24 QUESTIONS

TO HELP BUYERS OF SOCIAL MEDIA RESEARCH

ESOMAR
WORLD RESEARCH

24 QUESTIONS
TO HELP BUYERS OF SOCIAL MEDIA RESEARCH

INTRODUCTION

Social media research is a relatively new methodology within the market research space. The components, which may include content analysis, text analysis, sentiment analysis, scaling, norms, sampling, and weighting, have been in use for many decades. It is the integration of these processes in concert with the relatively new opportunities provided by the internet and social media which defines this new technique.

These questions are intended to help users of social media research consider issues which influence whether a social media listening tool is fit for purpose in relation to a particular set of objectives whether qualitative, quantitative, or both. The context notes help provide an understanding of the reasons why the questions should be asked. They will help the user ensure that what they receive from a social media data provider is what they expected.

FOR QUICK REFERENCE THE AREAS
COVERED IN THIS GUIDE ARE:

COMPANY PROFILE

1 What is the expertise and history of the company with both market research and social media research?

This will help you form an opinion about the relevant experience of the provider and whether it focuses on market research, website analytics, information technology, or some other aspect within social media. How long has the company provided market research services? How long has the company provided social media research services?

2 What specific services are offered?

Social media research companies offer a selection of outputs including portal or dashboard access, data downloads, consulting, and/or self-serve/full-service products including unlimited types of reports and presentations. Describe each of the components the company provides.

3 What is the main purpose of the system or service?

It is useful to know if the system was designed for key word monitoring, customer service, market research, or some other purpose. The main consideration for selecting a tool is the business objective so the service selected must suit the purpose.

DATA SOURCES

4 Does the company collect their own data, rely solely on a third party supplier, or use a combination of the two?

Many social media analysis companies employ third party suppliers that specialise in the collection of social media data. Some of these suppliers may have special permissions with website owners to collect data and/or collect extra data. This may impact what data a social media analysis company has access to.

5 From which and how many websites does the company collect data including websites reflecting international or specific markets/languages?

It is helpful to have a wide range of data sources to ensure broad coverage of different groups of people. Does the company include data from major websites, e.g., Facebook, Twitter, Flickr, YouTube, Blogger? Does the company also include data from thousands or millions of other websites? Can the company add new websites that were not previously collected?

6 Does the company provide historical data from none, some, or all website sources?

You may want to put the data into a time perspective and must be aware of those websites that do not provide historical data. As such, some data are only available if a company has been collecting and storing data from the source themselves. Also, given the reach and growth of the internet, what is the company's view on how far back in time the data should go to provide a valid time perspective?

DATA MANAGEMENT

7 If used, describe the sentiment analysis system. Is it fully automated, fully manual, or a combination of the two? Is it based on a dictionary, Natural Language Processing, or other system? Is the coding categorical, continuous, or some other form?

Remember that all systems, whether automated or manual, cannot code sentiment 100% perfectly. Are research clients able to report errors and to correct them? Are research clients able to test the validity themselves?

8 What is the process for categorising data into specific content areas (groupings)?

Users may want to group data by topic areas, psychographics, or geography depending on their research objectives. The processes used for grouping can impact the interpretation and validity of the data. Is the process manual or automated? Can clients create their own groupings? Are clients able to report errors and have them corrected?

9 Does the company provide all data or a subset of the data?

Extremely popular topics may generate a dataset of millions of records, making it difficult to collect, store, and process every record. Describe any strategies the company uses for reducing the dataset to a smaller size.

10 If used, describe the system's sampling processes. Can the client specify the sampling process?

Sampling is the process of determining which websites or pieces of data out of those available will be included in a dataset and is determined by the research objective. Can clients specify which websites or portions of data, e.g., topic areas or keywords, they wish to receive data from?

11 If used, describe the system's weighting processes. Can the client specify the weighting matrix?

Weighting is the process of determining how much each website or type of data will contribute to the analysis and is determined by the research objective (e.g., a marketing campaign for a specific activity on Twitter, general internet, or specific type of person). Weighting can also help ensure that sections of the data providing a disproportionate share of the total data are weighted down to more natural levels (e.g., a website that is used by 10% of the population but contributes 50% of the data.)

12 Given that demographic information, including age, gender, income, education, geography, and more, is not widely available for social media data, how extensive is the demographic information? What validation processes are applied?

A systematic approach to estimating demographic features will help you to assess the data. What processes does the company use to provide this type of data, and how much of the data are populated? What proportion of the demographic data is actual versus inferred versus unpopulated? How are inferred methods validated?

DATA QUALITY AND VALIDATION

13 What is the method for identifying spam?

Biases might be created if spam is included in the data set. What types of automated and/or manual processes are in place for identifying spam or commercial messages that are misrepresented as consumer data? Are clients able to flag and identify such data and delete it themselves?

14 If used, how accurate are the sentiment scores and what method is used to validate them?

There are no standard validation processes which means the validity results from one company may not compare to those of another. A poor quality validation process may overestimate the reliability of the system. What is the specific validation process in terms of how messages are selected, how many are selected, and how often is the process conducted?

15 If used, how accurate are the content groupings and what method is used to validate them?

There are no standard validation processes for content groupings which means that measurements provided by different companies may not compare with each other. What is the specific validation process in terms of how messages are selected, how many are selected, and how often the process is conducted?

16 How is duplicate data defined and how, if at all, is it dealt with?

For instance, if one person says the same thing in two social media platforms or if two or more people say the same thing in the same social media platform (e.g., re-tweets, shares), it is useful to know if this is considered to be unique data. At what point would this be considered duplicate data?

POLICIES AND COMPLIANCE

17 What is the company's stance on engaging with social media users? When engagement does take place, what is the process for ensuring transparency about the researcher's role and presence?

Does the platform encourage interaction between the researcher and social network users? If researchers interact directly with users, they must identify themselves and explain their purpose to ensure they do not misrepresent themselves as a normal user of that social media space. What processes are in place to ensure social network users are treated respectfully (e.g., interactions do not take place without consent)? Is there a clear line between market research, marketing, and customer service activities?

18 How does the company approach user privacy or make content unidentifiable?

Some people post information that discloses their identity and have a diminished expectation of privacy while others are less aware that the services they are using are open for others to collect data from. What processes are in place to protect social media users in cases where their comments may be offensive or embarrassing to themselves or others? What guidelines does the company provide regarding the use of personal data such as name, username, and full links for messages? Given the risks to individuals and the industry associated with a failure to account for privacy, what guidelines are followed in this regard (e.g. **The ESOMAR Guideline on Social Media Research**)?

19 What processes are in place to obtain consent from social media users if identifiable data are shared?

Researchers should not report data which is identifiable. Identifiable comments should either be masked or used in reports with prior consent. If sending an email requesting consent to report identifiable data, researchers must remain mindful of concerns about privacy and intrusion. What guidelines are followed in this regard?

20 When it is known that a social media user is a child or a young person, how is their data treated?

Many social media platform users are children. Researchers must obtain permission from a parent or legal guardian to collect and report identifiable data or they must take special care that children cannot be identified. What processes are in place for cases when the company has identified that a contributor is a child or a young person? Is the data left untouched, deleted, or anonymised?

21 How is data masked/cloaked for reports?

If consent has not been obtained (directly or under the Terms of Use/Service) researchers must ensure that they report only depersonalised data. What techniques are employed to make it more difficult for research report users to find message contributors via an internet search? What recommendations does the company provide in terms of how to mask data for inclusion in reports? In handling this data, it is important that users are aware of and respect inadvertent disclosure of personally identifiable information. What guidelines are followed in this regard?

22 How can social media users request that their information be deleted from the research database?

What processes are available for contributors of messages to use should they wish to remove their tweets, blogs, status updates, video comments, etc., from the research platform?

23 How does the company comply with social network agreements and privacy statements, terms of service and robot exclusion standards etc.?

Legal conditions may apply to the social media content researchers use and researchers must respect any requests for privacy (including robot.txt file requests, secure pages, etc). Companies should describe their processes to ensure that they and their third party providers are compliant with such requirements.

24 What technological and manual processes are in place to ensure the protection of data?

Some sensitive and confidential information may be collected and stored that needs to be properly secured. Are servers maintained in secure locations? Is identifiable information retained in safe places? Are technical and organisational controls in place to limit access to the information on a strict need-to-know basis and are data retention policies in place such that personal information is destroyed once the purposes for which it was collected have been fulfilled?

GUIDANCE ON PROFESSIONAL STANDARDS

Maintaining consumer trust is integral to effective market, social and opinion research. ESOMAR through its codes and guidelines promotes the highest ethical and professional standards for researchers around the world. Providers of social media research are expected to follow all relevant ESOMAR codes and guidelines.

The ICC/ESOMAR Code on Market and Social Research, which was developed jointly with the International Chamber of Commerce, sets out global fundamentals for self-regulation for researchers. It has been undersigned by all ESOMAR members and adopted or endorsed by more than 60 national market research associations worldwide.

The ESOMAR Guideline on Social Media Research is of particular relevance to researchers using social media data and should be read in conjunction with these questions for more explanation of the legal and professional responsibilities of researchers who are collecting and analyzing social media data.

In addition, ESOMAR has issued the following guidelines to provide more detailed advice on how to address the legal, ethical, and practical considerations of conducting specific areas of research.

Guideline on Research via mobile phone
Guideline for Online research
Guideline on Distinguishing market research from other data collection activities
Guideline on Passive data collection, observation and recording
Guideline on Interviewing children and young people
Guideline on Customer satisfaction studies
Guideline on Mystery shopping
Guideline on How to commission research
ESOMAR/WAPOR guide to opinion polls

PROJECT TEAM

Annie Pettit, Vice President, Conversition and Editor of the text
Manila Austin, Vice President Research, Communispace
Pete Cape, Global Knowledge Director, SSI
Mike Cooke, Director: Global Panel Management, GfK
Jeffrey Henning, Chief Marketing Officer, Affinnova, Inc.
Adam Phillips, Chair, ESOMAR Professional Standards and Legal Committees

ESOMAR

WORLD RESEARCH

ESOMAR is the essential organisation
for encouraging, advancing and elevating
market research worldwide.