Questions on Ethics for Research in the Virtually Come...	2010	Article	Information Syste	Law	Not Applicable
16: Tell, M, Fleanu, F. & Hoeken, D. (2007) The Internet as a Library-of-People - For a Cylb...	2007	Article	Information Syste	Not Applicable	Not Applicable
17: Langer, R and Beckman S C (2005) Sensitive Research Topics - Netnography Revisited	2005	Article	Communication	Journalism	Not Applicable
18: Short, B F (1999) Beyond Netiquette - The Ethics of Doing Naturalistic Discourse Resear...	1999	Article	Communication	Not Applicable	Not Applicable

I have previously encountered discussions regarding researching on the Internet and therefore had several topics of interest into which I expected coded data would fall (for example the ethics of covert observation), but I had no preconceptions as to what overall results the coding would yield, hence my approach was between an inductive and deductive one. The first nodes created were those I intended to develop into tree nodes, these I

stored in the folder 'Review nodes' before I started coding. During coding I added child nodes to the initial tree nodes as the work progressed (Fig 2).

Fig 2: Sample of 'Review nodes'

Name	Sources	References	Created On	Modified On
Consent and announced presence	0	0	18/04/2012 19:18	10/04/2012 20:55
Archive use vs current postings	1	1	18/04/2012 19:18	17/04/2012 17:46
Effects of consent and announced presence	6	9	18/04/2012 19:18	17/04/2012 17:46
Examples	2	3	18/04/2012 19:18	10/04/2012 21:19
Sustained consent	2	2	18/04/2012 19:18	10/04/2012 20:55
When covert study is legitimate	1	1	18/04/2012 19:18	10/04/2012 20:55
Context of the debate on ethics	0	0	18/04/2012 19:18	10/04/2012 20:55
Development of research online	5	15	10/04/2012 19:18	11/04/2012 18:21
Ethnography in general	0	22	18/04/2012 19:18	11/04/2012 18:58
Insufficient guidance	3	4	18/04/2012 19:18	17/04/2012 17:46
Legality	4	11	18/04/2012 19:18	17/04/2012 17:46
Offline techniques are problematic online	11	18	18/04/2012 19:18	17/04/2012 17:46
Relationship of technology to research	5	10	18/04/2012 19:18	10/04/2012 20:55
Semantics	2	9	18/04/2012 19:18	11/04/2012 18:43
Subject areas	1	1	18/04/2012 19:18	10/04/2012 20:55
Exploitation	1	2	18/04/2012 19:18	10/04/2012 21:31
Fan studies methods	3	6	18/04/2012 19:18	14/04/2012 14:02
General history of technology	3	6	18/04/2012 19:18	10/04/2012 20:55
Interpretation	0	0	18/04/2012 19:18	10/04/2012 20:55
Considerations for representation	3	6	10/04/2012 19:18	15/04/2012 10:26
Potential limitations	6	7	18/04/2012 19:18	14/04/2012 14:47
Researchers construct the field	5	15	18/04/2012 19:18	15/04/2012 10:26
The position of the researcher	5	14	18/04/2012 19:18	14/04/2012 11:09
The process of analysis	4	7	18/04/2012 19:18	11/04/2012 21:02
Methodological recommendations	0	0	08/04/2012 19:36	15/04/2012 12:32

Tutorials and early practices with the software taught me that more efficient conceptualisation is possible if data are coded into relatively detailed nodes from the start, rather than broad coding requiring double-handling of data at a later stage¹. This was analogous to the manual analysis method with which I am familiar, and also complemented the deductive/inductive approach described above. I opted for descriptive names for nodes, preferring to avoid ambiguity when possible. For example, where there may have been one node within a tree about confidentiality called "Public access" I opted instead to specify the different meanings of the content in several nodes:

¹ Though this may not be the case for projects with a larger data set when wider categories may be more useful to begin with in order to get more of a handle on the data before the analysis proper starts, as suggested by Richards (2005: 97).

“Public accessibility = no consent needed”, “Public accessibility = perceived privacy”, “Public because public interest” and so on.

With context paramount to a literature review I coded entire paragraphs when necessary to ensure adherence to the original meaning. When analysing literature for a review, I believe line by line coding runs the risk of obscuring the author’s overall meaning since comments taken out of context can often appear to contradict the author’s conclusion (that may be coded elsewhere). There were occasions where I coded only a few words that were part of a list, for example the relevant words from the following extract were coded at “Anonymity of subjects”, “Private vs public” and “Size of data source”:

In a recent report on the ethical and legal aspects of human subject research on the Internet Frankel and Siang (1999) point to three features of the Internet that pose considerable difficulties in this respect: *the blurred distinction between the private versus public domain, the ease of anonymous and pseudonymous communication, and its global reach* (Bakardjieva and Feenberg 2000: 233, my emphasis).

Themes emerged from the start of the coding process, and during this initial pass some nodes were merged – for example two different advantages of online research were merged into one node for all advantages of online research¹.

To avoid the overall sense of the literature being lost via a heavy reliance on detailed coding I created a memo for each source into which I pasted what I considered to be the key quotes. These memos were referred

¹ And later separated again during my node reviewing stage.

to whenever I needed reassurance that analysis wasn't deviating from authorial intent. Crucial to this procedure was a process of constant review and revision as recommended by Richards (2005: 85-103).

Reviewing

The first pass of coding illuminated several areas requiring further consideration and deeper analysis, specifically the need to justify research online, ways of conceptualising the Internet that required definition, and varying approaches to appropriating data 'freely available' on the Internet. However, a brief review of the first pass also revealed that in order to more usefully focus on the final literature review some nodes would need to be set aside. Rather than overwrite work already done, and risk losing potentially necessary data, I created a new folder of nodes called "Structure nodes", into which I copied only the information pertinent to the structural core of the literature review. Again I found it more convenient to organise nodes in trees, gathering together data relating to significant sections of the emerging analysis.

Pertinent review nodes were copied into the structure nodes folder and reviewed in detail. All references were checked to ensure the coding was true to context and the references belonged to the particular node, and to create more descriptive nodes where necessary. For those nodes at which several points of view had been coded I created sibling nodes where opposing arguments could be coded. An example of this can be seen in the structure node "Concepts to define\Text or space", which contains nodes

summarising three emerging arguments: “Hybrid or cyborg because...”, “Space or human because...” and “Text or object because...” (Fig 3). The references in these nodes originated in the review node “Obstacles...\Internet as ‘space’ vs ‘text’” that was created originally to code a discussion by Bassett and O’Riordan (2002) on whether the Internet should be viewed as a site of humans or a site of human productions, but at which I coded several other sources that, while not specifically referring to the space/text dichotomy as described by Bassett and O’Riordan (2002), defined their conceptualisations of the Internet in similar terms.

Fig 3: Sample of structure nodes

Name	Sources	References	Created On	Modified On
Bringing ethnography online	0	0	12/04/2012 1	12/04/2012 19:23
Concepts to define	0	0	12/04/2012 1	12/04/2012 20:33
Confidentiality	0	0	15/04/2012 1	15/04/2012 17:48
Data recording	1	1	15/04/2012 1	15/04/2012 10:17
Privacy	0	0	15/04/2012 1	15/04/2012 13:24
Text or space	1	1	15/04/2012 1	15/04/2012 20:13
Emails are a hybrid	1	1	16/04/2012 1	16/04/2012 19:20
Hybrid or cyborg because...	10	29	16/04/2012 2	21/04/2012 13:16
Between real world and online conventions	6	9	16/04/2012 2	19/04/2012 18:48
Both culture and artefact	5	7	16/04/2012 2	18/04/2012 19:24
Connected to the body	1	2	16/04/2012 2	16/04/2012 20:58
Fuller understanding is gained	3	5	16/04/2012 2	16/04/2012 20:44
Virtual is insufficient	1	2	16/04/2012 2	16/04/2012 21:03
If hybrid then you are dealing with people	1	3	16/04/2012 2	16/04/2012 20:46
If Internet is space then censorship needed	1	1	15/04/2012 1	16/04/2012 20:14
If it is text recognise the author	1	1	16/04/2012 2	16/04/2012 20:15
Questions to answer	4	4	16/04/2012 1	16/04/2012 20:34
Space or human because...	5	10	16/04/2012 2	21/04/2012 13:16
Text or object because...	3	13	16/04/2012 2	21/04/2012 13:16
The argument is unnecessary	1	1	16/04/2012 1	16/04/2012 17:57
Justifying online research	0	0	12/04/2012 1	14/04/2012 13:43

To finalise the coding process, and again to retain the overall meaning of the literature sources, I created a folder of “Summary nodes” at which I coded entire sources. To code at these summary nodes I triangulated my decisions using both the visualise tool to see the structure nodes at which the source had been coded, and the key quotes memos I had created. The purpose of summary nodes was to create a reference to the content of the source, as demonstrated in Fig 4 where I can see Bakardjieva and Feenberg’s (2000)

key points were concerns about the effects of research on participants, and methodological recommendations in favour of collaborating with participants.

Fig 4: Summary nodes

Source notes		
Name	Nodes	References
Bakardjieva, M and Feenberg, A (2000) Involving the	70	127
Bassett, E and O'Riordan, K (2002) Ethics of Internet	84	166
Christians, C G and Chen, S-L S (2004) Introduction -	31	42
Dicks, B et al (eds) (2005) Qualitative Research and	27	36
Gray, J and Mittell, J (2007) Speculation on Spoilers -	9	12
Hadas, L (2009) The Web planet - How the changing I	26	37
Hine, C (2000) Virtual Ethnography	37	58
Hoser, B and Nitschke, T (2010) Questions on Ethics	72	167
James, N and Busher, H (2009) Online Interviewing	84	140
Kendal, L (2004) Participants and Observers in Onlin	47	79

Node	Percentage coverage
Nodes\4 Query nodes\Construction of knowledge	100.00%
Nodes\4 Query nodes\Flexible and reflexive approaches	22.17%
Nodes\4 Query nodes\Free	8.90%
Nodes\4 Query nodes\Justifying online research	100.00%
Nodes\4 Query nodes\Rigid	7.61%
Nodes\5 Summary nodes\Concerned about affects on participants	100.00%
Nodes\5 Summary nodes\In favour of collaboration	100.00%

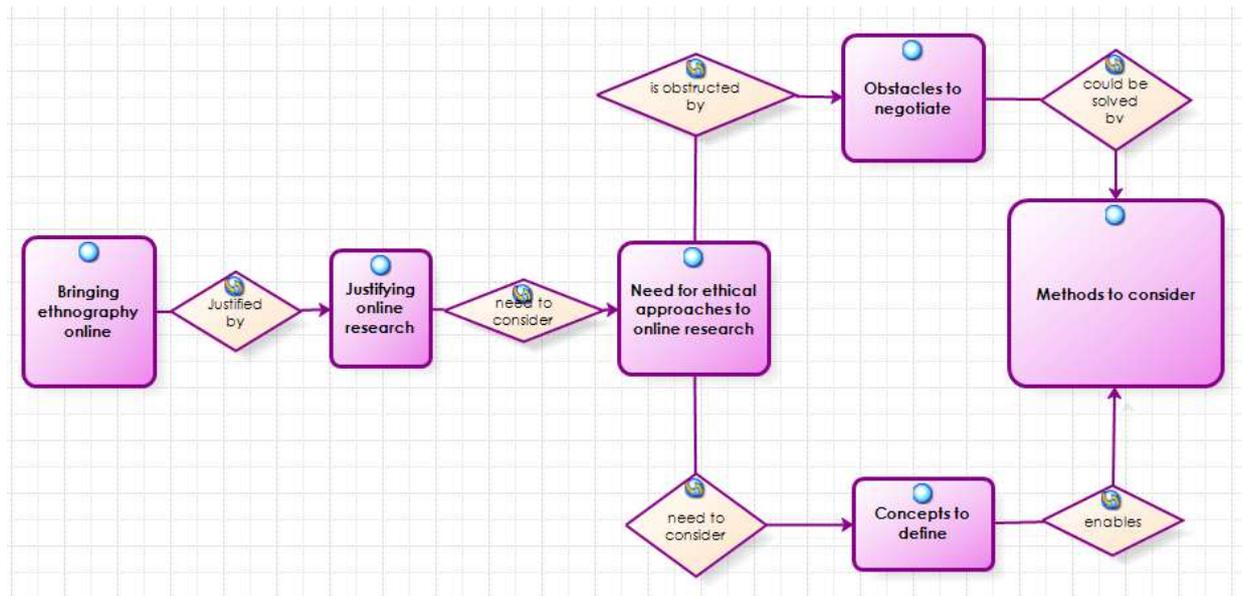
Modelling and querying

Using the node attribute for year of publication I ran queries to investigate potential chronological patterns across key issues brought to light by the coding, and saved the results of these queries as nodes in the folder "Query nodes". Except for those basic queries I have been reluctant to rely upon the query tool due to the potential for ambiguous results yielded by the relatively small number of sources in the project. Had the sources been literature rather than notes I am sure word frequency queries would have been illuminating, but as the sources were written mainly in my own words I deemed this to be of little value to the analysis¹. However I can see how this

¹ To illustrate, the top three most coded words of five letters or more were 'research', 'online' and 'internet'.

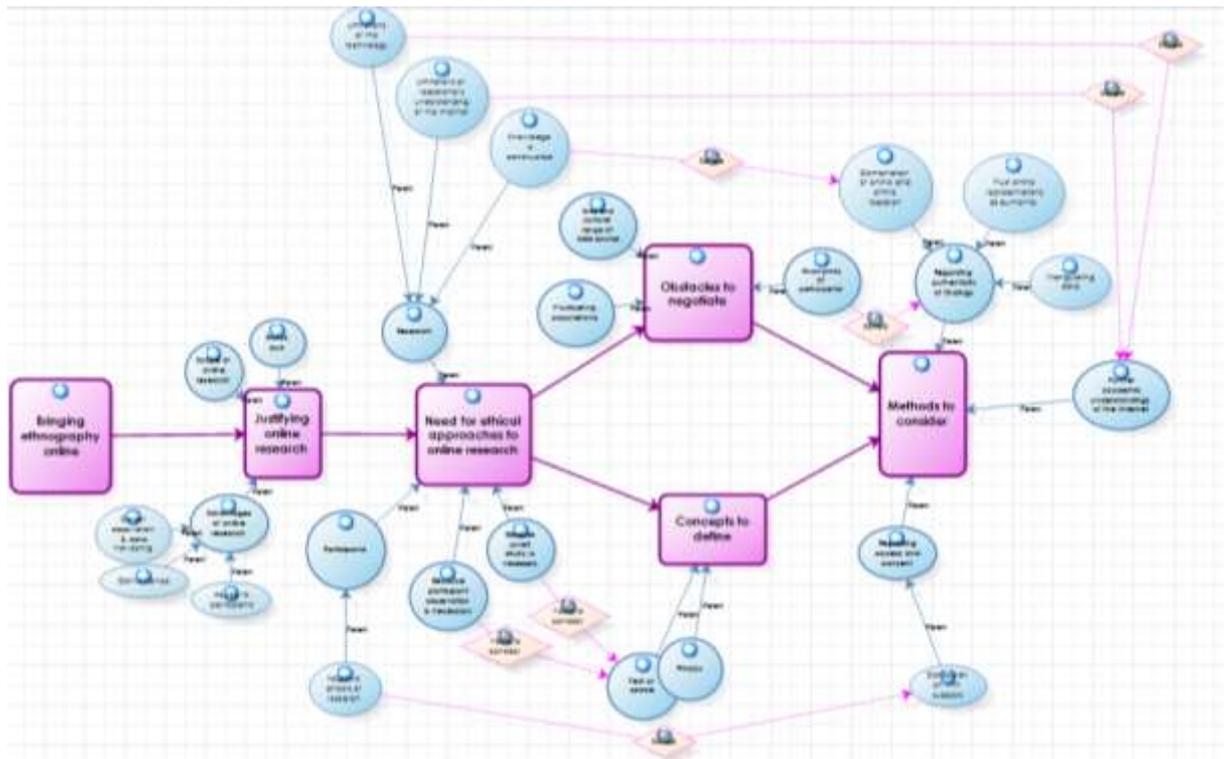
tool may be invaluable to the analysis of primary sources such as content analysis of interview transcripts.

Fig 5: Simple flow chart of the ethical consideration process



I found the most useful analytic capability of the software to be the model tool, with which I created two models: a simple flow chart (Fig 5) and a more detailed chart including the principal child nodes for each stage, and relationships suggesting how each issue may be countered with selected methodological approaches (Fig 6). Modelling is a process I have previously done by hand, and using NVivo to explore relationships in this way was a comparatively efficient use of time and allowed experiments not encouraged by hand drawing.

Fig 6: Detailed flow chart of relationships between stages of the ethical consideration process



Reflections on NVivo

This early in my experience of coding I find it difficult to reflect on the process except to outline the general apprehensions I attempted to counter. While coding I had a constant worry that too concentrated a focus on the minutiae would obscure the running themes and reasoning processes of the theorists. To counter this I used general nodes at which to summarise the essence of a source and memos into which I pasted important points. In this way NVivo was able to complement my usual manual process by enabling me to 'zoom' in and out of the data as necessary.

From an analytical point of view, my tendency was to code an extract at an existing node and only create new ones when it was absolutely necessary. During the first pass of coding I began to feel certain data were

being forced into categories rather than fitting more easily into new ones. When I reviewed the initial coding and created structure nodes I found a balance between these opposing techniques and saw much more descriptive patterns emerging.

Because the software displays quantitative information alongside qualitative I was conscious of the temptation to use numbers to substantiate claims. While figures were useful in highlighting areas of interest, I often noted to myself how 'popular' certain nodes were. To avoid drawing conclusions that weren't necessarily accurate I often checked nodes to get a feel for how often an author may repeat the same argument or quote another author's similar point of view. This, however, seems more a concern with my beginner-level understanding and use of the software, rather than an inherent problem of the programme, and it certainly encouraged constant reflexive engagement.

The detail with which the software allows one to sort through data cannot fail to be useful for handling large amounts of raw data, when used appropriately and without reliance on the technology to do everything. My overall conclusion on the use of NVivo though is concerned with efficiency of time. This was my first experience of formal coding and I found the process slow-going. While I am led to believe by colleagues and authors (Dey 1993; Richards 2005) that this is necessarily the case, I cannot help but conclude that for a literature review of so few sources my usual manual processes would have been more efficient. The structure of the programme was useful,

particularly the memo function and 'see also' links¹, and this leads me to recommend NVivo as primarily an organisational tool, secondary to its function for analysis, when working on a literature review.

Literature review

Context

Thriving on reflexivity (Hine 2000: 13), ethnography is a method through which researchers aim to gain a complete understanding of the context of a social subject in order to produce a deep and detailed description (a thick description) of that subject. Ethnographers can use a selection of qualitative and quantitative research and analysis methods. 'Participant observation is the most common component of this cocktail, but interviews, conversational and discourse analysis, documentary analysis, film and photography all have their place in the ethnographer's repertoire' (Kozinets 2010: 59).

It has been noted across disciplines that there is a growing tendency towards social life online. Cultural, sociological and communication theorist Howard Rheingold, writing in 1993, describes how those from the television era were 'beginning to migrate to CMC² spaces that better fit their new ways of experiencing the world' (2000/1993: xxvii). Two decades later, in the fields of information systems and law, Hoser and Nitschke point towards the

¹ I often found myself wishing there was the capacity to link a selection of text to another selection of text in another source, rather than linking a selection to a whole item. There may be a way around this need using nodes, but I haven't yet discovered it.

² Computer mediated communication.

availability of the Internet and vast data streams enabling a trend towards life in social network sites (2010: 181). It is this shift in focus from society in situ to society created by interaction (Markham 2004: 144) that has necessitated an evolution of ethnographic methods in order to account for the Internet as a cultural context.

The application of offline methods to online situations is not, however, a simple one¹. Discourse at the beginning of the last decade was apprehensive about ethical concerns, such as privacy and confidentiality, already issues in offline research, being 'brought into particularly broad relief by the ambiguous nature of the electronic medium' (Sharf 1999: 245). Socio-technological theorists Bakardjieva and Feenberg described how an 'online data rush which treated every content found on the Net as open to downloading, analysing and quoting has been countered by an ethical perfectionism leaving almost no space for research on virtual forums' (2000: 233). In addition to these early concerns, more recently theorists are wary of early explorations of online social subjects that ignored the offline context of those subjects, viewing the Internet as a subject privileged over activity overlapping offline actions².

¹ For a persuasive discussion about the need to refer to online research as 'cyberethnography' rather than 'virtual ethnography' see Teli et al (2007).

² See for example Teli et al's (2007) critique of Christine Hine's seminal *Virtual Ethnography* (2000).

Justifying online research

With concerns about the potential negative effects of online ethnographic research on both the research participant and the research results, it is paramount that the relevance and benefit of such research be established. Of primary significance is the wider scope of research into online activity. If cyberspace and the relationships formed within it are becoming (or indeed already are) a part of everyday life (Kozinets 2010: 38), then online ethnography offers the potential for understanding the complexity of the emerging social worlds in the terms of sociology (Dicks et al 2005; Hine 2000), marketing (Kozinets 2010) and communication studies (Sharf 1999). Indeed, social and media theorist Annette Markham states:

If we do not grapple with natural and necessary changes (in both the social structures we live in and study and the research norms we practice) wrought by our growing connections with communication technologies in an epoch of decentred authority and multiauthored realities, our research will not reflect the complexities we strive to understand (Markham 2004: 154).

In other words we cannot hope to answer the bigger questions about what it means to be on the Internet without adapting our research to it.

Once we have justified online research in terms of the questions it can answer, we are able to outline the advantages of research using the Internet rather than more traditional methods. By far the most frequently cited advantages of online research relate to its practical ease in accessing participants, its convenience and the ability for covert observation and data harvesting.

The ease with which researchers can access marginal social groups is particularly highlighted in Langer and Beckman's (2005) critique of online ethnographic methods used to research cosmetic surgery. They outline how the Internet allowed them to find online discussions between participants who would be firstly almost impossible to locate for face to face discussions and secondly difficult to engage in formal discussions about a topic often seen as taboo (Langer and Beckman 2005: 191-192). Related to this ease in locating research subjects is the capacity of the Internet for data harvesting. We can "'just" use log files' (Hoser and Nitschke 2010: 3) and harvest 'readily available logs of conversation' (Christians and Chen 2004: 19) that are potentially free from the taint of the researcher's presence¹.

In addition to the convenience of covert observation and data harvesting, the advantages of Internet technology over traditional methods concern efficiencies of time for data capturing, recording and handling. Of particular interest websites automatically store archives of online conversations and interactions, providing 'a near complete record of online social interactions [...] far easier than the surreptitiously recorded fieldnotes and fragmented recollections of the in-person ethnographer' (Kozinets 2010: 72). One could suggest that easy access to such a volume of data, and the time saved compared to face to face interactions make online ethnography the ideal tool for social research. However, adapting traditional ethnographic methods to the online environment provokes debate about

¹ It is outside the scope of this paper to go into detail regarding the effects of the researcher's presence on the research. For an entertaining and descriptive discussion about this see Markham (2004).

how to resolve ethical dilemmas in this new context (James and Busher 2009: 56).

Issues of ethics in online research

Ethical approaches to research are primarily concerned with potential negative effects of the research on the research participant, and with the potential implications for the research findings of an approach that has not been considered rigorously. While some of these concerns may be routine to experienced researchers, the potential online for exploitation of participants and misuse of the technology's more convenient facilities has re-ignited debate.

Much of this debate occurred relatively early in the use of the Internet for social research (see Bakardjieva and Feenberg 2000; Langer and Beckman 2005; Sharf 1999), though it is still an important contemporary consideration (Hoser and Nitschke 2010; James and Busher 2009). A frequent concern is the neglect of confidentiality and privacy rights, and a subsequent loss of the individual's control of the self – taking personal stories and representing and analysing them out of context, effectively distancing an aspect of the self from its author and altering its intended meaning – what Bakardjieva and Feenberg refer to as 'alienation' (2000). This is highlighted particularly in online ethnography because, as Markham notes in her argument for context sensitivity, 'In text-based computer mediated social spaces, Self and Other are constructed through interaction more obviously' (2004: 147) and, 'While we are taught to think of language as an abstraction,

online, language is the reality. Online, discursive practices create place, self, and embodiment' (2004: 152-153).

Other possible effects on the research subject that are particularly relevant to online research include the perpetuation of the underrepresentation of minority groups through the suppression of alternative media (Bassett and O'Riordan 2002: 244), and the potential impacts of increased visibility on minority groups who see their online activity as contained within a safe environment, regardless of the open accessibility of the websites they frequent.

When considering the validity of research findings, ethnography using the Internet, and particularly in relation to participant observation, raises similar issues to ethnography offline. Sociologists and media theorists are concerned with the construction of knowledge being subject to the researcher's position, from the planning stage (Dicks et al 2005) through to the effects of their presence on research subjects (Bakardjieva and Feenberg 2000; Markham 2004), and their interpretation and presentation of results (Kendal 2004; Markham 2004):

Frankly, whether or not the researcher participates or simply observes, the construction of the research report will present a particular reality of the object of analysis that is influenced by the identity and participation of the researcher. Thus, the effort or unconscious decision to absent oneself from the field will not remove the researcher from the process and product (Markham 2004: 145).

What complicates the issue for social and cultural research online is that knowledge is traditionally viewed as constructed in relation to the body.

Kozinets asserts that mediated communication is not new; that communications from letters through to phone conversations 'suffer from the same dubious anonymity and lack of embodiment as textual communications and interactions online' (2010: 69). Markham goes further, suggesting that the online environment 'restricts many of the senses that would traditionally help the scientist make sense of place, Other and context' (2004: 151). Concordantly, if the integrity of research is reliant upon the researcher's ability to verify results, verification is problematized when the technology grants anonymity and distance to participants. Markham warns that integrity is endangered in studies of online contexts if the site of the authentication of 'essentially disembodied relationships and cultures' is essentially the body of the researcher (2004: 151-152, my emphasis).

Obstacles and definitions

As there are ethical dilemmas that are especially pertinent when considering researching online, there are also aspects of the Internet that complicate the process of resolving these dilemmas to create the ideal ethical method. If the medium obscures the messenger (Kozinets 2010: 70), how are we able to find and fully comprehend that messenger? Offline we are able to ask participants to sign forms giving their informed consent, we are able to use our visual and aural senses to judge their comfort level in interviews and inform our understanding of utterances. On the Internet we only have participants' textual input from which to draw our conclusions.

How can the researcher balance the traditional scientific impulse to know what is real (read: physical, authentic, embodied) through meaningful but disembodied interactions among personae who may or may not correspond to their physical counterparts? (Markham 2004: 142).

Academics also highlight that even were we able to trust data gathered online as authentic, the amount of data available can lead to 'data-overload' (Dicks et al 2005: 118) and hence a prohibitive number of participants for whom risk would need to be evaluated and consent obtained (Bassett and O'Riordan 2002: 235-236). Due to the fluctuating nature of online communities we must also consider that participants may have even ceased posting on the website under study (Bassett and O'Riordan 2002: 244).

Research is further complicated by inter- and intra-disciplinary inconsistencies in the way text-based interaction on the Internet is conceptualised. A primary difference is whether we should view online discussions on publicly accessible forums as private interaction and subject to consent for data harvesting, or as public interaction and hence free from constraints on use. Arguments fall loosely into two attitudes. The first is that no consent is required for observation because the discussion is publicly accessible (Kleinman 2004), because seeking consent would be prohibitive and could adversely affect the research findings (Langer and Beckman 2005) and because the law is likely to consider the academic use of the discussion as fair use (Kleinman 2004: 57). The second presumes consent is required if the discussion is understood by contributors as private (or at least quasi-

private (Sharf 1999: 246)), despite being publicly accessible. Hoser and Nitschke relate this to the social convention of the private sphere, to which individuals are entitled as long as they adhere to the rules of society (2010: 180-181).

Proposed methods

Answering concerns about privacy, Bassett and O’Riordan suggest that a resolution is required on whether the Internet should be treated as a ‘space’ of human interaction where the text is ‘all but conflated with the mind or indeed the soul of the author’ (2002: 237) and is hence subject to the ethical constraints of the human subjects research model, or as ‘text’ – a ‘piece of property belonging to, yet distinct from the author’ (2002: 238) and subject instead to the constraints of copyright law. Against a prohibitive resolution that would require seeking consent from all participants, Bassett and O’Riordan highlight that ethical constraints must avoid the inadvertent erasure of minority groups from accounts of the online social world by considering all ‘textual output as private social interaction’ (2002: 245). They argue that a way needs to be found to account for the hybrid nature of the Internet – containing objects that are neither virtual selves, nor objects that are distinct from their authors (2002: 245). This call is usefully answered with Teli et al’s conceptualisation of the Internet as a cyborgic entity, something between humanity and humanity’s product (2007).

Sociologists are major proponents of an approach to research that involves collaboration with research subjects in order to gain a clearer

understanding of the group's perception of privacy and to avoid unintentional alienation and misinterpretation. James and Busher go so far as to suggest that developing personal relationships with participants in order to convey the competence and integrity of the researcher is an alternative to rigid ethical frameworks (2009: 59). Collaboration enables participants to engage critically with and influence the research process (Bakardjieva and Feenberg 2000: 238) and perhaps most importantly apprehensions about the research can be allayed with information about how the research project can be beneficial to the group as well as the researcher (James and Busher 2009: 64; Sharf 1999: 253). To support the veracity of the research some researchers encourage participants to check academic interpretations of their contributions (James and Busher 2009: 67; Kendal 2004; Sharf 1999: 251), though as Kendal reminds us, the research findings may not be meaningful to them:

[M]ost of what I have to say about BlueSky participants is more or less irrelevant to how they see themselves and to what they find important about their online experience. My translation of their culture into terms that are meaningful to me as an ethnographer, and then again into terms that are meaningful to various audiences results in stories that are often no longer particularly meaningful to the participants whom those stories purport to be about (Kendal 2004: 139).

As an alternative method of ensuring the authenticity of results, Markham answers concerns regarding the reliance on the senses to verify data by suggesting:

in cultural spaces negotiated by information exchange, the ethnographer might be wise to go native; that is, to trust information as representation rather than use traditional senses as the most authentic filters for understanding and analysis (Markham 2004: 151).

However, if we are to consider the Internet as a cyborgic entity to resolve privacy issues, then this trust in information is insufficient in achieving a complete understanding of the context of an online group that is inescapably associated with bodies 'left behind' (Rheingold 2000/1993: *vxii*)¹. A more comprehensive approach of combining online and offline methods in order to triangulate data has been recommended through the last decade by Hine (2000: 155), Kleinman (2004: 60) and Kozinets (2010).

Throughout these considerations is it important to remember that despite users' assumptions of privacy they are voluntarily interacting on websites that usually encourage openness and public access (Bassett and O'Riordan 2002: 242-243). Collaboration with research subjects may answer issues about privacy in the short term for the website under scrutiny, but does not answer wider concerns about privacy online. As Hoser and Nitschke note:

[Internet users] conduct themselves using the implicit rules for interaction in the 'real world', disregarding the impact their behaviour may have, especially disregarding the fact that anybody can come across the informational traces they are leaving. This makes their personal sphere very easy to intrude (Hoser and Nitschke 2010: 181).

¹ It is possible that Markham's suggestion may be of more help if our object of study is an online community, rather than a community online, as defined by Kozinets (2010: 63-65), because the questions we aim to answer would be those which interrogate human social interactive elements that are unique to the online environment.

Responsibility therefore lies with the academic community to raise awareness and encourage discussions on the risks of Internet use (Bassett and O’Riordan 2002; Hoser and Nitschke 2010; Rheingold 2000/1993: xxvii; Sharf 1999: 252).

Conclusion

In this paper I have outlined the scope of research on the Internet, the need to develop ethical guidelines for that research and the obstacles that require negotiation before those guidelines can be ratified. The recommendations discussed above serve to answer some of the ethical concerns raised over the last two decades, particularly concerns about privacy and authenticity. However, this still falls short of answering all the demanding questions of the ethics board; so much is still to be decided upon. We are left with a catch 22 situation where more grounded academic conceptualisations of the Internet are needed before apposite ethical guidelines can be outlined, but in order to achieve these conceptualisations research is required that can use only existing ethical guidelines that are insufficient in ensuring authentic and verifiable results. I would have to agree with Markham that in the meantime the important factors of research design require ‘Context sensitivity, flexible adaptation, internal consistency, and reflexivity’ (Markham 2004: 146).

Bibliography

Bakardjieva, M. and Feenberg, A. (2000) 'Involving the Virtual Subject'. *Ethics and Information Technology*. 2: 230-240

Bassett, E. and O'Riordan, K. (2002) 'Ethics of Internet Research: Contesting the Human Subjects Research Model'. *Ethics and Information Technology*. 4: 233-247

Christians, C. G. and Chen, S-L. S. (2004) 'Introduction: Technological Environments and the Evolution of Social Research Methods'. In *Online Social Research: Methods, Issues, & Ethics*, edited by Mark D. Johns et al. 15-23

Dey, I. (1993) *Qualitative data analysis: A user-friendly guide for social scientists*. London: Routledge

Dicks, B. et al (eds) (2005) *Qualitative Research and Hypermedia: Ethnography for the Digital Age*. London: Sage

Frankel, M. S. and Siang, S. (1999) *Ethical and Legal Aspects of Human Subjects Research on the Internet*. Washington: American Association for the Advancement of Science

www.aaas.org/spp/dspp/sfrr/projects/intres/main.htm [accessed 29.04.12]

Hine, C. (2000) *Virtual Ethnography*. London: Sage

Hoser, B. and Nitschke, T. (2010) 'Questions on Ethics for Research in the Virtually Connected World'. *Social Networks*. 32: 3. 180-186.

<http://dx.doi.org/10.1016/j.socnet.2009.11.003> [accessed 04.04.2012]

James, N. and Busher, H. (2009) *Online Interviewing*. London: Sage

Kendal, L. (2004) 'Participants and Observers in Online Ethnography: Five Stories About Identity' in *Online Social Research: Methods, Issues, & Ethics*, edited by Mark D. Johns et al. Oxford: Peter Lang. 125-140

Kleinman, S. S. (2004) 'Researching OURNET: A Case Study of a Multiple Methods Approach' in *Online Social Research*, edited by Mark D. Johns et al. Oxford: Peter Lang. 47-62

Kozinets, R. V. (2010) *Netnography: Doing Ethnographic Research Online*. London: Sage

Langer, R. and Beckman S. C. (2005) 'Sensitive Research Topics: Netnography Revisited'. *Qualitative Market Research: An International Journal*. 8: 2. 189-203. <http://dx.doi.org/10.1108/13522750510592454> [accessed 04.04.12]

Markham, A. N. (2004) 'Representations in Online Ethnographies: A Matter of Context Sensitivity' in *Online Social Research: Methods, Issues, & Ethics*, edited by Mark D. Johns et al. Oxford: Peter Lang. 141-155

Rheingold, H. (2000) *The Virtual Community: Homesteading on the Electronic Frontier* (2nd Ed). London: The MIT Press

Richards, L. (2005) *Handling Qualitative Data: A Practical Guide*. London: Sage

Sharf, B. F. (1999) 'Beyond Netiquette: The Ethics of Doing Naturalistic Discourse Research on the Internet' in *Doing Internet Research*, edited by Steven Jones. London: Sage. 243-256

Teli, M., Pisanu, F. & Haaken, D. (2007) 'The Internet as a Library-of-People: For a Cyberethnography of Online Groups'. *Forum: Qualitative Social Research*. 8: 3. Art 33.

<http://www.qualitative-research.net/index.php/fqs/article/view/283/621>

[accessed 27.02.2012]