MRS Market Research Standards Board

Online Data Collection and Privacy

Discussion Paper

Background

In recent consultations on the draft MRS *Guidelines for Online Research* and *Guidelines for Research with Children and Young People*, the MRS Market Research Standards Board (MRSB), received a number of comments and queries about the changing nature of online data collection and the extent to which personal data could be legally and ethically collected online without taking steps to obtain the consent of the individual concerned. In particular, the queries related to the use of social media and social networking services for this purpose.

MRSB is concerned that online research may, in some cases, be being conducted in a manner contrary to the long term interests of the research profession. MRSB believes that researchers in their work must have regard to the protection of respondents, the provision of good quality research services to clients, and to the reputation and integrity of the research profession.

MRSB has decided to issue this paper to share its current thinking on online data collection and to engage with researchers and other interested parties on this issue. The paper seeks to set out the legal and ethical parameters for online research not just in the UK but in the EU and beyond. MRSB understands that this is a global issue and invites comment and responses from researchers, clients and respondents to develop an agreed understanding of all relevant issues before drafting any new guidelines or rules for online data collection.

From the outset we should note that this is an issue that spans multiple jurisdictions, and it is not always clear which national law will apply to a particular data collection event. For this reason, the high level principles of both the data protection directive and the MRS Code of Conduct are of particular importance.

The online collection and processing of personal data in the EU is currently regulated by the Data Protection Directive 95/46/EC and the ePrivacy Directive 2002/58/EC. These have been transposed into UK law by the Data Protection Act 1998 and the Privacy and Electronic Communications Regulations 2003. For detailed information on the current legal situation please refer to *The Data*

Protection Act 1998 and Market Research: Guidance for MRS Members [1] and *Guidance on the Privacy and Electronic Communications Regulations 2003*[2] available from the MRS website.

Social Networking Services

A **social networking service** focuses on building online communities of people who share interests and/or activities, or who are interested in exploring the interests and activities of others. Most social networking services are web based and provide a variety of ways for users to interact, such as e-mail and instant messaging services. Well known services include Facebook, YouTube, Twitter, Flickr, Tumblr, Vimeo and, more recently, Google+.

Vast amounts of personal data are available online for anyone who wishes to look for them. In 2010 a single computer researcher compiled a list of more than 170 million Facebook users and the Web address of their profile page on the site and released it on a file-sharing site[3].

What is the problem?

Social networking data can appear to be a rich source of material ideally suited to researching habits, opinions behaviour and attributes of life in the early 21st Century. There are certain parallels with the origins of modern market and social research in the Mass Observation project started in the UK in 1937[4]. In one element of the project, a team of paid investigators went into a variety of public situations: meetings, religious occasions, sporting and leisure activities, in the street and at work, and recorded people's behaviour and conversation in as much detail as possible. While the archives of that project are a valuable historical resource, the methods employed would certainly not be acceptable today and would be likely to breach a number of provisions of data protection and privacy legislation.

In short, because information can be *found* it does not mean it should be *used* for research.

Legal and Ethical Framework

In this paper we look at the legal and ethical constraints within which researchers are obliged to work. MRSB has identified five key areas that it believes that researchers must consider when working in this area:

[1] http://www.mrs.org.uk/standards/dp.htm#guide
[2] http://www.mrs.org.uk/standards/mrs_guidelines.htm
[3] http://news.cnet.com/8301-27080_3-20012115-245.html#ixzz1CuJCMxMb
[4] http://www.massobs.org.uk/index.htm

1. Expectations of users and Terms of Use
2. The law of privacy
3. The law of copyright
4. The law of data protection
5. The principle of voluntary participation

Expectations of Users and Terms of Use

Each social networking site or service has their own terms of use. These documents set out the basis on which the service is offered to users, what happens to the data placed on the service, who owns it and who it will be shared with. It is the contract between the user and the service provider, and as such provides a statement of the reasonable expectations of users when they place data on the site.

Typically terms of use will provide conditions for the following:

- Access may be open or limited to certain age groups.
- Information supplied may be publicly available or limited to sub-networks or friends
- How other users can use data from profiles they have access to.

For example, the Facebook terms of use[5] set out clear conditions for the use of data from Facebook profiles. In short, they set a standard of informed consent:

> *If you collect information from users, you will: obtain their consent, make it clear you (and not Facebook) are the one collecting their information, and post a privacy policy explaining what information you collect and how you will use it.*

Further, terms of use may have special terms for particular jurisdictions. Again, Facebook stipulates that all disputes or claims for all users are subject to litigation in US Federal Court in Santa Clara, California. The only exception to this is for users in Germany whose use of the site is governed by German law.

The law of privacy

In many countries the law provides independent protection of personal information in addition to data protection. Where not expressly provided for in legislation or a national constitution, it has

[5] http://www.facebook.com/terms.php?ref=pf

often been found to be a fundamental or implied right of the citizen. For example in Ireland, the Supreme Court in *Kennedy and Arnold v. Ireland*[6] found that there was an implied right to privacy:

> *The right to privacy is one of the fundamental personal rights of the citizen which flow from the Christian and democratic nature of the State.... The nature of the right to privacy is such that it must ensure the dignity and freedom of the individual in a democratic society. This cannot be insured if his private communications, whether written or telephonic, are deliberately and unjustifiably interfered with.*

A similar implied right was found by the French Constitutional Court in 1995.[7]

Explicit rights to privacy are provided for international conventions. Article 8 of the European Convention of Human Rights provides:

1. *Everyone has the right to respect for his private and family life, his home and his correspondence.*

2. *There shall be no interference by a public authority with the exercise of this right except such as is in accordance with the law and is necessary in a democratic society in the interests of national security, public safety or the economic well-being of the country, for the prevention of disorder or crime, for the protection of health or morals, or for the protection of the rights and freedoms of others.*

This has been transposed into UK law by the Human Rights Act 2000.

Since the entry into force of the Lisbon Treaty, the Charter of Fundamental Rights of the European Union has had legal effect. Article 7 of the Charter provides that:

> *Everyone has the right to respect for his or her private and family life, home and communications.*

Historically, the development of the right to privacy has been in response to technological change. The first publication advocating privacy in the United States was the article by Samuel Warren and Louis Brandeis, *The Right to Privacy*[8]. It was written largely in response to the increase in

[6] [1987] IR 587
[7] Décision 94-352 du Conseil Constitutionnel du 18 Janvier 1995
[8] http://groups.csail.mit.edu/mac/classes/6.805/articles/privacy/Privacy_brand_warr2.html

newspapers and photographs made possible by printing technologies. In 2005, a group of Israeli researchers[9] proposed the following definition of right to privacy in the face of new technologies:

> *The right to privacy is our right to keep a domain around us, which includes all those things that are part of us, such as our body, home, thoughts, feelings, secrets and identity. The right to privacy gives us the ability to choose which parts in this domain can be accessed by others, and to control the extent, manner and timing of the use of those parts we choose to disclose.*

In the UK in June 2011, an early day motion was tabled in the House of Commons calling for the creation of an Internet Bill of Rights. This was in direct response to reports in the *Independent* newspaper of the creation of a database of internet users for the purposes of targeted advertising, which was characterised by Richard Halfon MP as "secret monitoring of internet users." As privacy becomes a political priority, particularly in the wake of disclosures in July 2011 by News International, researchers should expect that there will be increased requirements for transparency and consent online.

The law of copyright

Nearly all modern copyright laws are based on the Berne Convention for the Protection of Literary and Artistic Works. 168 countries are parties to the Berne Convention and almost all of its provisions are extended to other countries by membership of the World Trade Organisation and the Agreement on Trade-Related Aspects of Intellectual Property Rights (TRIPS). The relevant law in the UK is the Copyright Patents and Designs Act 1988[10].

The key aspect of copyright law is that copyright exists from moment of creation by the author and generally lasts for at least 50 years. It does not require registration (unlike a patent or a registered trademark). Generally copyright should be assumed to exist for the following online materials:

- Literary works – blog posts, profiles, updates
- Photographs
- Video Clips
- Sound recordings

[9] Yael Onn, et. al., *Privacy in the Digital Environment* , Haifa Center of Law & Technology, (2005) pp. 1-12
[10] http://www.legislation.gov.uk/ukpga/1988/48/contents

unless there is a clear indication to the contrary. Terms of Use may transfer the copyright to the provider of the social networking service. Copying of copyright works without permission of the author or the rights holder is a breach and may be the subject of legal action.

The law of data protection

Most information on a social networking service will be personal data, that is, data which relates to identifiable living natural persons. As such, the collection and use of information will engage the provisions of national data protection acts. In the EU and EEA these acts are based on Directive 95/46/EC. Under Article 7 of the Directive, personal data must be fairly and lawfully processed and six conditions are provided. The first condition is that of unambiguous consent of the data subject. It is this condition that researchers almost exclusively rely on to provide a legal basis for data collection from respondents.

The principle of voluntary participation

Research codes of conduct set down a clear standard of voluntary participation in research. The historical origin for this is the Nuremburg Code[11] and is part of market, social and opinion research's scientific heritage. The first point of Code is:

> ***The voluntary consent of the human subject is absolutely essential****. This means that the person involved should have legal capacity to give consent; should be so situated as to be able to exercise free power of choice, without the intervention of any element of force, fraud, deceit, duress, over-reaching, or other ulterior form of constraint or coercion; and should have sufficient knowledge and comprehension of the elements of the subject matter involved as to enable him to make an understanding and enlightened decision. This latter element requires that before the acceptance of an affirmative decision by the experimental subject there should be made known to him the nature, duration, and purpose of the experiment; the method and means by which it is to be conducted; all inconveniences and hazards reasonable to be expected; and the effects upon his health or person which may possibly come from his participation in the experiment. [emphasis added]*

This continues to this day in the principles of the research Codes themselves. The first principle of the MRS Code of Conduct states:

[11] The Nuremberg Code is a set of research ethics principles for human experimentation set as a result of the Subsequent Nuremberg Trials at the end of the Second World War

Researchers shall ensure that participation in their activities is based on voluntary informed consent.

The ICC/ESOMAR Code provides:

Respondents' cooperation is voluntary and must be based on adequate, and not misleading, information about the general purpose and nature of the project when their agreement to participate is being obtained and all such statements shall be honoured.

Indexing and Searching

All of the issues raised above are in some ways heightened in the online environment by the electronic trail that is left behind by users. Every comment, interaction, posting and profile will be replicated across multiple terminals, servers and accounts. Further this trail is automatically indexed in most cases by the service providers themselves and by search engines such as Google and Bing. Further the search engines may keep snapshots (caches) of webpages so that they can be found and searched even if the original webpage has been amended or deleted.

In the research context, this provides a wealth of information and opportunity to allow clients or other third parties, to identify respondents online. Quite frequently, a verbatim comment will present unique sequence of words that can be used to identify the source within seconds, obliterating any expectation of privacy or assurance of confidentiality offered by a researcher.

While data masking or cloaking may be offered as a privacy solution, from a methodological point of view it is rather unsatisfactory. This approach could distort results if the original wording was materially altered (and even a single word change can do this). Moreover, an unscrupulous researcher could claim they had used data masking when in fact they had just made up the quote themselves. There would be no way of verifying the authenticity of the data used in the research.

Applying the framework

MRSB believes that the most fundamental issue for social media research is the principle of voluntary participation. In every other context or medium, researchers are required to obtain the informed consent of respondents and to respect their right to withdraw from a research project at any stage. In social media research, there is a strong incentive for researchers to bypass the issue of voluntary participation and simply collect information that they can access or find.

Some researchers have expressed the view that information may be freely collected from "public internet areas" without the consent of the individuals to whom it relates. There is a problem in defining these areas. Is this an area online that can be generally seen by other internet users? This would appear to be problematic. Drawing parallels with the real world, there are many areas that can be seen by third parties, such as gardens, or rooms with windows facing a street that we would instinctively regard as private.

Where a researcher is being asked to collect information from or about individuals who have profiles on a social media service, they should ask the following questions:

- Do the respondents know that they are participating in a research project and do they have the right to withdraw?
- What have users been told about the why the data is being collected and the purposes for which is will be used?
- What expectations do they have about their data being accessed by third parties?
- Do the Terms of Use of the site permit the presence of a researcher on the network for the purpose of this research?
- If so, what information is the researcher required to disclose about themselves, their organisation and the purpose of the data collection?
- In addition what information needs to be disclosed and consent obtained to meet national data protection laws?
- Who owns the right to copy and reproduce photographs, videos and postings?

Virtual life is real life

The online environment is not a separate and distinct space with different rules. Virtual actions have real impacts in real life – just think of how information taken from social networking sites has been used (controversially) by employers or university admissions tutors in deciding on applications for employment or education. That is why as the real and virtual worlds converge further, MRSB believes that researchers must apply the same ethical standards to research online as they do in real life. That is:

- Obtaining the informed consent of all persons from or about whom data is collected;
- Clearly stating the purpose for which data is being collected (research);
- Clearing specifying what data is being collected and who will have access to it.

What do you think?

This paper sets out MRSB's current position on online data collection and privacy. In time their position may inform new guidelines or amendments or additions to the MRS Code of Conduct.

MRSB is now inviting you to contribute your thoughts, comments and ideas to the debate about ethical standards for online data collection and research.

Do you agree with the concerns set out by MRSB in this paper?

Do you agree with MRSB's interpretation of the relevant law?

Should the same ethical standards be applied to virtual and real life or can a persuasive case be made for applying different standards? If so please provide details.

MRSB welcomes responses from MRS members, Company Partners and all other interested parties. Please send your response by email to codeofconduct@mrs.org.uk by Friday 16 September 2011.

MRSB will publish its response to this consultation in the autumn of 2011.