World Research Codes and Guidelines

ESOMAR GUIDELINE ON SOCIAL MEDIA RESEARCH

CONTENTS

BACK TO CONTENTS

1. INTRODUCTION

The evolution of social media in recent years has changed the way that hundreds of millions of people share information about themselves around the world. The concept of consumers generating their own content on the internet has become ubiquitous. This has created new opportunities for researchers to observe, interact and gather information. Already many techniques have been developed to leverage social media such as community panels, crowd-sourcing, co-creation, netnography, blog mining and web scraping. Moreover it is likely that many more will evolve over the coming years as the internet continues to change.

The ICC/ESOMAR International Code on Market and Social Research requires that the same fundamental ethical and professional principles which govern face to face, mail and telephone research, are also applied to all types of online research (see Appendix 1). This document aims to provide guidance for the use of social media in market, social and opinion research.

It builds upon previously issued guidelines concerning Online Research and Passive Data Collection and supports ESOMAR's mission to ensure effective self-regulation to foster public confidence in our profession and industry. It aims to provide social media researchers with an awareness of the issues and guidance on how they can best apply the Code's fundamental principles of respect for consumers, trust, transparency and professionalism. This guideline also supports ESOMAR's consistent position to maintain a clear distinction between market, social and opinion research and marketing and PR activities.

Recognising that online research generally, and social media specifically, is continually evolving, this guideline conveys ethical and professional principles rather than being prescriptive about methodology.

Social media data often includes personally identifiable information. Most regulations in this area were developed before it was possible for one person to communicate with many on a publicly accessible online platform. Updates in privacy and data protection laws are still being developed and often lag changes in practises that have become generally accepted. This guideline is based on the principles underlying relevant laws and regulations presently in operation, especially with respect to data protection and intellectual property. However it also aims to propose pragmatic solutions that work within the spirit of existing laws and the ICC/ESOMAR Code, while aligning with currently accepted usage of online information around the world.

1.1 Scope

This guideline covers the collection of social media data for market, opinion or social research purposes. It recognizes that there are many different activities enabled by social media - of which market, opinion or social research is just one – and that these varied activities (including customer service and customer relations management) necessarily have different implications for consumers and those who make use of social media.

Researchers must not allow personal data they collect in a market research project to be used for any other purpose than market, social and opinion research. If it is intended to collect personalised social media data for other purposes, they must clearly differentiate this activity from their research activities and not misrepresent it as research. See section 2.1 for more on this issue.

This guideline shall be read in conjunction with the ICC/ESOMAR International Code on Market and Social Research and other ESOMAR guidelines available at www.esomar.org.

1.2 Definitions

Social media is defined as internet based platforms and technologies that permit users' interaction and/or facilitate the creation and exchange of user generated content. Whilst the scope is evolving, currently the most frequently used examples include:

- Online forums/discussions, communities, blogs, social networks (e.g. Facebook)
- Video/photo sharing (e.g. YouTube)
- Multi-person/group communication and/or collaboration platforms (e.g. Twitter).

Social media data refers to the information (photos, comments, etc.) that users generate or share while engaged in or with social media. It often includes personally identifiable data.

BACK TO CONTENTS

Social media research (SMR) covers all research where social media data is utilised either by itself or in conjunction with information from other sources. Examples of current social media research include:

- Monitoring or crawling social media platforms (from automated monitoring of brand sentiment through to ad-hoc desk research)
- Ethnographic research (from observing online social behaviour to participating and collecting primary data in various forms, including 'friending' users). This includes netnography
- Co-creational techniques used for research purposes
- Online communities that generate or deliver consumer opinions, reactions, feedback on a regular, formal or systematic basis.

Note that SMR excludes behavioural tracking research, even when conducted on social media websites.

Throughout this guideline a number of specific terms will be used whose meanings are as follows:

Data collection is the process of extracting data from social media data for analysis, and is sometimes referred to as scraping or crawling. This can be automated or done manually.

Masking is a technique whereby the original social media data such as comments, photos or videos is altered to a point that it cannot be traced back or attributed to the original user (using a search engine for example). See Section 2.4.

MROC (Market Research Online Community) is one of the more frequently used terms used to describe an online community created specifically for the purposes of market, social and opinion research. Others include DORC (Dedicated Online Research Community).

Personally identifiable information (referred to as personal data in EU legislation) is information that can be used to uniquely identify, contact, or locate a single person or combined with other sources to uniquely identify a single individual

ToU is the Terms of Use policy that a website or online service requires its users to accept.

Walled garden is an online service which requires users to register or apply for membership before being permitted to participate. A walled garden can only be accessed after the user has obtained a login and/or password, even if entry is automatic.

The following general terms are taken from the definitions in the ICC/ESOMAR Code:

Client is any individual or organisation that requests, commissions or subscribes to all or any part of a market research project.

Market research is the systematic gathering and interpretation of information about individuals or organisations using the statistical and analytical methods and techniques of the applied social sciences to gain insight or support decision making. The identity of respondents will not be revealed to the user of the information without explicit consent and no sales approach will be made to them as a direct result of their having provided information.

Researcher is any individual or organisation carrying out, or acting as a consultant on a market research project, including those working in client organisations.

User is any individual or organisation from which information is collected for the purposes of a market research project, whether they are aware of it or not, or is approached for interview. In the context of SMR, the user may also be referred to as an Author, Member or Poster. (Note: this term is equivalent to a Respondent or Participant in other market, opinion or social research modes.)

BACK TO CONTENTS

2. KEY PRINCIPLES

All the core fundamental principles of the ICC/ESOMAR Code (see Appendix 1) apply to social media research. The following section explains the implications of them for this context.

2.1 Distinguishing market, social and opinion research as the purpose

Researchers must not allow personal data they collect in a market research project to be used for any other purpose than market, social and opinion research. This principle can cause particular issues when investigating social media.

Under Article 7c of the ICC/ESOMAR Code, if quotes containing personally identifiable information are passed to the client, unless national provisions require stricter regulations, the following conditions must apply:

i) the respondent has explicitly expressed this wish and/or

ii) the respondent has given their explicit consent

iii) on the understanding that no commercial activity will be directed at them as a direct result of their having provided information.

In some cases those taking part in research-based communities could be exposed to sales and PR messages as part of the research process. This is permissible under the ICC/ESOMAR Code provided the purpose is for research (see section 3.3 for more guidance on how best to inform users about the purpose).

The ICC/ESOMAR Code requires researchers to be transparent in their dealings and not to misrepresent as market research any project which has another purpose. To aid clarity and protect the reputation of the researcher and of market research, the research services and the organisation or company carrying them out must be presented in such a way that they are clearly differentiated from any non-research activities. To ensure the public is not confused when social media data is being used by an organisation which is involved in both research and non-research activities, it is recommended that:

- the company's privacy policy and promotional literature must differentiate the different services that are being offered and separate market research from other activities;
- it must be easy for users and others to contact the researchers carrying out market research and those making enquiries must not be confused by having apparently to get in touch with a non-research organisation or deal with non-research staff to raise queries or complaints about market research activities;
- the introduction used when contacting a user must clearly define the purpose and they must not be left with the impression that the exercise has a research purpose if it does not.

These requirements do not prevent researchers from being involved in non-research activities providing the purpose of collecting personally identifiable data is not misrepresented. Nor do they in any way restrict the right of the organisation to promote the fact that it carries out both market research and other activities providing they are clearly differentiated and that they are conducted separately and in accordance with the relevant laws and local professional rules of conduct. More guidance in this area is given in the Distinguishing Market Research from Other Data Collection Activities.

2.2 Conforming to the law

Researchers must conform to all relevant national and international laws. There are three main legal aspects to consider here.

First, social media research must comply with national and international data privacy legislation and relevant requirements for notice, consent, accuracy, security and access when personally identifiable data is collected and stored. See section 2.3 for more on 'Consent'. There are also legal issues and measures (e.g. Safe Harbour requirements) relating to international transfer of personal data where data from which personal identifiers have not been removed is transmitted across national borders as well as considerations as to whether the country to which the data is transferred offers an adequate level of protection from a data protection perspective.

The second legal issue is that by accessing virtually all online services, researchers will be subject to the service owners' Terms of Use (ToU). Most ToU have intellectual property rights clauses that explicitly forbid the unauthorised copying of material. Many go further to bar all forms of social media data collection.

Subject to 'fair use' exceptions in certain countries, such ToU could prevent a researcher from even copying material to their computer for further analysis and forbid any form of selling on of that information to their clients, without permission.

For example, the following popular social media ToU can be located on the web at:

Facebook: www.facebook.com/terms.php

Twitter: twitter.com/tos

LinkedIn: www.linkedin.com/static?key=user_agreement

Researchers should therefore check what conditions apply to the content they use from social media and respect any requests for privacy (including robot.txt file requests, secure pages, etc). They must seek permission to scrape content from any source where this might breach the ToU and to abide by that service's ruling. Where permission is not granted, reading of such information and summarising the issues without copying anything is permissible, subject to the guidelines in section 3. Note, researchers should not engage in efforts to circumvent web sites' protections of the data they hold (eg IP spoofing, fictitious user ID's, etc).

Where researchers use third party aggregators for data collection services, the onus is on the researcher to check with their supplier that permissions have been obtained and the data has been sourced lawfully (see Appendix 2).

The third legal aspect concerns copyright. The ToU will usually also address copyright issues and the use of material on the website. Often the website owner will also be the copyright owner. However copyright laws are complex. They also vary across jurisdictions and researchers must acquaint themselves and observe the relevant laws on this topic.

2.3 Consent and notification

The ICC/ESOMAR Code states that users' co-operation must be based on adequate information about the purpose and nature of the project and their agreement to participation obtained. In addition, in some countries, existing data protection laws may also require users to be informed when personally identifiable data is collected.

Although it is potentially easy to obtain consent from members of market, social and opinion research communities, it poses more issues for other social media where users will generally not have been informed in advance or have consented to its use for research unless this is covered in the ToU.

As noted in the ESOMAR Online Research guideline, researchers must remain mindful of concerns about privacy and intrusion if sending an email requesting such consent. They must reduce any inconvenience such an email might cause to the recipient by clearly stating its purpose in the subject heading and keeping the total message as brief as possible.

If consent has not been obtained (directly or under the ToU) researchers must ensure that they report only depersonalised data from social media sources. If researchers are using automated data collection services, they are recommended to use filters and controls to remove personal identifiers such as user names, photos, links to the user's profile, etc. Where this is not possible or they are manually collecting data from websites, their analysis must only be with depersonalised data and no attempt should be made to identify people – see section 2.4 for a discussion on when identifiable quotes can be potentially used.

If depersonalized social media research data is passed on to another researcher or client, the researcher must have a contract with the recipient of the data which requires them not to attempt to use technical means to re-identify quotes or their posters and use such data for a non research purpose and to observe the ICC/ESOMAR Code and the provisions of this guideline.

When researchers, or their automated agents, interact directly with users within social media, they must also convey to them their purpose, role and how they will use any comments. They must also seek permission from users and the service owners or their representatives to conduct their work. Furthermore when taking part, researchers must ensure they do not misrepresent themselves as a genuine member of that social media space.

When researchers, or their automated agents, are working in social media spaces, they must provide an email address and a telephone number and/or mail address to facilitate contact and verification. For additional transparency, and to meet data collection notice requirements, researchers must publish a privacy policy on their website which explains how any personal data they collect is handled, offering appropriate measures to ensure that users can exercise their rights to require that their personal data are deleted or rectified, or that their personal data are not made available to others.

2.4 Protecting identifiable data

Social media platforms offer many opportunities to view personally identifiable data. Some people post information that overtly discloses their identity, are aware of this and have a diminished expectation of privacy. Others are not aware that the services they are using are open for others to collect data from or think that they have disguised their identity by using a pseudonym or username. However, online services are now available that make it possible in many cases to identify a "poster" from their username or comments and can link that to many other aspects of personally identifiable data including their address, phone number, likely income and socio demographic data.

Given this, data cannot always be 100% anonymised on the internet by merely removing the username and linked URL from the comment. Therefore if researchers wish to quote publicly made comments in reports or to pass these on to people not bound by the ICC/ESOMAR Code (or a contract linked to this), they must first check if the user's identity can be easily discoverable using online search services. If it can, they must make reasonable efforts to either seek permission from the user to quote them or mask the comment to such an extent that the identity of the user cannot be obtained.

Masking is a technique whereby raw data is so changed that it becomes very difficult for others to find the data online with a search service and thereby identify the person from whom it originates. It is a useful technique to ensure that the anonymity of people making comments is preserved where

1. the researcher has not sought their permission and
2. the comment would be easily traceable with a search service.

The degree of masking required will depend on the nature of the comment and its author. Masking can be applied in varying degrees ranging from just changing the odd word through altering key features of a comment to précising. It is the responsibility of the researcher to decide which is most appropriate. Factors to take into account include:

- If the topic being discussed is sensitive or personal
- If abusive language is used
- If it includes anything against the law
- If it includes anything embarrassing or is likely to impact career opportunities
- If it includes any personally identifiable information
- If it includes any data about others that is not already public.

In the case of public pictures or videos, consideration should be given to techniques such as pixelation of faces, where masking is required. Note, masking is unlikely to be sufficient in many B2B contexts or when researching very small groups, as identification is highly likely even if masked. When a researcher passes a masked comment on in a report to a client, they should clearly indicate if it has been masked.

If a researcher decides to seek users' permission to quote them, they must abide by relevant country laws, be sensitive to users' concerns about being observed and explain clearly and honestly the purpose of their work. The user should be given the opportunity to check the bona fides of the researcher, if they so wish, before deciding what to do.

Where researchers use services to enhance comments with demographic profiles (eg personal profiles), they should only use this information for research classification purposes. Since such services often provide other identifiable personal information (eg phone number, address or email), they could permit unintentional linkage of research data to personal data and this must not be used in analysis or passed on.

BACK TO CONTENTS

2.5 Ensuring no harm

Another key principle of the ICC/ESOMAR Code is that the rights of users as private individuals shall be respected and that they shall not be harmed or adversely affected as the direct result of participating in research.

The greatest risk in social media research relates to inadvertently revealing the identities of users, who did not realise they were participating in research and so would not expect to be identified. Again to ensure users are not harmed by research activities, the abiding principle must be one of caution, removing any personal identifiers in data as soon as possible taking into account necessary quality controls.

2.6 Children

Researchers must take special care when carrying out research among children and young people (see ESOMAR Guideline on Interviewing Children and Young People).

This is a particular issue with many social media platforms as their users can include children. Where data is likely to be from a child, researchers must take particular care in masking responses to ensure that the user cannot be identified or obtain permission from a parent or legal guardian to collect and use identifiable data (see ESOMAR Guideline for Online Research).

Where a market research online community (MROC) is being conducted with young people, permission must be obtained from a parent or legal guardian.

2.7 Reputation of the industry

Researchers must not do anything that might damage the reputation of market, social and opinion research.

Given its specific nature, working with social media requires additional care since any mistakes or misunderstandings can be spread virally within minutes across the network. Social media researchers must therefore be mindful of the core principles of the ICC/ESOMAR Code in the work they and their companies conduct and avoid activities and practices which could undermine public confidence in market, social and opinion research.

2.8 Reporting

The ICC/ESOMAR Code requires that projects are reported and documented accurately, transparently and objectively.

As social media research is a relatively new area, care needs to be taken by researchers to explain the impact that this sample source may have had on the results and their validity and reliability. This is necessary to ensure transparency for all involved and to educate users who may not be familiar with this type of research data rather than to devalue the information obtained.

3. SOME SPECIFIC RECOMMENDATIONS FOR CERTAIN SOCIAL MEDIA

3.1 Defining social media areas

There are three areas in which social, opinion and market research could be conducted. These are defined as follows:

- **Public social media:** This covers the majority of social media. It includes all places where access has been set by the website or the user to 'public' and entry is without any form of entry barrier. It can also include those where a username or password is required, but these are required for identification or site revenue reasons, rather than to protect the privacy of the data posted. Examples include public profile pages of social media networks; public micro-blogging posts; and many forums (including those where a username may be required, but is automatically granted, that is they are not moderated).
- **Private social media:** This covers areas where the user or the website do not want the data to be publically accessible. All require username identification for access, although this is not a distinguishing feature. These are sometimes referred to as 'walled gardens'. Examples include: 'private wall to wall' or individual communications on social media networks; protected posts on micro-blogging sites; or forums/groups areas where admittance is controlled by an administrator or moderator.

- **Market research social media:** This covers any online place specifically created for market, social and opinion research purposes where users have been informed of its function and the use to which their comments might be put. Typically (but not always) these are also private spaces. Examples include Market Research Online Communities (MROC's), certain blogs, online ethnographic and co-creational techniques which utilise social media platforms.

In addition to the guidelines in Section 2, researchers need to consider the following when using data derived from private and market research social media spaces.

3.2 Private social media areas issues

Researchers can only access these areas with the permission of the service operator or their agents. They should make it clear in their profile and preferably also in their avatar and/or username that they are a researcher, who they work for and their purpose. In interacting with members of a private area they should include a reference to their role, so that members are left in no doubt who they are talking to.

As a general rule, researchers should not copy or scrape content within private areas, even if they have permission of the site owner. If researchers do so, it should be made clear to all users that this is happening and they should provide individuals with a process to be excluded from such data collection.

Researchers must observe great sensitivity interacting with people in private spaces. When this needs to be done, they must follow the guidelines in section 2.3 on consent and notification and 2.4 on protecting identifiable data.

3.3 Market research social media areas issues

Market, social and opinion research spaces are normally private walled gardens where members must agree the purpose for which data is being collected and the terms and conditions for participation before they sign up to take part in the community, blog or co-creational project.

These terms must be simple, clearly worded and easy to understand. Members must be fully aware of:

- The purpose of the space - that it is for research, but in the case of a MROC, that they may be exposed to marketing information for research purposes only. This might include incentives for panel participation noting that using the client's products could be regarded as marketing in some countries
- That all data may be shared with the client – this is especially important since some members will actively share real names and photos
- How it could be used
- The rules for interacting (i.e. no cyber-bullying, defamatory comments etc)
- The site privacy policy including requirements listed in the Online Research guideline.

Unlike public internet areas, content can be copied and scraped and utilised for any research purpose, subject to members being fully aware of these applications. However the personal identity of those making comments must be protected. Some research communities offer clients opportunities to meet members and interact with them directly but this must be only with the consent of members. Clients must agree to abide by the ICC/ESOMAR Code, in particular that such interaction will be for research purposes only.

Where spaces are communal, members should be warned about contributing personally identifiable information. For example, members in a community should be given the option of using a pseudonym and uploading a photo of an avatar, a pet or inanimate object, rather than requiring that they use their real name and a photo of themselves. In addition on sign up, researchers should provide reassurance to community members that they will never ask for information that could create a risk of identity theft if lost, misused or disclosed to an unauthorised party, such as credit card numbers, social insurance/social security numbers or bank account details.

The role of moderators and clients should be unambiguously identified as such in all their interactions and communications with members. However care needs to be taken to protect them from being easily contacted outside the project, so it may be necessary to withhold full names and/or emails of such people and to utilise secure email systems within the relevant software.

BACK TO CONTENTS

Where MROC's are used to test products, advertising and/or communication of messages it is important that users are made aware of this. If users are taking part in a simulated sales test, it must be clear that they are helping in a research project and not in some form of direct marketing or sales exercise. No personal data collected during the course of an MROC may be used for any non-research purpose such as subsequent direct marketing or promotion to the individuals taking part.

4. FURTHER INFORMATION

Members who are unsure about the application of the Guideline in specific circumstances can seek advice by contacting the ESOMAR Professional Standards Committee, professional.standards@esomar.org

Project Team

- *Adam Phillips, Committee Chair, Managing Director, Real Research and Chair of ESOMAR Professional Standards and Legal Committees*
- *Ulf Andersen, MD of Synovate Scandinavia*
- *Pete Comley, Founder, Join the Dots (previously known as Virtual Surveys)*
- *Ed Keller, CEO, Keller Fay Group, co-founder of WOMMA Word-of-Mouth Marketing Association*
- *Peter Milla, Member of CASRO Task Force*
- *Annie Petitt, Chief Research Officer, Conversition*
- *Niels Schillewaert, Managing Partner of InSites Consulting*
- *Kristin Sharp, CEO of Ipsos Understanding UnLtd*
- *David Stark, VP, Compliance and Privacy, GfK, member of ESOMAR Professional Standards and Legal Committees*

BACK TO CONTENTS

APPENDIX 1: KEY FUNDAMENTALS OF THE ICC/ESOMAR CODE

1. Market researchers shall conform to all relevant national and international laws.
2. Market researchers shall behave ethically and shall not do anything which might damage the reputation of market research.
3. Market researchers shall take special care when carrying out research among children and young people.
4. Users' 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.
5. The rights of users as private individuals shall be respected by market researchers and they shall not be harmed or adversely affected as the direct result of cooperating in a market research project.
6. Market researchers shall never allow personal data they collect in a market research project to be used for any purpose other than market research.
7. Market researchers shall ensure that projects and activities are designed, carried out, reported and documented accurately, transparently and objectively.
8. Market researchers shall conform to the accepted principles of fair competition.

APPENDIX 2: CONTRACT/POLICY ADVICE WITH SUB-CONTRACTORS/THIRD PARTY SUPPLIERS OF SMR

When sub-contractors are used, ESOMAR recommends the researcher checks if they follow appropriate practices and procedures, especially with respect to observing any legal requirements and also about the privacy and protection of identifiable user data. Researchers should:

- Carry out due diligence when identifying and selecting sub-contractors
- Execute written Non-Disclosure agreements
- Execute written contracts that outline duties, obligations, and responsibilities of the sub-contractors that address all parts of the research process, especially privacy and data protection; parties involved and address non-disclosure requirements
- Engage in on-going oversight of sub-contractors and their activities and
- Ensure any data provided to them by sub-contractors is provided legally and in accordance to service ToU.

Policies and contracts relating to the research process and privacy are available from CASRO (http://www.casro.org) through the CASRO Privacy Protection Program (CASRO 3P). The CASRO 3P program has been designed to address the needs of various geographies, including the US and the EU.

Model contracts for transfer of personal data from the EU are available from the European Commission

http://ec.europa.eu/justice/policies/privacy/modelcontracts/index_en.htm

End of document