Web 2.0 as a Social Science Research Tool

Helene Snee
University of Manchester

ESRC Government Placement Scheme
The British Library

November 2008

CONTENTS

Executive Summary

This report explores Web 2.0 and social research, concentrating on the utilisation of platforms like blogs, social network sites and wikis in data collection. The primary foci of the project were the methodological implications of such research, for example the continuing relevance of 'offline' evaluation criteria in an online context, and the potential ethical issues of data collection in this new environment.

The aims of the research were: to publicise the benefits and issues of new technologies for the social science research community; and to report on current practice and thinking in the use of Web 2.0 technologies as a social science research tool. A literature review was conducted to examine methodological and ethical issues identified in online research literature; this then informed a series of qualitative interviews with fifteen researchers.

The term 'Web 2.0' is associated with 'social software' and user generated content, and greater participation and interaction between Internet users and the web. Interviewees reflected on the democratic potential of Web 2.0 and the blurring of boundaries between online and offline lives.

Illustrative examples of how Web 2.0 has been utilised show a variety of methods and applications employed by social scientists. Interviewees highlighted a number of advantages and rationales for such studies. These include: practical advantages, such as relative ease of access to user content; the wealth of data about everyday life that is increasingly available online; and the implications for the quality and proximity of relationships between researchers and participants. There are certain practical issues that can pose a challenge in Web 2.0, such as accessing and understanding these youthful environments, and the overwhelming amount of data that can be generated.

Whilst Internet use has become more widespread, the use of Web 2.0 is not ubiquitous amongst the wider population, and a potential methodological challenge is the representativeness of social research samples that draw solely upon Web 2.0 users. Other features of the Web 2.0 environment, such as the lack of boundaries to some populations, could result in sample biases. However, sampling from set populations and the larger sample sizes enabled by web research mean that such errors can be minimised.

Difficulties in verifying the identity of research subjects in Web 2.0 studies may affect the validity of data, although interviewees suggested that the 'identity play' associated with early Internet use is less of a concern than in the past. Whilst this may continue to be an issue that needs consideration if researchers wish to generalise from their findings, interviewees also commented on relying on participants to tell the truth in 'offline' research.

The interviewees fell into two camps on the use of evaluation criteria like representativeness and authenticity, and its application to Web 2.0 data. Either offline standards should continue to apply; or, alternatively, they would not apply such criteria to their research. Of greater concern to the interviewees was that, as with any social science method, research design should be 'fit for purpose'. Examining new and innovative techniques raises just how important it is for methods to be appropriate for the research question under investigation.

The blurring of public and private spaces on the Internet was identified as a key concern for ethical decision-making in Web 2.0. The nature of the Web 2.0 environment has further eroded the distinction between online and offline spaces. Personal lives are increasingly exposed in Web 2.0 applications as part of a broader cultural shift towards openness and changing notions of privacy. The traceability of Internet data means that it can be difficult for researchers to ensure that participants remain anonymous. However, interviewees raised the issue that some participants may have deliberately put their content into the public domain, and thus should be recognised as authors. This issue impacts on when social scientists need to obtain informed consent for the use of Web 2.0 content. Some interviewees felt that putting information online should not imply consent for the use of this material. An additional challenge was determining the right of researchers to collect data without the owners' permission.

There was some variation amongst interviewees with regards to where responsibility lies to protect personal information, but there was a general agreement that context was important. It was noted that the public / private issue was a 'grey area', and therefore interviewees favoured a context-driven approach to reflect the contingency of such issues, which is consistent with current social and Internet research guidelines. Factors that could determine the decisions made were: assessing the potential harm if participants are identifiable; whether the data is presented in aggregate; and how the research is to be disseminated. A final point with regards to privacy in Web 2.0 was the concerns expressed over the use of user content and information in some government and commercial research.

Interviewees suggested that existing guidelines could be more helpful to enable ethical decision-making, and that ethics committees may have a lack of knowledge about new Internet technologies. However, there was resistance to over-bureaucratic and restrictive approaches to social science research ethics. The changing relationships between researcher and researched in Web 2.0 were not seen by interviewees to suggest a need for new ethical principles, but that their application may need to be rethought in an online environment. Making sure there are safeguards which mitigate the potential ease of conducting online research without scrutiny needs to be balanced against constricting social scientists from conducting valuable work.

Considering the ethics of Web 2.0 research sheds light on existing research practices, such as the regulation of ethics; the need for research to be sensitive to context; and that guidelines need to be useful but not restrictive.

As well as a tool for data collection, Web 2.0 can also be utilised by social scientists to disseminate research in innovative ways, such as via online video and libraries in virtual worlds. An openness to sharing information means that social scientists can be more transparent about their practices and publicise their work to wider audiences. There is also potential for teaching and collaborative working applications of Web 2.0, although such uses are yet to be fully realised.

Interviewees felt that the potential of Web 2.0 as a research tool had yet to be fully explored. Future research could be hindered by the speed of technological developments; a perceived generation gap between researchers and 'Web 2.0 natives'; and the technical skills and knowledge of social scientists. Collaboration with other disciplines was one suggested way of exploring Web 2.0 further.

1. Introduction

This report forms the output of a three-month research project conducted as part of a joint initiative between the Economic and Social Research Council (ESRC) and the British Library to enable ESRC-funded research students to work with the Library on short Fellowships. The study was designed in response to the following remit from the ESRC / British Library guidance for applicants:

> *How are web 2.0 technologies being used by the UK social science research community? What are the implications for government and scholarly authority and research?*

The research placement took place 12 May - 15 August 2008.

The first section of the report outlines the objectives of the study and the data collection method. It then goes on to cover:

- Defining Web 2.0
    Definitions of Web 2.0 in the literature reviewed and interviews conducted
- Web 2.0 Research
    Current and emerging research in the field
- Methodological Issues
    Challenges of Web 2.0 methods
    Criteria used for evaluating social research
- Ethical Issues: Privacy in Web 2.0
    The problem of defining what is public and what is private
    The impacts of privacy issues upon ethical decision-making
- Ethical Approaches to Web 2.0 Research
    Formalised ethical procedures
    Wider implications for social research ethics
- Other implications of Web 2.0
    Social science and Web 2.0 beyond data collection
- Future research
    Interviewees' reflections on moving Web 2.0 research forward
- Conclusion
    Areas for exploration
- Appendices
    Participant details, interview topic guide and consent form, and glossary of Web 2.0 terms

2. Research Scope and Method

Two objectives were set at the beginning of the study: to publicise the benefits and issues of new technologies for the social science research community, so that researchers may be introduced to Web 2.0 or to new uses of such methods; and to report on current practice and thinking in the use of Web 2.0 technologies as a social science research tool.

The research consisted of two stages, a literature review and interviews with social science researchers. The literature review examined methodological and ethical issues in the online research literature, and looked at studies that discuss or utilise Web 2.0 as a research tool. This informed the topic guide used for the interviews; qualitative interviews were carried out with fifteen researchers.[1]

The focus of the research was to ask interviewees to reflect upon the evaluation criteria, such as data validity, applied to Web 2.0 research, and the potential ethical issues associated with this work. A semi-structured approach was taken, which meant that as well as the topics from the interview guide, conversations developed that touched upon a number of additional areas of interest.

All interviewees were given an information sheet and signed a consent form (see appendix 3) which gave them the option of remaining anonymous. Two participants requested this option and are referred to in the report as 'R1' and 'R2'; those happy to be identified are referred to in the report by their initials. Any quotes or specific comments that were ascribed to interviewees were referred back to them to ensure that they were happy with the presentation and context. The interviews were recorded and summarised, although not completely transcribed verbatim. The summaries were coded on the basis of the topic of discussion and then key points transferred into an Excel workbook to allow sorting into common themes and issues.

To avoid confusion, the following conventions are used: when referring to people interviewed as part of this study, the term 'interviewees' is used. References to 'researchers' should be taken to mean the wider research community, and 'participants' to those who take part in, and are the subject of, social research.

Given that the project was relatively small scale, the following report offers a high level overview of the subject and the topics raised in the interviews, and identifies some areas for further, more detailed examination.

---

[1] See appendices 1 and 2 for participant details and the interview topic guide.

3. Defining Web 2.0

Before discussing the ways in which Web 2.0 might be utilised by the research community, it is useful to briefly outline how it has been conceptualised in social science literature.

The development of the term 'Web 2.0' is usually ascribed to the American media company O'Reilly Media Inc. It was used by the company and its founder Tim O'Reilly to identify common features of a set of innovative Internet companies and their business characteristics, rather than describe a group of technologies. However, the term has come to be associated with 'social software' and user generated content, which share some of the features identified by O'Reilly, such as participation, the user as contributor, harnessing the power of the crowd, and rich user experiences (Anderson 2007: 6).

Social software "let[s] people interact with people and data in a fluid way" (boyd 2006: 17); commonly referenced examples included the social networking site Facebook[2] and the social bookmarking site Digg[3]. A recent OECD report on user-generated content notes that changes in how users produce, distribute, access, and use information, knowledge, culture and entertainment is seen to lead to increased autonomy, participation and diversity (OECD 2007 page 63). Bruns (2008) has called such changes a shift toward 'produsage', with the distinction between producers and users of content disappearing. According to Cooke and Buckley (2008), the key feature of Web 2.0 is that it enables collaboration and sharing online.

So alongside the technical move to the Internet as a platform, there is a sense of greater participation and interaction between Internet users centred on web content. The significance for the social sciences has been discussed by Hardey (2007) who suggests that this user agency brings Web 2.0 within fundamental debates because it is "*inherently social* so that users are central to both the content and form of all material and resources" (Hardey 2007: 869; italics in original). Despite the fact that it is a problematic term, and that the boundaries of the concept are unclear, it is "good enough" to talk about the new technologies and the behaviours that have "captured the imaginations of millions of users worldwide" (boyd 2006: 18-19).

The interviewees were asked to define Web 2.0 in an attempt to understand how it has been employed in the research community. These definitions broadly reflected the issues identified above. For example, a number of them commented on the use of Web 2.0 as a label. One issue raised here was the

---

[2] http://www.facebook.com
[3] http://digg.com

danger of subsuming different activities and applications under one term, thus losing a nuanced understanding of their differences; however it was also suggested that applying such a label as convenient shorthand enables some common trends to be discussed. Most identified one such trend as the growth in user generated content. This tied in to ideas about the more 'democratic' elements of Web 2.0; it is easier to put content on the Internet as the technological barriers have been removed, and the interviewees referred to a culture of collaboration, sharing and participation:

> [Web 2.0 is] a social medium which creates and facilitates interactions between people, in a way that the web was supposed to be in the early days. But [then] it was too high level, people had to have a degree in computer science to create anything [or] to start a conversation (AK).

As well as the technological developments that enabled more users to contribute to web content, many interviewees commented on the interactive facilities of Web 2.0 applications, and the growth of the 'online desktop'. Some also referred to the blurring of boundaries between online and offline lives, particularly how the everyday and the personal are becoming central to our Internet presence. As will be seen later, this was discussed in relation to (and has implications for) the opportunities for online research and the ethical decisions that social scientists must deal with.

4. Web 2.0 Research

The following discussion of how Web 2.0 has been utilised for data collection is not intended to be an exhaustive examination of this field but to highlight the variety of methods and applications that are currently being employed. It includes a summary of the uses, benefits and potentials of such methods as identified by the interviewees. Some of the work referred to here is drawn from the interviews and from unpublished studies. As noted in a number of the interviews, much of the work in this area is in progress, including postgraduate research and studies working their way through the publishing system[4]. Examples have been included from recent research in the following areas:

- social networking;
- blogging;
- use across Web 2.0 applications;
- virtual worlds; and
- the application of Web 2.0 in a specific discipline.

4.1. Recent articles

4.1.1. Research into social networking

Liu, H. (2007) "Social Network Profiles as Taste Performances", Journal of Computer-Mediated Communication 13(1) 252-275.

Liu analyses MySpace[5] social networking profiles, arguing that they are performances of taste that display the user's status and distinction. A qualitative pilot study identified four different types of 'taste statements' within the 'interests' section of these profiles, conveying: prestige; differentiation; authenticity; and theatrical personas. This is followed by a statistical analysis of 127,477 profiles which suggests that prestige and differentiation were the primary types of taste statements; differences in tastes are found to be aesthetic in nature but the evidence for socioeconomic factors to classify profiles is inconclusive.

Lange, P.G. (2007) "Publicly Private and Privately Public: Social Networking on YouTube" Journal of Computer-Mediated Communication 13(1) 361-380.

This study is based on an ethnographic analysis of videos, comments and profiles on YouTube[6], along with interviews with users, to look at the social networks created and negotiated. Lange considers the public/private nature of YouTube content, and suggests that rather viewing this as a dichotomy,

---

[4] See section 7 for the potential impact of Web 2.0 on publishing mechanisms.
[5] http://www.myspace.com
[6] http://www.youtube.com

user behaviour is more nuanced. Some behaviour is 'publicly private': identity information is disclosed but access to the videos or their personal context is limited; and some is 'privately public', as users connect with others freely but withhold their personal details.

Thelwall, M. (2008) "Social networks, gender, and friending: An analysis of MySpace member profiles" Journal of the American Society for Information Science and Technology *59(8) 1321-1330.*

Thelwall uses an automated method of data collection ('webcrawling') to download three samples of profiles from the social networking website MySpace. The main data set for the research was a random sample obtained using a program that generated random user ID numbers. The profiles sampled were automatically scanned to extract demographic information about the users, which was then analysed to provide descriptive statistics about the MySpace population. Findings from the study included a profile of a 'typical' MySpace user (who is female, 21, single and interested in online friendship) and a commentary on the different friendship dynamics within the community.

Hogan, B (2008) "Analyzing Social Networks via the Internet (Pre-release version)" in Fielding, N, Lee, R and Blank, G (eds). The Handbook of Online Research Methods*. Sage: Thousand Oaks, CA.   Available from http://individual.utoronto.ca/berniehogan/Hogan_SAGE_Internetworks_RC1.pdf.*

In a preprint of a chapter for an online research methods textbook, Hogan considers how the methodology of social network analysis (SNA) can be applied to online data. SNA is an established 'offline' method that attempts to map the structure and nature of relationships between people, organisations and other entities. To illustrate the application of SNA, Hogan conducts an explanatory analysis of the social bookmarking and link-sharing website Digg, which has been criticised for being dominated by particular individuals who set the agenda and vote for each other. His analysis examines the top submitters to the site, looking at the number of times their stories made it to the front page and the ratio of popular stories to those less successful. The analysis found that success was related to the network structure so that having a fan among top submitters carried more weight than having a non-top fan, but that sheer number of friends did not guarantee whether a story would be popular.

4.1.2. Research into blogging

Herring, S.C. and Paolillo, J.C. (2006) "Gender and genre variation in weblogs" Journal of Sociolinguistics *10(4) 439-459.*

This study examines the relationship between linguistic variation, gender and genre using quantitative analysis of a random sample of weblogs. The analysis builds on previous work by Herring and colleagues that suggested three sub-genres of blog writing: filter (which 'filter' content, often news stories, from the rest of the web); personal journals; and k(nowledge)-logs (which record information around a specific, often technical, activity). Herring and Paolillo's research concluded that the genre of a blog will affect its linguistic style, but that the gender of the blog author does not. They also reflect on how certain genres are considered to be gendered (personal journal writing is more associated with females, for example) which impacts on the prestige of the blog, as personal writing is seen as 'less important'.

Hodkinson, P (2007) "Interactive online journals and individualization", New Media and Society, 9(4) 625-650.

Hodkinson conducted an ethnographic case study of the experiences and perspectives of users on LiveJournal[7], a blogging and social networking site. He set up his own account (making it clear he was a researcher), wrote a blog, established contact with users and made 'friends', predominantly with those interested in the 'goth' subculture. The study examines the online goth community's transition from discussion forums to blogs in the context of debates about individualisation and identity. Hodkinson concludes that patterns of communication within the community are on the whole individualistic, but that existing group attachments do remain.

4.1.3. Research across Web 2.0 applications

Beer, D (2008) "Making Friends with Jarvis Cocker: Music Culture in the Context of Web 2.0" Cultural Sociology, 2(2) 222-241.

Rather than looking at a particular website, Beer focuses on the presence of a popular music performer across a number of Web 2.0 websites and applications. The article examines the implications of Web 2.0 for music culture, and the part that music plays in the development of Web 2.0 media. Beer also reflects on the capacity of Web 2.0 applications and content as data archives available for sociological analysis. The discussion looks at the ways in which users can engage in research about the performer through YouTube and Wikipedia[8]; form friendship networks around the performer through MySpace; and how relationships between performers and their audience have changed in the current environment.

---

[7] http://www.livejournal.com
[8] http://www.wikipedia.org

4.1.4. Research in Virtual Worlds

Chesney, T, Chuah, S-H and Hoffman, R (2007) "Virtual World Experimentation: An Exploratory Study" Available from:
http://www.nottingham.ac.uk/economics/cedex/papers/2007-14.pdf

This study examines whether virtual Web 2.0 worlds like Second Life[9] can be used for experimental research with human subjects. It reflects on the associated methodological issues and assesses the suitability of the environment to conduct such tests, concluding that they can simulate physical labs at a lesser cost. Incentivised games, which test economic decision-making, are replayed in 'virtual labs' with Second Life users and the results are compared with those from offline games. The behaviour of people in Second Life was typical compared with 'real life' studies conducted with university students (a common research pool for these sort of experiments), and any differences could be assigned to demographic differences rather than a different cultural environment. In fact, participant values were more in line with the general population than the average university campus[10].

4.1.5. Example of the application of Web 2.0 in a specific discipline

Cooke, M and Buckley, N (2008) "Web 2.0, social networks and the future of market research" International Journal of Market Research 50(2) 267-292.

The article proposes that Web 2.0 can be utilised by market researchers as collaborative tools not only to examine trends and consumer communities, but also to establish new interactive research methodologies such as 'participatory panels'. The use of such a panel in a research situation is described in a project that aimed to engage local business in consultations on community strategy. This employed social networking software to allow threaded discussions, surveys, e-mail communications and bulletins, Q&As, content publishing and user profiles. In such a process, respondents, researchers and clients have co-creation roles; this not only utilises the technical capabilities of social software but also incorporates its participatory elements, changing the nature of the relationship between researchers and researched.

4.2. Emerging research

As discussed at the start of this section, the relative novelty of Web 2.0 technologies means that work in this area of study is only just beginning to be published, or is the subject of ongoing postgraduate research. One interviewee noted that while social network sites have attracted a lot of

---

[9] http://www.secondlife.com
[10] See Giles, J (2007) for other uses of Second Life by social scientists.

attention, "less work is being done on wikis, folksonomies and mash-ups" (RB).

Work in these latter areas is beginning to emerge. For example, a project entitled "Tags, Networks, Narrative" at De Montfort University asked volunteers to tag 40 sites in Del.icio.us[11], a social bookmarking website, and analysed the resulting folksonomy. The output has been presented at a number of conferences (ICOT 2007). Two interviewees (MT and RB) had PhD students who were conducting work on folksonomies and wikis: one is investigating why people tag in various ways and if it is possible to encourage them to tag in ways that are more useful for others[12]; another is analysing user data downloaded from Wikipedia to map patterns of use and facilitate understandings of engagement with the technology[13]. Wikis have also been discussed with reference to their implications for knowledge production and dissemination (see section 7).

Work that examines the specific methodological and ethical implications of Web 2.0 research is also being published. A recent article in *Information, Communication and Society* examines some of the sampling problems when studying blogs and bloggers and suggests strategies to combat these (Li and Walejko 2008). Also with reference to blogging, Hookway (2008) uses his research into morality and everyday life that drew upon self-reflections in online journals as a case study to discuss the use of these sources in social research.[14]

4.3 Why use Web 2.0 in social science research?

As illustrated by the above examples, a variety of methods and approaches have been employed to utilise Web 2.0 in social science research. The interviewees were asked to reflect upon why researchers might want to employ these strategies, and three key themes emerged.

Firstly, the practical advantages of these methods were mentioned by a number of interviewees. The content generated by users in such environments is relatively easy to access; the researcher does not have to leave their office to investigate the social lives of individuals and communities. One interviewee commented on how "…*the standardised format of the content, for example how all MySpace profiles follow a particular template, means that data can be more easily collected using automated methods*" (TE).

---

[11] http://del.icio.us
[12] At the University of Wolverhampton.
[13] At the University of York.
[14] The issues raised in both of these articles will be discussed in more depth in the following sections.

Secondly, everyday life is now increasingly played out and documented online. The sheer wealth of such data was commented on: people now "…*express unprecedented information about themselves online*" (LP), and "…*millions of conversations and interactions are recorded*" (VM).

Thirdly, such methods have implications for the quality and proximity of the relationship between researchers and participants. "*The interactive nature of Web 2.0 allows subjects to have a more active role in research*" if desired, with researchers able to share results and get feedback (EM). Conversely, the fact that the content is already freely available means that unobtrusive research can be conducted, which can be seen as an advantage in an environment where over-surveying has led to a kind of research fatigue.

Web 2.0 technologies clearly offer a number of interesting avenues of research for social scientists; the following sections examine the potential methodological and ethical issues of such approaches, and present the outcome of discussions with the interviewees reflecting upon these concerns.

5. Methodological Issues

Although Web 2.0 offers social scientists opportunities for innovative research, it does come with some challenges. Some of these are on a practical level; for qualitative researchers, accessing and understanding Web 2.0 communities can be difficult. Youthful applications like Facebook can be "…cold spaces for outsiders doing research" (DB). Another issue interviewees raised was that the wealth of data available can be overwhelming as well as advantageous. Interviewees were, therefore, asked to comment on the various methodological concerns of online research.

5.1. Sampling and representation

One of the criteria that social scientists use to assess the validity of their research is the 'representativeness' of a sample. This is the extent to which generalisations about the wider population can be made; quantitative researchers attempt to avoid sample bias in which certain parts of the population are more strongly represented than others.   Although generalisations are less applicable to qualitative research, more in-depth studies still aim to say something about wider social trends.

The concept of the 'digital divide' suggests that certain groups are under-represented online: "…..access to the Internet is a matter not only of economics, but also of one's place in the world in terms of gender, culture, ethnicity and language….." (Mann and Stewart 2000: 31). However, it has now been argued that concerns such as this are slightly outdated and that representativeness in online research methods has become less of a problem as Internet use has increased in scope; a lot has changed since 2000. However, despite the growth in access to the Internet, recent studies show that use of Web 2.0 applications may not be widespread across all sectors of the population. According to the *Oxford Internet Survey 2007*, there has been a rise in the use of social networking sites in the UK but still less than one fifth of Internet users maintain a profile (Dutton and Helsper 2007: 57); UK blogging has actually declined slightly in recent years to around 12% of users (Dutton and Helsper 2007: 53). Use is also structured around age, for example; students are three times as likely as employed Internet users to have a social networking profile but almost no retired people do (Dutton and Helsper 2007:56).

Specific Web 2.0 practices have their own methodological considerations. Li and Walejko (2008) discuss the problems associated with sampling blogs and their authors. Firstly, there is no central 'list' to draw a representative sample from. Secondly, there are characteristics of the 'blogosphere' which mean that samples can be skewed, such as the presence of spam blogs ('splogs')

that are not genuine blogs but forms of advertising with often garbled content designed to direct readers to a particular website.

Interviewees suggested that representativeness could be a concern in Web 2.0 research. The nature of social networking profiles may result in sampling biases. For example, studies that utilise the publicly accessible data on Facebook will only capture a subsample of the Internet population who use Facebook and then a further subsample who make their information public. It could be argued that it is not only that some people do or do not choose these privacy settings, but that it is a particular 'type' of person who makes certain decisions: "*There may be many millions of people who use Second Life, but this is still only a small subsection of the population*" (JW). Even if the goal of research is not to generalise about the wider population:

> *Web 2.0 data sets still miss certain aspects of the phenomenon itself, such as the practices and behaviours of people who deliberately aren't involved or are indifferent (DB).*

Despite the points made above, some interviewees referred to ways in which Web 2.0 research may improve quantitative sampling techniques. For example, if sampling from a set population like users of a social networking site, it is possible to get a "*…genuine, mathematically random sample*" (MT). Due to the ease in collecting more data when it is in an electronic format, even if error rates are higher, having larger sample sizes minimises these errors. A number of interviewees also stated that mixed methods approaches that combine Web 2.0 data with other methods may address issues of representativeness and give a more rounded picture overall.

5.2. Authenticity

Another concern that is often discussed in relation to Internet research is that it is difficult to verify the identity of research subjects:

> *Anonymity in text-based environments gives one more choices and control in the presentation of self, whether or not the presentation is perceived as intended (Markham 2005: 809).*

Early Internet studies discussed how users experimented with their identity online, and were liberated from their offline embodied self; for example, one participant in Turkle's (1995) influential study had 4 different personas in online settings which switched genders, personalities and even species. On the other hand, anonymity may encourage candidness, so that people are more 'truthful' (Hewson et al 2003: 44).

Overall the interviewees suggested it was debatable whether the authenticity of participants was of considerable concern in relation to Web 2.0 research.

There was a sense that ideas about identity experimentation were a little outdated, especially considering the expansion of Internet use, and some doubts were expressed whether this identity play was "*…overstated to begin with*" (LP). Interviewees felt that users of Web 2.0 applications did so to present their everyday lives, even if this was a particular version of themselves. The only real issue of concern expressed was from a more ethical perspective, "*…for example if underage or vulnerable people were included in a sample*" (JG).

For studies aiming to provide some generalisation about populations, the authenticity of data may have a more significant impact on research results, for example with the demographic profiling of MySpace users in Thelwell's (2008) study discussed in section 3. In this case, there were concerns that some users entered false ages in their profiles, often as a joke or to get around the site's age restrictions. The data was, therefore, treated cautiously, although the pattern of user ages seemed to be credible, and consultations with previous studies found that fewer than one in ten profiles had incorrect age information. "*I've taken the attitude that... if it's reasonable... I'm assuming that in everyone's MySpace ages they're telling the truth*" (MT).

It was also noted by a number of interviewees that even in offline settings, researchers have to rely on participants to tell the truth or put some effort into carefully reflecting on their responses. This questioning of why online research should be less trustworthy than offline research has been discussed in the methodological literature. Hookway refers to Silverman's concept of an 'interview society' in which the only path to authenticity is through a face to face interview, and the stubborn tendency for other forms of research to be less trusted (Silverman 2001, cited in Hookway 2008: 97-8). One interviewee commented that the novelty of Internet research may mean that it comes under greater scrutiny, but in fact there are authenticity issues with offline research also (R2); online interviews are therefore not necessarily less 'authentic' than face-to-face.

5.3. Evaluating social research

Interviewees were asked to reflect upon whether evaluation criteria like representativeness and authenticity should be applied to Web 2.0 research. Responses generally fell into two positions: that offline standards should continue to apply to online criteria, or they questioned the use of such criteria in the first place; these interviewees would not apply such criteria to their research. This latter position reflected the position of qualitative researchers who may have a different viewpoint on what can be learnt about the social world, for example rejecting concerns of representativeness, generalisability and credibility and instead looking for strong, in-depth readings in their data (Denzin 1999: 115). None of the interviewees thought that new methodological criteria were required for Web 2.0 research.

Of greater concern to all the interviewees was ensuring that the methods employed match the research questions. Web 2.0 research, as with any method, "…*needs to be fit for purpose, in terms of design quality, process quality and reporting quality*" (R1). Although this does apply to any method, it was felt that there was a danger that as Web 2.0 is new and exciting, it might result in research being conducted just because the data is there or that it is "…*driven by the technology rather than the research goals*" (JW).

This idea of 'fit for purpose' can be applied to issues of representativeness and authenticity. For example, "…*if the research is examining MySpace itself, it is not an issue if a sample is not representative of the general population*" (MT). The importance of whether a blog is authentic or not depends on "whether a researcher is looking at how blogs work to produce particular effects or whether they are looking at how blogs correspond with an 'offline' reality" (Hookway 2008: 97). Some researchers also pointed to the importance of keeping in mind the context of data collection. One interviewee voiced a concern that "…*the ability to collect massive amounts of data means researchers may lose sight of the cultural setting it is drawn from*" (YL). In addition, online research has been put forward as a method which enables international research, but questions have been raised as to "…*whether the methods can be translated into non-Western contexts*" (R2).

In conclusion, although there are certain methodological concerns with conducting Web 2.0 research, reflecting upon these issues also highlights questions for social science as a whole, namely ensuring that the methods employed are appropriate. It may seem self-evident that research design is important, however, looking at new and innovative methods raises just how significant 'being fit for purpose' is. The other substantive topic of discussion in the interviews was the ethical considerations of Web 2.0 studies, and is summarised in the following sections.

6. Ethical Issues: Privacy in Web 2.0

> *The ability of both researchers and their subjects to assume anonymous or pseudonymous identities online, the complexities of obtaining informed consent, the often exaggerated expectations, if not the illusion, of privacy in cyberspace, and the blurred distinction between public and private domains fuel questions about the interpretation and applicability of current policies governing the conduct of social and behavioural research involving human subjects (Frankel and Siang 1999: 2)*

Many of the ethical issues raised in the interviews were centred on this idea of a 'blurred distinction' between private and public spaces on the Internet, a concern that is also reflected in online research literature. Eysenbach and Till (2001) identify this difficulty of separation as the "main problem with using Internet communities for research" (Eysenbach and Till 2001: 1105).

6.1. Web 2.0: the personal as public

> *"[A] log of daily life...   Reading people's diaries would have been impossible 5 years ago; you would have had to burgle their home. But if you want to read a million diaries today, you can do in MySpace" (MT).*

Whilst defining what is public and what is private may have been a concern for Internet researchers for some time, the nature of the Web 2.0 environment complicates this process further.   Social networking sites encourage users to share their lives with their networks, and personal weblogs document the minutiae of everyday life as well as sensitive or significant events. Interviewees noted the trend of the personal becoming public as something relatively new. The reference to burglary in the above quote is telling, as it suggests what has traditionally been kept behind closed doors is now out in the open. This is not something that is exceptional to the Internet however, but is:

> *… part of broader cultural shifts, alongside reality TV for example, that celebrate mundane life and the desire to put this in the public domain, signifying changes to the notion and value of privacy (DB)*

Interviewees noted that researchers may be more concerned about privacy in certain contexts than participants, given how open people are in Web 2.0. However, the issue of what can be classed as public on the web is important because it impacts on the ethical decisions that researchers must make.

6.2. Subject or author?

Social research nearly always involves studying human subjects, defined as:

> *… a living individual about whom an investigator… conducting research obtains data through intervention or interaction with the individual, or identifiable private information (Frankel and Siang 1999: 16).*

Ethical guidelines require researchers to protect subjects from harm; disclosure of an individual's identity represents potential harm, and measures such as anonymity should be employed to avoid this. (SRA 2003: 38-9). A number of interviewees highlighted that the difficulty with making Internet data anonymous is its 'traceability'; if a quotation is put into a search engine, the origin of the text can be easily found and subjects identified. This was seen to be less of a concern if data is presented in aggregate rather than referring to specific instances.

The blurring of the private / public distinction adds another dimension to the matter of anonymity. Bassett and O'Riordan (2002) argue in "Ethics of Internet Research: Contesting the human subjects research model" that not all Internet users necessarily want to remain anonymous, and may have chosen to deliberately publish in the public domain; their texts should, therefore, be cited in the same way as traditional print media. This issue was referred to in the interviews, such as the "…*intellectual property problem if original content is not recognised*" (VM) or "…*if people did not want to be associated with the content they post online, they would do this anonymously*" (WD).

The Association of Internet Researchers (AoIR) has suggested in its ethical guidelines that differentiation of subject or author-generated content is setting-dependent. For example, those participating in intimate chatroom settings may be best considered as human subjects, while those who write publicly accessible weblogs should perhaps be treated as authors (Ess 2002: 7). Where the research location falls on the public / private continuum thus impacts on the need to anonymise data; as discussed in more detail below, the ethical decisions that social scientists must make often depend on the particular research circumstances.

6.3. Informed consent

The lack of clarity about the public nature of Web 2.0 data also has an impact on obtaining informed consent. Bruckman (2004) notes that human subject research norms such as informed consent do not apply to material that is published, but as can be seen above, it is difficult to establish this consent online (Bruckman 2004: 103). Some interviewees did not think that implied

consent could be assumed just because information was online and thus in the 'public domain'. One interviewee sent messages to 4000 people who had been included as participants-by-proxy in her study (they had been named as contacts by her survey respondents) as she felt it was "…*at least important for people to be aware they were part of the research*" (AK). Again, it was suggested that the outcome of ethical decision-making depends on the specific context of the research. This is reflected in the AoIR guidelines, which note that informed consent considerations need to include: timing (not only if, but when, to ask); the medium of obtaining consent; and how the material is to be used (direct quotation or paraphrased? Attributed or anonymised?) (Ess 2002: 6).

Some interviewees cited examples from their own research of when they did not feel they needed to obtain informed consent. This included distinctions between technically public and technically private[15] social networking profiles studied in aggregate and a content analysis of blogs that did not reveal URLs or refer to specific cases. It can be seen that such decisions rest on the public / private distinction once again. There are also instances in which asking for informed consent may be considered unethical, as found by Hudson and Bruckman (2005) in their study of attitudes to privacy in a specific online setting. They experimented with different broadcast messages and studied the responses to a variety of consent conditions:

> *Our data indicated that chatroom participants kicked us out roughly two-thirds of the time when we attempted to obtain informed consent. Which is the greater harm – annoying two-thirds of the potential subjects or not obtaining consent? (Hudson and Bruckman 2005: 299)*

Interviewees mentioned that unsolicited contact with participants may be viewed as spam, for example, especially if large numbers are contacted at the same time.

6.4. Ownership of information

As well as needing to consider the Web 2.0 user's position as author, interviewees referred to the possibility of the host website objecting to data being used in this way, especially with regards to automated methods of collection. One noted the "…*potential copyright issues about content on websites which require permissions to be sought and which could be prohibitive for storing and using the data*" (YL). Allen, Burk and Ess (2008) point to the ethical questions of these data gathering methods, as they can place excess load on servers, potentially in contradiction to expected use, and violate the company in question's legal prerogatives. Interviewees who

---

[15] Profiles are technically public if anyone with access to the Internet can view them.

conducted these methods said that they dealt with this issue by following the 'robots.txt' protocol. This is an agreement where companies who do not want their websites to be accessed by webcrawlers put a file entitled robots.txt on the site; this issues an instruction to the automated programs that accessing the data is not allowed (Allen, Burk and Ess 2008: 19). For one interviewee, the absence of a robots.txt file on MySpace was the "…*deciding factor*" in using data from that particular social networking site (TE).

6.5. Publicly available data?

It could be argued that any Web 2.0 data that does not have restricted access is in the public domain, and consequently there should be no ethical concerns with collecting and analysing this data despite the fact that users may be posting personal content. No interviewee went as far as stating that all available content on the web should be used without considering the implications of its use, but there was a sense that perhaps researchers were over-cautious:

> *If people are making public postings about material it should be within our warrant as social scientists to be able to download that data and analyse it for its shape and form, as long as we try to minimise individual harm (RB)*

Another felt that users of social networking sites needed to "…*be aware that they sacrifice a certain level of privacy when they sign up, and should make sure they read the terms and conditions*" (YL). This does not necessarily mean that this information should be exploited, but that users should take greater care about the information they disclose, and that responsibility for privacy does not always fall on the researcher. Similarly, the technological functions of blogs typically allow restrictions to access. Interviewees stated that if these restrictions were in place, it was fairly clear if the blogs were private; and if not, that they were expected to be read. This consideration is reflected in the AoIR Ethical Guidelines, which suggest researchers examine the "acknowledged publicity of the venue" (Ess 2002: 6).

However, interviewees also recognised that users may not know how to use these functions, and that they are perhaps "…*not as aware as they should be*" (TC). One interviewee highlighted how the legal and ethical aspects of an issue might not be hand in hand; a Web 2.0 site "…*might be public space legally, but users' perceptions might be different*" (JW). Even if users are aware of the technological and legal privacy of a space, their expectations might be very different, which some interviewees thought should be respected. This is reflected in the online methods literature that proposes researchers should think about who is the intended audience of Internet content, even if it is located in a public space (Mann and Stewart 2000: 46). Empirical research has suggested that there is certainly a lack of awareness

about who can access the postings of Internet users. For example, Viegas (2005) discusses bloggers who lose their jobs because they write about their working lives and who do not have a good idea of (or think about) who their audience actually is; Hudson and Bruckman (2005) suggest that attitudes such as these could be due to perceived anonymity online and the invisibility of readers (Hudson and Bruckman 2005: 299).

It was suggested that "…*people are becoming more aware that data is being collected*" from Web 2.0 applications and sites (AK). Recent media stories have highlighted privacy concerns on Facebook in terms of identity theft, and earlier studies showed that there was some resistance to Internet spaces being used as research locations and tools. When discussing the attitudes of mailing list users towards researchers, Chens, Hall and Johns (2004) noted that animosity was not uncommon. The list moderators and users expressed a desire to protect the space from the 'research paparazzi' (Chens, Johns and Hall 2004: 160).

More recently, the founder of a company that set up a dedicated social networking site for young travellers and Gypsies stated that "The kids were very clear that they didn't want to be in a goldfish bowl where sociology students or voluntary workers could examine their lives" (Benjamin 2008). This could either be seen to contradict the idea that users are less concerned about privacy, or alternatively, that they set their own privacy boundaries as necessary. There was some variation between the interviewees over where the responsibility to protect personal information lies, but there was an agreement that context was important.

6.6. Importance of context

Questions of user awareness, and the public / private debate more generally, are indicative of the blurring of boundaries in Internet research and the extension of these in the new Web 2.0 environment. A number of interviewees described these issues as a 'grey area'; "…*you could argue the privacy issue either way*" (TC). Internet research literature notes the researcher's philosophical position on ethics will influence the decisions they make, and that there are international differences in standards. A utilitarian approach will weigh risk against benefits, and is more common in the US, whereas the deontological approach of European countries stresses the protection of individual rights as paramount (Markham, 2005: 813). For example, a utilitarian view might state that if no harm comes to Internet users from not obtaining informed consent then the research is ethically justifiable; deontological views on this would see the right to consent as being violated and the research is unethical even if the users would never discover that their content had been used in this way (Hudson and Bruckman 2005: 299).

As a result of these debates, interviewees favoured a context-driven approach to ethics reflecting the contingency of the issues. This is consistent with the ESRC's Research Ethics Framework (2005) that suggests general principles across disciplines but that their application is context-dependent; and also the AoIR ethical statement, which advises guidelines rather than recipes and notes that given the philosophical differences to the approaches described above, there is often more than one defensible course of action (Ess 2002: 4).

Some of the contingent factors identified by the interviewees when debating the issue of privacy included assessing the potential harm if the participants are identifiable; whether the data is to be presented in aggregate; and where and how the research will be disseminated: "…*you would have to be a lot more careful on intellectual property grounds if it was published*" (VM). A number of interviewees advised that one way of making a decision about the public / private issue was to consider:

> *...what is reasonably expected to be private and what is reasonably expected to be public; if you know you are reading a private discussion amongst friends then you should introduce yourself as a researcher… and be clear about what you're doing (WD)*

Public blogs, on the other hand, were felt to be more freely available data. The output from an early workshop on Internet research ethics proposed that researchers should consider the technological privacy of the environment and then the psychological understanding of the research participants (Frankel and Siang 1999: 11); the interviewees advocated a similar approach in response to the challenges of the Web 2.0 environment. Additional useful advice can be found in the ethics research literature. Bruckman (2002) recognises the situational nature of the public / private divide in Internet research and considers the right for authors to receive credit for their work needs to be balanced with the protection of vulnerable subjects. She therefore puts forward guidelines for the level of disguise of subjects' names when reporting such research on a "continuum of possibilities" (Bruckman 2002: 230). This framework was used in Hookway's (2008) study of morality in everyday life that utilises personal blogs as data. The protection of identity is privileged over credit to author, and blogs are not attributed to their source. Hookway (2008) argues the case for this in terms of 'reasonable' non-attribution under Australian copyright guidelines (Hookway 2008: 106)

It seems, therefore, that ethical guidelines that offer a series of 'questions to consider' rather than trying to advocate a blanket approach to Internet research ethics are more useful when conducting Web 2.0 research. As will be seen in the next section, a major concern for interviewees was to avoid the development of a rigid, bureaucratic ethical framework that did not address the needs of social scientists.

6.7. Government and commercial research

An additional point was made by the interviewees regarding the commercial and government research that is conducted using Web 2.0. It was felt that the practices of social scientists were ethical on the whole, even to the point of actually hindering research.

> *Anyone who is part of mainstream scholarship, there's enough checks and balances… I think the ethical issues are more pertinent when you talk about how companies and governments do research on people, also private security firms (EM).*

References were made to Facebook applications that collect users' information; targeted advertising by Google; and anti-terrorism surveillance by the UK government. It was felt that such activities were hidden from users of Web 2.0 but that it is likely that it is unwanted. This can be seen by the reaction to the Beacon program on Facebook, which published user activity from other websites on their profiles without their consent. Shortly after it was launched an online campaign and the resulting controversy resulted in a change to Facebook's approach and users instead had to opt-in to the program (Wikipedia Contributors 2008). Some interviewees mentioned that it was the role of social science to actually expose these privacy concerns, either by evaluating the ethical implications of this surveillance or by conducting research that demonstrates how easy it is to collect data utilising Web 2.0 technology.

7. Additional Ethical Issues

Interviewees were asked what they felt were the key ethical issues; many referred to the governance of research ethics in the social sciences as well as the difficulties of negotiating what is public and private in Web 2.0.

7.1. Ethical Scrutiny

One topic of conversation was existing ethical guidelines, which some interviewees did not find particularly helpful for Web 2.0 study. The AoIR recommendations were seen to perhaps be slightly out of date and a little too general; one interviewee found they were "…*not much practical help*" when faced with an ethical dilemma (TE). Similarly, the British Sociological Association's (BSA) Ethical Statement was described as a little "…*vague*" with reference to Internet research (VM). Researchers noted that ethical approaches were being "…*work[ed] out… that are appropriate for society, as people are making vast amounts of data publicly available*" (R1).

In addition, it was felt that research ethics committees in institutions were unlikely to have much knowledge about Web 2.0 research. One interviewee stated when her proposal for research in Second Life was submitted, they "…*didn't know what I was talking about*" (AK). Another interviewee who has regular discussions with ethics committees suggested that they are "…*not that aware or haven't seen it*" (WD). Given that this is a relatively new area of social science research, it is perhaps not surprising that there is a lack of knowledge about Web 2.0; however, the ethical dilemmas identified by interviewees mean that institutional guidance and support would be beneficial. One task that was recognised was educating the research community: "…*researchers need to ensure that their proposal applications are transparent and aim to inform ethics committees as much as possible*" (JG).

Despite some criticism of existing guidelines, and the need to educate ethics committees, there was a general sense amongst the interviewees that constraining, over-regulative procedures should be avoided. Some felt that movements to a bureaucratic ethical framework in the social sciences had already begun, and was preventing potentially valuable research. One factor in this trend was seen to be the 'medicalisation' of social science ethics, for example using medical research ethics committees and procedures as a template. The ESRC Research Ethics Framework notes that a social science specific framework is required, rather than using those from other fields (ESRC 2006: 1). Overall, there was resistance to any approaches which may restrict social scientists unnecessarily:

> *I think that it's a big issue – ethics as guidelines or procedural ethics… sometimes the two get confused. Procedural ethics can restrain ethical research. (R2)*

There was also evidence of differences in approaches to ethics between disciplines. For example, BSA guidelines simply note the various problems with conducting Internet research, and suggest "erring on the side of caution" (BSA 2002: 6). The Market Research Society (MRS) has more detailed documentation, which interpret the existing MRS rules with reference to their application in online research settings (MRS 2006). These differences were noted by some of the interviewees, and one gave an example of how this resulted in contrasting practices; an information sciences researcher who carried out collaborative research with a social psychologist had to "…*gain ethical clearance before we could do anything from a research ethics committee*", but this was not the case for a project with an economist, even though this involved human subjects (TC).

7.2. An Ethics of Web 2.0

Web 2.0, and Internet research more generally, was seen by the interviewees to alter the relationship between the researcher and the researched. This has some practical implications for ethical practice; it is much harder to debrief participants and assess harm, for example, than with face-to-face contact. There is also the more intangible issue of trust, which was seen by some interviewees to be the basis of an ethical relationship with research participants in the online environment, and that it was particularly important to be open:

> *The only currency you have in any internet interaction is trust – that's it… If you're in a space that actively promotes the use of a pseudonym, you need to make very clear you're not there to take the mick (AK)*

It was noted by one interviewee that given the public nature of Web 2.0, researchers needed to be honest about the uses they make of data and the fact that the participants might be traceable; this allows them to make informed decisions: "*Transparency is the key, not confidentiality any more*" (R1).

The changing nature of the research relationship might be read as evidence that a 'Web 2.0 ethics' is required, but most interviewees did not think that revised guidelines were necessary. Instead, it was felt that existing principles like minimising harm are still relevant, although the application may be need to be rethought in the online environment. It was important to avoid the constricting procedures outlined above:

It's the same as non-Web research, the need for a set of questions to be applied to particular context, not straight-jacketing researchers into inappropriate ethical restrictions. Guidelines should facilitate good ethical practice rather than make the process bureaucratic (JG).

The need for an ethical standpoint that is sensitive to the research context is therefore reiterated. Rather than having new ethical debates around privacy, for example, existing debates are "…*brought to another level*" (YL) in Web 2.0; the difficultly could be determining what is public and what is private, rather than changes to how social scientists deal with private data. It was also pointed out by one interviewee that the speed of online developments meant that "…*by time you have an ethics of Web 2.0, it would have changed anyway*" (DB).

What was a cause for concern for interviewees was that the ease of doing online research may mean that it can be carried out with little ethical scrutiny. One interviewee stated:

The ethical issues aren't different but anyone can collect data now. It lowers the bar to entry and they might not think about the issues (EM).

Another noted how "…*the ease and speed of conducting empirical research online*" meant that "…*ethically questionable*" research could be being carried out under the radar screens of ethics committees (WD). Madge (2007) notes the research community has to balance avoiding over-regulation with some sort of 'baseline' to prevent "shoddy cowboy" research practices (Madge 2007: 665).

7.3 Conclusion: Ethical issues in Web 2.0 research

As noted by Hine (2005) the growth of literature discussing computer mediated communication as a tool for social science research marks the question of research methods and new technologies as one of considerable interest but also concern; there is an air of innovation but also anxiety (Hine 2005: 5). The interviewees did not seem to show 'anxiety' as such with regards to Web 2.0 research but did highlight potential ethical challenges, particularly around changing notions of privacy. There were some differences of opinion over the level of responsibility that researchers had towards those who display their life publicly online, however all the interviewees suggested that context was important rather than blanket rules.

As noted by the AoIR guidelines, there is usually more than one ethically-defensible position on any given situation, and the blurring of boundaries that Web 2.0 is part of means that it becomes more complicated to establish clear-

cut standards. There was a strong resistance from the interviewees to bureaucratised ethical procedures, but some also referred to difficulties in making ethical decisions, or that they did not find existing guidelines particularly useful. As more social scientists utilise Web 2.0, it can be hoped that precedents will be established that may aid other researchers. Some researchers have argued for proactive work into these areas: Hudson and Bruckman (2005) noted that the idea of 'reasonable expectations' of privacy can be difficult, and consequently conducted a study into participant expectations in a particular setting. They argue the case for more empirical research like this (Hudson and Bruckman 2005: 301). It was also highlighted that existing ethical standards are built upon the historical dilemmas and the practical experiences of researchers:

> *I think there's a problem with a priori ethical development in social research anyway. Most ethical principles are based on case law, and most of the ethical principles that have been devised in the history of social research have been based on instances when people did something, there was an implication, and they thought it through (RB).*

As with the discussions about methodological issues, considering the ethics of Web 2.0 research sheds light on existing social science practices, such as the regulation of ethics, the need for research to be sensitive to context, or the overall feeling that ethical guidelines need to be useful but not restrictive to support valuable and rigorous social science research.

8. Other implications of Web 2.0

The developments termed 'Web 2.0' have additional applications as research tools beyond their utilisation for data collection. Although this was not the focus of the interviews, these uses were touched upon and outlined below.

A number of interviewees talked about using Web 2.0 to disseminate research in innovative ways. One interviewee mentioned an online peer-reviewed journal they were studying that features videos of scientific protocols rather than articles (LP)[16]. Another interviewee has "…*published something in Second Life that is not available elsewhere*" (TC). The potential for Web 2.0 to change current academic publishing practices has also been picked up in the literature. For example, Black (2008) argues that the current system of peer review in journals is susceptible to bias and delays the publishing of content; he suggests the use of wikis to move the process of peer review into the public arena.

The use of wikis is part of the trend towards openness and sharing that is linked to Web 2.0, and one interviewee noted "…*the move to making academic content more open, with peer review potentially expanding to the entire web community*", but reserved judgement on which was more worthy (AK). Peña-Lopez (2007) suggests that academics should set up a personal research portal, containing their profile, personal work, weblog and other interactive features. This enables academics to open up their research, foster international research networks and share knowledge in a low-cost, highly flexible space. In addition, a forthcoming book chapter by one interviewee argues the case for researchers to post their proposals online for public review. By utilising Web 2.0 interactive features in this way, researchers would be "…*more accountable for the ethics of their research*" (WD).

Beaulieu (2005) examines the use of blogs in documenting the research process, and how they can also be used to interact with the research population and authenticate the researcher (Beaulieu 2004: 151). The interviewees did not mention the use of blogs in terms of documenting a specific study, although many had a professional blog that discusses their current activities and areas of interest. Two interviewees (LP and EM) were studying the use of blogs by academics and referred to chemists blogging their lab books and archaeologists blogging their digs. Another interviewee (R2) has a PhD student who is keeping a research diary online and will incorporate this into their thesis. As well as being open and transparent, an online presence also helps researchers to expose their work to wider audiences:

---

[16] *The Journal of Visualized Experiments (JoVE)* http://www.jove.com

> *There's a self-promotion angle; if, as professionals, we're putting our profiles and publications on the web, we want as many people as possible to find them (LP).*

Beer and Burrows (2007) suggest that Web 2.0 can be used to teach the social sciences, for example students collaborating to produce wikis, running seminars online through social networking software, and having students use folksonomies to tag and collate content online (Beer and Burrows 2007: 4.7). None of the interviewees discussed using Web 2.0 in this way, although one teaching project set up its own social networking site and blog. Interestingly, the interviewee involved in the project has conducted some research on Facebook with undergraduate students and found they were "…*quite resistant to its use as a teaching space*" (JW) because they saw it as a primarily social environment.

Web 2.0 technologies also have the potential for collaborative working. In studying how scientific collaborations use new technologies, one interviewee was interested in "…*the implications for the science that can be done, the questions that can be asked and answered, the way in which science, social science, the humanities can be organised*" (EM). Some interviewees had experimented with this with mixed results; for example, using a wiki to jointly write a paper did not work, but "…*it worked better if one person wrote and others edited, rather than all trying to write at same time*" (TC). As trials of the ways to utilise these technologies continue, it can be expected the most useful applications for teaching and working will emerge.

As well as the ways in which Web 2.0 can be used as a research tool, there are also implications for the development of knowledge, and interviewees mentioned the debates about its democratising potential. These debates suggest the decline of top-down, expert knowledge, and ever-growing sources of freely accessible information created by networks of users through open participation and communal evaluation. This trend has been criticised; Andrew Keen talks about the 'Cult of the Amateur' and proposes that information is no longer reliable and is of lower quality (Keen 2007). Whether a positive or negative evaluation of these shifts is taken, it was suggested that they are part of "…*a huge cultural and intellectual change, that we're only just at the beginning of and don't really understand yet*" (RB).

9. Future research

> *I have a feeling that Web 2.0 research has not reached a breakthrough yet …people need to be aware of the technologies and competent in handling [them] as well in order to do more substantial breakthrough research (YL)*

The interviewees also discussed how Web 2.0 might be put to use in the future, and it was felt that there were a number of potential applications in the social sciences that have yet to be fully explored.

9.1. New methods and research practices

Some interviewees discussed their plans for using Web 2.0 in forthcoming studies. One proposed method was to set up a secure server "…*with our own wiki software… people can access that wiki and edit, and build up a definition… [and] keep their own blog*" (TC).   Another interviewee was debating the use of Facebook to conduct interviews. The potentials for using Web 2.0 were anticipated by a number of interviewees, especially in conjunction with geographically-located data. One was inspired by the ability to "…*connect very very different data sources together*" for example twitter[17], feeds on Google maps[18], linked to pictures on flickr[19] (TE). The integration of content with mobile technologies such as GPS signals, which can be automatically collected, was seen by one interviewee to be a "…*fundamentally new way of gathering data*" (R1).

It was also felt by a number of interviewees that future research could be enabled by working with computer scientists:

> *I think it's helpful to have collaborative research with someone who has the technical expertise to gather this data. I think it's an opportunity for social scientists who want to research topics that involve the internet to have a little bit of help from computer scientists or information scientists (MT).*

9.2. Problems

Although Web 2.0 offers social science potentially innovative ways of gathering data, problems were identified by the interviewees that may hinder future research. Firstly, the velocity of Internet innovations means it can be difficult to "…*keep up to speed as a research community with developments in this field*" (JG). There was also a feeling that there was a generation gap

---

[17] http://twitter.com/
[18] http://maps.google.com/
[19] http://www.flickr.com/photos/

between 'native' users of Web 2.0 and social researchers, so that social scientists do not understand or lag behind the communities: "…*things that seem innovative are ubiquitous and mundane to 13 and 14 year olds*" (RB). As younger people enter the research community, the problem may be combated. However, one might question if this is a perennial problem, as these 'natives' may not be familiar with whatever is beyond Web 2.0.

Finally, it was suggested that social science researchers lacked the technical skills and knowledge to really get to grips with Web 2.0: "*Social scientists should educate themselves better [and] try to understand the advantages and disadvantages*" (YL). This could be combated by collaborative research; one interviewee noted that if he was "...*conducting a large project, [he] would try to collaborate with someone computer orientated*" (VM). Collaborative research has its own problems, however:

> *I know everyone talks about interdisciplinarity, working on projects that are supposed to bring together computer scientists, economists, political scientists, but it's really hard… This is a major problem for the future, [social scientists] need to expand their toolkit (TE).*

It can, therefore, be argued that social scientists need to have a greater understanding of the technology itself if they wish to study these trends and collect Web 2.0 data.

10. Conclusion

The preceding discussion aimed to outline some key issues with regards to Web 2.0 research. Although its scope did not extend to drawing any definitive conclusions about the implications of this technology, some key findings, along with a number of questions and areas for further examination have emerged.

Web 2.0 enables social scientists to explore everyday life in new and innovative ways. This research can be completely unobtrusive; it can also be more interactive and participatory. However, the differences between online and offline research can be overstated as Internet use expands. As with any method, online research design needs to be 'fit for purpose', and Web 2.0 data should not be used for its own sake.

A key issue that impacts on ethical decision-making is deciding what is public and what is private in Web 2.0. As attitudes towards expectations of privacy can vary from website to website, social scientists need to be aware of the context in which their research is conducted. Further study of the utility of existing ethical guidelines for researchers would be of interest, given the need to balance sensitivity to context with practical help for social scientists to act ethically. Many of the interviewees referred to the medicalisation of social science ethical procedures, and were resistant to the over-bureaucratisation of research ethics. Finally, both research ethics committees and the wider research community could benefit from a greater knowledge and understanding of the online environment, and consideration should be given to how social scientists can work with other disciplines to explore Web 2.0 further.

Bibliography

ALLEN, G.N., Burk, D.L., and Ess, C. (2008) "Ethical Approaches to Robotic Data Gathering in Academic Research", *International Journal of Internet Research Ethics* 1(1): 9-36

ANDERSON, P. (2007) "What is Web 2.0? Ideas, technologies and implications for education", *JISC Technology and Standards Watch*, Feb 2007. Available from:
http://www.jisc.org.uk/publications/publications/twweb2.aspx

BASSETT, E.H. and O'Riordan, K. (2002) "Ethics of Internet Research: Contesting the human subjects research model", *Ethics and Information Technology*, 2: 233-247.

BEAULIEU, A. (2004) "Mediating Ethnography: Objectivity and the Making of Ethnographies of the Internet", *Social Epistemology* 18(2-3): 139-163.

BEER, D. (2008) "Making Friends with Jarvis Cocker: Music Culture in the Context of Web 2.0" *Cultural Sociology,* 2(2) 222-241.

BEER, D. and Burrows, R. (2007) "Sociology and, of and in Web 2.0: Some Initial Considerations", *Sociological Research Online*, 12(5). Available from:
http://www.socresonline.org.uk/12/5/17.html

BENJAMIN (2008) "Safe space online for the real 'chavs'", *The Guardian,*
30 July 2008.

BLACK, E.W. (2008) "Wikipedia and academic peer review: Wikipedia as a recognised medium for scholarly communication?" *Online Information Review* 32(1): 73-88.

BOYD, d. (2006) "The Significance of Social Software", in Burg, T.N and Schmidt, J. (eds) *BlogTalks Reloaded*. Norderstedt: Books on Demand. Available from: http://www.danah.org/papers/BlogTalkReloaded.pdf

BRUCKMAN, A. (2002) "Studying the Amateur Artist: A perspective on disguising data collected in human subjects research on the Internet", *Ethics and Information Technology* 4: 217-231.

BRUCKMAN, A. (2004) "Opportunities and Challenges in Methodology and Ethics" in Johns, M.D, Chen, S-L. S, Hall, G.J. (eds) *Online Social Research: Methods, Issues and Ethics* New York: Peter Lang.

BRUNS, A. (2008) *Blogs, Wikipedia, Second Life and Beyond: From production to produsage* New York: Peter Lang

BSA (2002) *Statement of Ethical Practice for the British Sociological Association.* Available from:
http://www.britsoc.co.uk/equality/Statement+Ethical+Practice.htm

CHEN, S-L S., Hall G. J., and Johns, M.D. (2004) Research Paparazzi in Cyberspace: The Voices of the Researched, in Johns, M.D, Chen, S-L. S., Hall, G.J. (eds) *Online Social Research: Methods, Issues and Ethics* New York: Peter Lang

CHESNEY, T., Chuah, S.-H., and Hoffman, R (2007) "Virtual World Experimentation: An Exploratory Study" Available from http://www.nottingham.ac.uk/economics/cedex/papers/2007-14.pdf

COOKE, M. and Buckley, N (2008) "Web 2.0, social networks and the future of market research" *International Journal of Market Research* 50(2) 267-292.

DENZIN, N. (1999) "Cybertalk and the Method of Instances", in Jones, S. (ed) *Doing Internet Research: Critical Issues and Methods for Examining the Net* Thousand Oaks: SAGE.

ESRC (2005) *Research Ethics Framework*, Available from:
http://www.esrcsocietytoday.ac.uk/ESRCInfoCentre/Images/ESRC_Re_Ethics_Frame_tcm6-11291.pdf

ESS, C. and the AoIR ethics working committee (2002) *Ethical decision-making and Internet Research: Recommendations from the AoIR Ethics Working Committee.* Available from: http://aoir.org/reports/ethics.pdf

EYSENBACH, G. and Till, J. (2001) "Ethical issues in qualitative research on internet communities", *British Medical Journal* 323(7321): 1103-1105.

FRANKEL, M. and Siang, S. (1999) "Ethical and Legal Aspects of Human Subjects Research on the Internet", *American Association for the Advancement of Science Workshop Report.* Available from:
http://www.aaas.org/spp/dspp/srfl/projects/intres.main.htm

GILES, J. (2007) "Life's a Game" *Nature* 4 January 2007, 445: 18-20.

HERRING, S.C. and Paolillo, J.C. (2006) "Gender and genre variation in weblogs" *Journal of Sociolinguistics* 10(4) 439-459.

HARDEY, M. (2007) "The City in the Age of Web 2.0: A new *synergistic* relationship between place and people" *Information, Communication and Society* 10(6): 867-884.

HEWSON, C., Yule, P., Laurent, D., and Vogel, C. (2003) *Internet Research Methods: A Practical Guide for the Social and Behavioural Sciences* London: SAGE.

HINE, C. (2005) "Virtual Methods and the Sociology of Cyber-Scientific Knowledge" in Hine, C (ed) *Virtual Methods: Issues in Social Research on the Internet* Oxford: Berg

HODKINSON, P. (2007) "Interactive online journals and individualization", *New Media and Society*, 9(4) 625-650.

HOGAN, B. (2008) "Analyzing Social Networks via the Internet (Pre-release version)" in Fielding, N, Lee, R and Blank, G (eds). *The Handbook of Online Research Methods*. Sage: Thousand Oaks, CA.   Available from http://individual.utoronto.ca/berniehogan/Hogan_SAGE_Internetworks_RC1.pdf.

HOOKWAY, N. (2008) "'Entering the blogosphere:' some strategies for using blogs in social research" *Qualitative Research* 8(1) 91-113.

HUDSON, J.M. and Bruckman, A. (2005) "Using Empirical Data to Reason about Internet Research Ethics", *Proceedings of the Ninth European Conference on Computer-Supported Cooperative Work.*

HUTTON, W. and Helsper, E.J. (2007) *The Internet in Britain: 2007* Oxford: Oxford Internet Institute.

IOCT (2007) "Dissemination", *Tags Network Narrative Project,* Available from: http://www.ioct.dmu.ac.uk/tnn/dissemination.htm

LANGE, P.G. (2007) "Publicly Private and Privately Public: Social Networking on YouTube" *Journal of Computer-Mediated Communication* 13(1) 361-380.

LIU, H. (2007) "Social Network Profiles as Taste Performances", *Journal of Computer-Mediated Communication* 13(1) 252-275.

MADGE, C. (2007) "Developing a geographers' agenda for online research ethics", *Progress in Human Geography* 31(5): 654-674.

MANN, C. and Stewart, F. (2000) *Internet Communication and Qualitative Research.* London: SAGE.

MARKHAM, A.N. (2005) "The Methods, Politics and Ethics of Representation In Online Ethnography", in Denzin, N.K. and Lincoln, Y.S. (eds) *The SAGE Handbook of Qualitative Research* Thousand Oaks: SAGE.

MRS (2006) *MRS Internet Research Guidelines* Available from: http://www.mrs.org.uk/standards/internet.htm

OECD (2007) *Participative Web and User-Created Content: Web 2.0, Wikis and Social Networking* OECD: Paris.

PEÑA-LOPEZ, I (2007) "The Personal Research Portal: Web 2.0 driven individual commitment with open access", Knowledge Management for Development Journal 3(1): 35-48. Available from: http://km4dev.org/journal

SRA (2003) *Social Research Association Ethical Guidelines.* Available from: http://www.the-sra.org.uk/ethical.htm

THELWALL, M. (2008) "Social networks, gender, and friending: An analysis of MySpace member profiles" *Journal of the American Society for Information Science and Technology* 59(8) 1321-1330

TURKLE, S. (1995) *Life on the Screen: Identity in the Age of the Internet* New York: Simon and Schuster.

VIEGAS, F.B. (2005) "Bloggers' expectations of privacy and accountability: An initial survey", *Journal of Computer-Mediated Communication*, 10(3): article 12.

WIKIPEDIA CONTRIBUTORS (2008) "Facebook Beacon," *Wikipedia, The Free Encyclopedia,* Available from: http://en.wikipedia.org/w/index.php?title=Facebook_Beacon&oldid=23194458 0 (accessed August 15, 2008).

Appendix 1: Participant List

David Beer (DB), Sociology, University of York.

Roger Burrows (RB), Sociology, University of York

Thomas Chesney (TC), Information Systems, Nottingham Business School.

William Dutton (WD), Oxford Internet Institute, University of Oxford.

Tobias Escher (TE), Oxford Internet Institute, University of Oxford.

Jenny Graham (JG), Qualitative Research Unit, National Centre for Social Research.

Aleks Krotoski (AK), Psychology, University of Surrey.

Yuwei Lin (YL), National Centre for e-Social Science, University of Manchester.

Eric Meyer (EM), Oxford Internet Institute, University of Oxford.

Vincent Miller (VM), Sociology, University of Kent.

Lucy Power (LP), Oxford Internet Institute, University of Oxford.

Mike Thelwall (MT), Information Science, University of Wolverhampton.

Jane Wellens (JW), Staff Development Centre, University of Leicester.

"Researcher 1" (R1)

"Researcher 2" (R2)

Appendix 2: Interview Topic Guide

- Professional background information
    - o Institution, discipline

- Defining Web 2.0
    - o Specific websites / types of websites?
    - o Principles?
    - o Differences to Web 1.0

- Ask to outline own research interests, and if these involve Web 2.0 in any way. If so
    - o How did they become interested in this area and when did they become interested?
    - o Ask about methods – what they did, how they did it, why they did it, what were difficulties, how did they assess the validity etc of data?
    - o Any ethical problems or deliberations – how did they solve them?
    - o Would they conduct this sort of research again? Why / why not?

- How can Web 2.0 be used by social scientists to collect data?
    - o NB – as site of research? As object of research?
    - o Think about own research, the current research of others, and potential future uses.
    - o Common issues across studies of blogs, social networks, folksonomies?
    - o Differences – in technologies utilised, in disciplines?
    - o Need for technological skills / understanding
    - o Example of particular study that has interested them

- Re-thinking evaluation criteria
    - o E.g. can we establish the 'authenticity' of data when research participants can easily mislead us?
    - o Validity? Reliability? Representativeness? Generalisability and sampling bias (if appropriate)?
    - o Positive effects, e.g. non-obtrusive; naturalistic settings.

- Are these methodological considerations any different from previous discussions of 'offline' or 'online' research methods?
    - o Something specific to Web 2.0 – are these issues unique?
    - o Continuations of debates in online research

- Ethical issues
    - o Public / private distinction?
    - o When do we need to obtain informed consent?
    - o When does research become intrusive?
    - o Confidentiality?

- o What is good practice?
- o Are these any different from 'offline' research?
- o Does Web 2.0 throw up any particular ethical problems that are different to previous debates in online research?

- Summary
    - o What do they think are the key issues?
    - o Where are they being discussed?
    - o What should I be focussing on?
    - o Anything else relevant that I haven't raised?

Appendix 3: Information and Consent Form

Date:

INFORMATION SHEET

Implications of the Digital Age: Web 2.0 and Social Science Research

I am currently working with the Social Science Collections and Research (SSCR) team at the British Library on a three-month doctoral placement co-funded by the ESRC.

The study investigates the use of Web 2.0 technologies as a social science research tool, as part of SSCR's activities in developing and running a hub of debate, networking and resources for the social sciences. The focus is on two key areas: the criteria for research evaluation (e.g. applying 'offline' standards of reliability and validity in an 'online' context) and the ethical issues associated with conducting research in this way. The study will not attempt to produce guidelines for Web 2.0 methodologies or an ethical checklist; rather, its intent is to map out current practice and thinking among social scientists in this emerging area of research.

An important part of the study is a series of interviews with researchers who have an interest in the use of these new technologies as a method of data collection, particularly those who have used Web 2.0 technology to interact with and study their research population, and those who consider research methodology and ethics in their work. I have approached you because your experience as social science researcher means that your views on such matters will be very valuable. I would be very grateful if you would agree to take part.

The interview will last between 45 minutes and an hour. During the interview I will ask you questions, and I would like to take notes and record the interview so that I will have fieldnotes to refer to and can transcribe the interview later.

You are free to withdraw from the study at any time. You may request your name and professional details to be kept confidential. It is envisaged that the output of the project will be a written report and a workshop. The workshop will be held at the British Library to disseminate the research findings to members of the social science research community, aiming to map out and publicise the benefits and potential issues of using Web 2.0 for social research. You will be kept up to date on the workshop arrangements if you wish.

If you have any queries about the study, please feel free to contact me at any time; my contact details can be found below

Helene Snee [Contact details provided]

CONSENT FORM

<u>Implications of the Digital Age: Web 2.0 and Social Science Research</u>

I have read and had explained to me by Helene Snee the information sheet relating to this project.

I have had explained to me the purposes of the project and what will be required of me, and any questions have been answered to my satisfaction. I agree to the arrangements described in the information sheet in so far as they relate to my participation.

I understand that my participation is entirely voluntary and that I may withdraw from the project or request to remain anonymous at any time before the publication and dissemination of the research findings.

I have received a copy of this consent form and of the accompanying information sheet.

Please tick as appropriate:

[ ]   I consent to my name and professional details to be referred to in the study

[ ]   I request that I remain anonymous and that my name and professional details are not referred to in the study.

Name:

Signed:……………………………………………………Date…………………

Appendix 4: Glossary

NB The definitions listed here are drawn from Web 2.0 glossaries and online resources available on the Internet; links to the source of the content are given for each term.

Blog / Weblog

"Blog is an abbreviated version of Weblog, which is a term used to describe Web sites that maintain an ongoing chronicle of information. A blog is a frequently updated, personal Web site featuring diary-type commentary and links to articles or other Web sites. Blogs range from the personal to the political and can focus on one narrow subject or a whole range of subjects."
http://www.web2fordev.net/glossary_of_web.html

Blogosphere

"The collective name for the millions of blogs on the Web, sometimes used by traditional media to gauge public reaction to an event."
http://www.saba.com/products/glossary/web_2dot0.htm

Folksonomy

"Folksonomy is used to categorize and retrieve Web pages, photographs, Web links and other Web content using open ended labels called tags. Typically, folksonomies are Internet-based, but their use may occur in other contexts as well. The process of folksonomic tagging is intended to make a body of information increasingly easier to search, discover, and navigate over time."
http://www.web2fordev.net/glossary_of_web.html

Online desktop

"A web desktop or webtop is a desktop environment embedded in a web browser or similar client application. A webtop integrates web applications, web services, client-server applications, application servers, and applications on the local client into a desktop environment using the desktop metaphor."
http://en.wikipedia.org/wiki/Online_desktop

Social bookmarking

"Social bookmarking is a Web-based service to share Internet bookmarks. The social bookmarking sites are a popular way to store, classify, share and search links through the practice of folksonomy techniques."
http://www.web2fordev.net/glossary_of_web.html

Social networking

"A method of connecting people with similar interests. Prime examples of this are LinkedIn and Plaxo in business environments, and Friendster, Facebook, and mySpace in social situations (although there has been some crossover). The power of social networking comes from the ability to search for and connect to people based on some aspect of how they describe themselves, and then gaining access to that person's friends and their interests."
http://www.saba.com/products/glossary/web_2dot0.htm

Social software

Social software is a type of software or Web service that allows people to communicate and collaborate while using the application.
http://www.web2fordev.net/glossary_of_web.html

Spam

"Spamming is the abuse of electronic messaging systems to indiscriminately send unsolicited bulk messages. While the most widely recognized form of spam is e-mail spam, the term is applied to similar abuses in other media".
http://en.wikipedia.org/wiki/Spam_(electronic)

Tags

"A tag is a (relevant) keyword or term associated with a piece of information (like a picture, article, Web site, or video clip), thus describing the item. Typically, an item will have more than one tag associated with it. Tags are chosen informally and personally by the author/creator or the consumer of the item - i.e. not as part of some formally defined classification scheme."
http://www.web2fordev.net/glossary_of_web.html

User-generated content

"User generated content… refers to various kinds of media content, publicly available, that are produced by end-users. The term entered mainstream usage during 2005 having arisen in web publishing and new media content production circles. Its use for a wide range of applications including problem processing, news, gossip and research reflects the expansion of media production through new technologies that are accessible and affordable to the general public. All digital media technologies are included, such as question-answer databases, digital video, blogging, podcasting, mobile phone photography and wikis. In addition to these technologies, user generated content may also employ a combination of open source, free software, and flexible licensing or related agreements to further reduce the barriers to collaboration, skill-building and discovery."

http://en.wikipedia.org/wiki/User_generated_content

Wiki

A wiki is a Web site that allows visitors to add, remove, edit and change content. It also allows for linking among any number of pages. This ease of interaction and operation makes a wiki an effective tool for mass collaborative authoring. The term wiki also can refer to the collaborative software itself (wiki engine) that facilitates the operation of such a site, or to certain specific wiki sites, like encyclopaedias such as Wikipedia.

http://www.web2fordev.net/glossary_of_web.html