
**Digital analytics and web analyses
for purposes of market, opinion and
social research — Vocabulary and
service requirements**

*Analytique numérique et analyses web pour les besoins d'études de
marché, études sociales et d'opinion — Vocabulaire et exigences de
service*

STANDARDSISO.COM : Click to view the full PDF of ISO 19731:2017



STANDARDSISO.COM : Click to view the full PDF of ISO 19731:2017



COPYRIGHT PROTECTED DOCUMENT

© ISO 2017, Published in Switzerland

All rights reserved. Unless otherwise specified, no part of this publication may be reproduced or utilized otherwise in any form or by any means, electronic or mechanical, including photocopying, or posting on the internet or an intranet, without prior written permission. Permission can be requested from either ISO at the address below or ISO's member body in the country of the requester.

ISO copyright office
Ch. de Blandonnet 8 • CP 401
CH-1214 Vernier, Geneva, Switzerland
Tel. +41 22 749 01 11
Fax +41 22 749 09 47
copyright@iso.org
www.iso.org

Contents

Page

Foreword	iv
Introduction	v
1 Scope	1
2 Normative references	1
3 Terms and definitions	1
4 Research project management requirements	6
4.1 Organization and responsibilities.....	6
4.1.1 Management of the research process.....	6
4.1.2 Project management responsibilities.....	6
4.1.3 Appointment of a research quality manager.....	7
4.2 Confidentiality of information.....	7
4.3 Documentation requirements.....	7
4.3.1 General.....	7
4.3.2 Control of documents (other than project-related documents).....	7
4.3.3 Control of research project documents.....	7
4.3.4 Control of records.....	7
4.4 Competence and training.....	7
4.5 Subcontracting/outsourcing.....	8
4.6 Reviewing the effectiveness of the research process requirements.....	8
4.6.1 Project management review.....	8
4.6.2 Problems and complaints management.....	9
4.6.3 Internal audits.....	9
5 Proposals and tenders	9
5.1 Proposals and tenders from service provider to clients.....	9
5.2 Other aspects to be established by service providers.....	10
5.2.1 General.....	10
5.2.2 Data cleaning and editing.....	10
5.2.3 Sentiment and/or text analysis.....	11
5.2.4 Website usage and measurement analytics.....	11
6 Execution of the project	12
6.1 General.....	12
6.2 Digital analytics and web analysis data collection.....	12
6.2.1 Data collection methodology.....	12
6.2.2 Validation of data collection process.....	12
6.2.3 Participant safeguards.....	13
6.2.4 Weighting.....	13
6.3 Protection of individuals.....	13
6.4 Device monitoring.....	13
6.5 Reporting.....	14
6.6 Data retention.....	15
6.7 Data security.....	15
6.8 Ownership and publication of results.....	15
Bibliography	17

Foreword

ISO (the International Organization for Standardization) is a worldwide federation of national standards bodies (ISO member bodies). The work of preparing International Standards is normally carried out through ISO technical committees. Each member body interested in a subject for which a technical committee has been established has the right to be represented on that committee. International organizations, governmental and non-governmental, in liaison with ISO, also take part in the work. ISO collaborates closely with the International Electrotechnical Commission (IEC) on all matters of electrotechnical standardization.

The procedures used to develop this document and those intended for its further maintenance are described in the ISO/IEC Directives, Part 1. In particular the different approval criteria needed for the different types of ISO documents should be noted. This document was drafted in accordance with the editorial rules of the ISO/IEC Directives, Part 2 (see www.iso.org/directives).

Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights. ISO shall not be held responsible for identifying any or all such patent rights. Details of any patent rights identified during the development of the document will be in the Introduction and/or on the ISO list of patent declarations received (see www.iso.org/patents).

Any trade name used in this document is information given for the convenience of users and does not constitute an endorsement.

For an explanation on the voluntary nature of standards, the meaning of ISO specific terms and expressions related to conformity assessment, as well as information about ISO's adherence to the World Trade Organization (WTO) principles in the Technical Barriers to Trade (TBT) see the following URL: www.iso.org/iso/foreword.html

This document was prepared by ISO/TC 225, *Market opinion and social research*.

Introduction

Analyses of digital behaviour and online digital statements by persons and companies have rapidly increased in importance. Examples are the measurement of the behaviour of website visitors, the measurement of behaviour by means of cookies, and the registration and measurement of statements and sentiments of users of social media.

This document provides insight into the working methods of service providers in the fields of digital analytics and web analyses research and, in this way, provides clients with transparency regarding the services they offer. This document is intended to supplement and be used in conjunction with ISO 20252.

Digital analytics and web analyses for the purpose of market, opinion and social research can be separated from the equivalent analyses carried out for non-research purposes. In both cases, the protection of privacy of the persons analysed is regulated by legal provisions that apply to the particular project and, furthermore, by the relevant professional codes of conduct and other ethical guidelines.

STANDARDSISO.COM : Click to view the full PDF of ISO 19731:2017

STANDARDSISO.COM : Click to view the full PDF of ISO 19731:2017

Digital analytics and web analyses for purposes of market, opinion and social research — Vocabulary and service requirements

1 Scope

This document specifies the terms and definitions, as well as the service requirements, for organizations and professionals that conduct digital analytics and web analyses for collecting, analysing and reporting of digital data for purposes of market, opinion and social research by various methods and techniques. It provides the criteria against which the quality of such services can be assessed and evaluated.

This document applies to digital analytics and web analyses conducted by service providers on their own initiative, commissioned by clients or conducted by clients themselves.

This document applies to digital and web analysis research activities such as:

- understanding the usage of websites via the use of cookies, page impressions and other means, navigation across sites, time spent by visitors and their actions;
- online metered panels, e.g. on-going measurement of web visitation via meters installed on panellists' desktop, mobile or tablet devices;
- tag-based solutions to measure online usage at universe level, which can be integrated with metered panel data to provide a hybrid measurement;
- social media analytics which collect, aggregate and analyse online comments, and user-generated content such as blogs, forums and comments on news sites or other sites.

NOTE Universe can also be known as population.

This document can be construed to cover all forms of digital data collection including from desktop computers, tablets, mobile devices and over the top (OTT) devices as well as internet of things (IoT) devices where applicable.

2 Normative references

The following documents are referred to in the text in such a way that some or all of their content constitutes requirements of this document. For dated references, only the edition cited applies. For undated references, the latest edition of the referenced document (including any amendments) applies.

ISO 20252, *Market, opinion and social research — Vocabulary and service requirement*

3 Terms and definitions

For the purposes of this document, the terms and definitions given in ISO 20252 and the following apply.

ISO and IEC maintain terminological databases for use in standardization at the following addresses:

- ISO Online browsing platform: available at <http://www.iso.org/obp>
- IEC Electropedia: available at <http://www.electropedia.org/>

3.1

ad impression

display of an advertisement on a device

**3.2
algorithm**

process or set of rules to be followed in calculations or other problem solving operations

**3.3
anonymization**

process of removing, obscuring, aggregating, or altering identifiers with the aim of preventing the identification of individuals to whom data originally related

**3.4
application programming interface
API**

set of definitions on the basis of which a computer programme can communicate with another programme or component, and which can also support access/data exchange internally or externally

**3.5
audit**

systematic, independent and documented process for obtaining evidence and evaluating it objectively to determine the extent to which *audit criteria* (3.6) are fulfilled

[SOURCE: ISO 20252:2012, 2.5]

**3.6
audit criteria**

set of policies, procedures or requirements used as a reference

[SOURCE: ISO 20252:2012, 2.6]

**3.7
bot**

autonomous software that operates as an agent for a user or a program or simulates a human activity

**3.8
client**

individual, organization, department or division, internal or external, that requests or commissions a research project

[SOURCE: ISO 20252:2012, 2.10]

**3.9
code**

numeric or alpha character or combination of characters associated with each response category

[SOURCE: ISO 20252:2012, 2.11]

**3.10
code frame**

list of categories with associated *codes* (3.9) for classifying observations or statements

[SOURCE: ISO 20252:2012, 2.12, modified]

**3.11
consent**

freely given agreement based on adequate information obtained prior to the collection of *participant* (3.27) data

Note 1 to entry: When conducting research with children, additional consent requirements apply, e.g. permission from a responsible adult.

3.12**cookie**

small piece of information (i.e. programme code) that is stored on a browser for the purpose of identifying that browser during audience activity and between visits or sessions

3.13**dashboard**

software application with which a number of mini-applications can be reviewed or managed and reported

EXAMPLE Offer data exports, allow API access.

3.14**data collection**

process for gathering information by different means

Note 1 to entry: This includes activities such as web monitoring.

3.15**device ID**

distinctive number or *code* (3.9) associated with a specific digital device which can be personally identifiable

Note 1 to entry: Such devices can have multiple device IDs, each used for different purposes, for example to enable services such as WiFi or Bluetooth, or to uniquely identify specific devices operating on a mobile carrier network.

3.16**digital analytics**

analysing and reporting of electronic data for the purpose of measuring and understanding humans and their behaviour

3.17**digital fingerprint**

technology that deploys *algorithms* (3.2) that analyse a large number of technical characteristics and settings on devices to generate unique identifiers that can identify a specific computing device producing a machine ID, and which can be personally identifiable

Note 1 to entry: This ID can change over time depending on actions taken on the device.

Note 2 to entry: Also known as device, machine, or browser fingerprint.

3.18**dongle**

small piece of hardware, often a USB device, that can be connected to other electronic equipment, to enable additional services such as access to the internet

3.19**incentive**

gift, payment, or other considerations offered to potential *participants* (3.27) to increase cooperation

[SOURCE: ISO 20252:2012, 2.31]

3.20**internal audit**

periodic checks carried out by a company's own trained employees as to whether projects within the company have been carried out in accordance with the described procedures

3.21**internet of things****IoT**

infrastructure of interconnected objects, people, systems, and information resources together with intelligent services to allow them to process information of the physical and the virtual world and react

3.22

machine learning

technology of getting computers to act without being explicitly programmed

EXAMPLE Speech recognition, effective web search.

3.23

malware

variety of forms of hostile or intrusive software, including computer viruses, worms, Trojan horses, ransomware, spyware, adware, scareware and other malicious programs

3.24

mobile device

portable electronic device which can collect data that can be uploaded to a third party either immediately or upon syncing with appropriate software

Note 1 to entry: Mobile devices might include smartphones, smart watches, fitness or health tracking devices, tablets, and geo-location devices.

3.25

nearfield communication

NFC

short-range wireless technology that enables communication between devices over a short distance

3.26

over the top

OTT

digital services delivered over the internet rather than via a service provider's own dedicated, managed internet protocol television (IPTV) network

Note 1 to entry: Examples of digital services include Amazon Fire, Apple TV, Chromecast, Roku, Netflix or Slingbox¹⁾.

3.27

participant

person from whom or about whom data are collected

3.28

quality manager

person who is responsible for the formulation of quality requirements within an organization and for the procedures and measurements aimed at the implementation of quality controls

3.29

robot instruction file

file that defines how a search engine should interact with the pages and files of a web site and is often used to define where automated systems are not allowed to go

3.30

sample

subset of the *target population* (3.36) from which data are to be collected

[SOURCE: ISO 20252:2012, 2.58]

3.31

sentiment

mood associated with, for example, a sound, image, or statement, usually on a continuum from positive to neutral to negative

1) This information is given for the convenience of users of this document and does not constitute an endorsement by ISO of these products.

3.32**service provider**

organization that conducts research projects or parts of research projects in market, opinion and social research

[SOURCE: ISO 20252:2012, 2.55, modified]

3.33**social media**

online technologies and practices that people use to share opinions, insights, experiences and perspectives with each other, transforming traditional one-to-many interactions into many-to-many interactions

[SOURCE: ISO 20252:2012, 2.63]

3.34**spyware**

devices or software that capture a *participant's* (3.27) behaviour without obtaining *consent* (3.11)

3.35**subcontracting**

outsourcing

passing responsibility for executing an element of the project to a third-party service provider

[SOURCE: ISO 20252:2012, 2.64, modified]

3.36**target population**

population of interest in the research project to which inferences are to be made

[SOURCE: ISO 20252:2012, 2.66]

3.37**text analysis**

method used to describe the characteristics of a message

Note 1 to entry: This may also be referred to as content analysis.

3.38**unique visitor**

unique user

inferred measure that signifies a distinct, unduplicated, individual requesting page from the website during a given period, regardless of how often they visit

3.39**validation**

process of using objective evidence to confirm that the requirements which define an intended use, application or outcome have been met

Note 1 to entry: Validation can be carried out under realistic use conditions or within a simulated use environment.

3.40**web analysis**

analysing and reporting of behaviour, statements and *sentiments* (3.31) from users/*participants* (3.27) of online platforms

3.41

web beacon

pixel tracker

web bug

piece of *code* (3.9), often a 1 × 1 pixel, on a website used to track website activity

Note 1 to entry: This differs from beacon, which is a low-powered transmitter which notifies nearby devices of its presence, and can be used to trigger an action.

3.42

weighting

calculation process in which different units or subgroups are recalculated by assigning numerical values, as necessary, to correct and/or improve the representativeness of *sample* (3.30) estimates

[SOURCE: ISO 20252:2012, 2.69]

4 Research project management requirements

4.1 Organization and responsibilities

4.1.1 Management of the research process

Service providers in the field of digital analytics and web analyses shall use for research projects a research process which covers all the requirements of this document.

The research process shall be documented. It can differ from one organization to another due to the following:

- the size of the organization and type of research activities;
- complexity of and risks associated with research projects and their interactions;
- the competence of personnel.

Procedures, instructions and methods required for completing the different tasks in accordance with the requirements of this document shall be documented, implemented, monitored, maintained and auditable.

Documents may address the requirements for one or more procedures. A requirement for a documented procedure may be covered by more than one document. The documentation can be in any form or type of media.

4.1.2 Project management responsibilities

The service provider, as part of its project management, shall take responsibility for:

- committing to quality of client service (including a statement of quality policy) appropriate to the purpose of research projects;
- documenting the organizational structure of the service provider, including the responsibilities of the people involved in the delivery of research services;
- reviewing and improving the research project;
- ensuring the provision of adequate and appropriate resources and information for the process, including the appointment of a research quality manager;
- ensuring that everyone involved in the provision of the research service is familiar with the applicable national and international ethical and professional codes, requirements of relevant legislation and documented procedures and methods which specifically affect their work.

4.1.3 Appointment of a research quality manager

A research quality manager shall be appointed who has enough authority to be responsible for the administration of the whole research process management and who is responsible for organizing internal audits in order to ensure that this document is applied.

NOTE In some circumstances, the research quality manager can be a part-time role, and in other circumstances, it can be more effective to appoint more than one research quality manager (i.e. share the role).

4.2 Confidentiality of information

All information supplied to the service provider by the client in order to conduct a research project shall be treated in the strictest confidence. It shall only be used in this context and shall not be made available to third parties without the client's authorization. Confidential information shall be stored securely. See also [6.7](#).

The research findings relating to a specific client that are obtained by a service provider shall not be used for other clients without authorization of the original client and shall be treated in the strictest confidence.

Participants' data, as well as their identity, shall be treated in the strictest confidence. All assurances given to participants, either directly or indirectly, shall be fulfilled.

EXAMPLE Assurances made by social media suppliers to their users.

4.3 Documentation requirements

4.3.1 General

Research records and documents (which can be paper based and/or electronic) shall be established and maintained to provide evidence of traceability. Service providers shall specify to what extent a research project deals with specific privacy legislation.

All electronic files shall be checked for malware by up-to-date malware detection software.

4.3.2 Control of documents (other than project-related documents)

Documents required as part of the research process shall be controlled. Documents shall be subject to a version control procedure that allows the current version to be clearly identified.

4.3.3 Control of research project documents

Each research project shall have a project file (which can be in electronic format) that contains or references the location of a project specification showing the basic requirements of the project.

In addition to the requirements of [4.3.2](#) and unique project identification, documents shall be uniquely identified to allow traceability and to ensure they can be located.

4.3.4 Control of records

Records established to provide evidence of conformity to this document and project requirements shall be kept secure from unauthorized access, useable and retained for defined periods (e.g. back-up).

4.4 Competence and training

The service provider shall:

- detail and document the required skills and competencies for each position in the research process;

- provide appropriate training, including easy access to suitable training materials, for all personnel involved in the collection and processing of research data;
- ensure that the training adequately covers all technological, ethical and legal considerations for the processing and collecting of personal data;
- ensure that all staff and subcontractors involved in the processing and collecting of personal data are subject to appropriate management supervision.

4.5 Subcontracting/outsourcing

The service provider shall remain entirely responsible for all services carried out in connection with a research project, including any part of the work relating to this document which is subcontracted and/or outsourced, except where the choice of the subcontractor is beyond the control of the service provider.

The service provider shall define procedures to select subcontractors in the field of research, establish contractual relations and control the quality of the service provided.

The service provider shall ensure that subcontractors understand the requirements of this document in relation to the processes subcontracted to them by providing a written specification of such requirements.

NOTE The level of detail of the specification considered appropriate can depend on whether the subcontractor has carried out similar work before for the service provider, and whether or not the subcontractor is able to provide evidence of conformity to this document, where necessary.

The service provider shall obtain confirmation from the subcontractor that research-related work is undertaken in conformity with this document. Appropriate confirmation shall include one or more of the following:

- a written contractual agreement to work to the requirements of this document;
- assessment and certification to this document;
- documentary evidence from the subcontractor of appropriate checking and validation of requirements;
- appropriate checking and validation by the service provider.

The service provider shall maintain relevant records, including records of unsatisfactory service received from a research subcontractor and actions taken. These records shall be traceable to the project.

4.6 Reviewing the effectiveness of the research process requirements

4.6.1 Project management review

The service provider shall monitor and manage the agreed research process and its management of the research project by the following means:

- monitoring and reviewing the functioning of the research process management against policy (e.g. internal audits);
- reviewing projects and outcomes against agreed client specifications;
- monitoring and reviewing client satisfaction at regular intervals, at least every year, in order to improve the quality of the service.

4.6.2 Problems and complaints management

When problems and complaints are identified (either in the research process or its outcome), they shall be rectified and steps shall be taken to prevent recurrence.

Records shall be maintained, including details of complaints, problems, causes, actions and solutions.

4.6.3 Internal audits

The service provider shall carry out, at pre-planned intervals, internal audits to establish conformity with the following:

- the research process requirements;
- project specifications agreed with clients, including documented amendments throughout the project.

The system of auditing shall include an audit schedule, scope and methodology. The research quality manager is responsible for organizing this system (see 4.1.3). All internal auditors shall be trained and, where the size of the organization permits, shall not audit their own work. Records of audits, follow up actions and outcomes shall be recorded and, where appropriate, referenced to client projects.

5 Proposals and tenders

5.1 Proposals and tenders from service provider to clients

Proposals and tenders shall be issued in written form, online or offline.

In the event of uncertainties in the briefing, the service provider shall contact the client regarding these.

Proposals and tenders shall clearly describe the responsibilities of the client and of the service provider with regard to the tools used and the implementation of the results.

In proposals and tenders, at minimum the following subjects shall be covered.

- The client's information requirements.

NOTE 1 It is possible that certain information needs can only be concretely defined during the execution of a project.

- The digital analytics or web analysis methods or techniques recommended by the service provider.
- The digital sources to be analysed, as well as the criteria that are to be employed in the selection of the sources. In so far as the service providers and subcontractors employ differing definitions with respect to the digital sources, the proposal or tender shall clarify which definitions are being used.
- A clear statement of the technical and sampling dimensions that are connected with the source to be analysed and which can determine the data quality.
- The exact period(s) of time which the digital analytics or web analyses will encompass.
- The scope, type and frequency of supply of the results.
- The manner in which data are to be stored by the party collecting the data, and the duration of storage. Both shall be agreed with the client.
- The ownership of the findings made available after conducting the project.
- The costs of the activities to be carried out, including a statement of any possible variable costs whose amount can only be determined after the execution of the analysis. The manner and time of invoicing shall be specified.

- Any individual steps which are to be or might be subcontracted to other external service providers. On request clients shall be told the identity of any subcontractor. If it transpires that, subsequent to the granting of the commission, certain services need to be purchased from subcontractors, without this having been specified in the proposal or tender, the service provider shall immediately notify the client of this.
- The relevant legislation and regulations as well as the applicable professional code of conduct.
- The declaration that in the execution of digital analytics and web analyses projects the service provider complies with this document.

NOTE 2 Certification is one way of providing evidence of conformity to this document.

The issuance of a proposal or tender requested by a potential client is an activity which shall be classified as 'confidential', even in the event that the proposal or tender does not lead to a commission.

5.2 Other aspects to be established by service providers

5.2.1 General

To facilitate transparency and understanding of the project and its capabilities and/or limitations, service providers shall prepare and have readily available for discussion (if requested) with clients, the information in [5.2.2](#), [5.2.3](#) and [5.2.4](#).

5.2.2 Data cleaning and editing

Service providers shall document and have readily available the data cleaning processes relevant to the research purpose.

- a) Depending on the objectives, data to be cleaned can include:
- homonyms (e.g. apple pie versus apple computer, target practice versus target store);
 - data sourced from clients and their competitors as opposed to that sourced from consumers (e.g. tweets by a client);
 - advertisements, coupons, promotional material, etc., through third parties (e.g. a shoe store advertising their own sale for a shoe brand);
 - spam (e.g. a blog comment might say "buy sneakers buy sneakers buy sneakers buy sneakers"), fake social media accounts;
 - data originating from news channels (e.g. a tweet linking to a magazine article).
- b) Where appropriate, service providers shall specify whether their processes will identify and remove invalid/non-human traffic such as:
- bots and spiders;
 - invalid browsers;
 - internal traffic;
 - incentivized browsing for the purpose of defrauding advertisers/publishers;
 - other sources of invalid traffic.
- c) Service providers shall prepare and have readily available for discussion (if requested) with clients the following information:
- details of the degree to which data cleaning takes place manually, using automated algorithms, or some combination of the two;

- explanation of the expected effects of the intended data cleaning strategy, degree and processes by which inappropriate records can be identified and cleaned, as well as who can do such tasks;
- criteria for removal of data from the active data set;
- criteria governing how to handle data with missing values filled in (e.g. “0” fill, “data neutralization”, inferring).

5.2.3 Sentiment and/or text analysis

Service providers shall prepare and have readily available for discussion (if requested) with clients the following information:

- whether the sentiment and/or text analysis is conducted manually (including coding assisted by software), automated using algorithms (whereby a human does not personally code every piece of data), or some combination (e.g. machine learning);
- an explanation of the type and method of sentiment analysis conducted;
- an explanation of the type and method of text analysis conducted;
- the types of data that will be analysed (e.g. short sentences, long sentences, non-sentences, emoticons, slang, profanity, sarcasm, acronyms, words with spelling errors);
- the frequency and process for evaluating the reliability, accuracy, and validity of the analysis (e.g. weekly or yearly, number and type of records validated, whether blind coding is used), including up-front validations;
- whether analysis is fine-tuned on a job-by-job basis;
- whether it is possible to identify possible coding errors and/or who will do the recoding of data;
- the languages included or excluded, multilingual messages, etc.
- the results of any quality measures with regards to sentiment or text analysis such as: Inter-coder reliability scores/indices (in case of manual coding) as well as manual re-coding of originally automated coding.

5.2.4 Website usage and measurement analytics

Service providers shall prepare and have readily available for discussion (if requested) with clients, the following information:

- the type of data that is collected (e.g. IP address, operating system, browser user agent, cookie UID, geo-location, data/time stamps for events recorded);
- whether data reflects unique cookies (unduplicated cookies that represent visits to internet content during a measurement period and can include multiple users on a machine using the same account), unique browsers (result from count of unique cookies after adjusting for cookie deletion), unique devices (result from count of unique browsers after multiple browser usage on an individual computer is accounted for) or unique visitors (unduplicated individual people visiting specific content);
- basis of reporting (e.g. daily, weekly, monthly), time zones as applicable and how visits/time spent are defined;
- where coding takes place, what rules are used to address e.g. abandonment, handling multiple tags, and improper tags, determinations of usage duration and support for such data adjustment;
- whether data collected is from the available universe or a projection from a sample, its composition, the level of demographic data that is available and if data matching/enrichment is offered;

- whether cookie sharing, deletion, or rejection is addressed and, if so, how;
- any limitations on data provided and/or ability to validate or audit, such as restrictions on data retention and/or accumulation periods.

6 Execution of the project

6.1 General

As soon as a client has accepted the proposal or tender, the service provider shall appoint a responsible project manager. This shall be communicated to the client since the project manager shall be the client's first point of contact. The project shall be conducted in accordance with the relevant terms and conditions of digital sources. Where changes in, or deviations from, standard procedures are expected, the service provider shall inform clients of the change and the anticipated result of that change in advance when possible or as soon as possible thereafter.

6.2 Digital analytics and web analysis data collection

6.2.1 Data collection methodology

The service provider shall document the details of the methodology used. This shall include:

- a) methodology for collecting online comments from sites and/or social media, for example:
 - the universe of websites and other social media forums included;
 - how the sites, individuals or authors within them, and their comments or posts were selected for inclusion in the research;
 - how the comments and/or posts were collected (e.g. from questionnaires or user-generated content);
 - where appropriate, a description of the search algorithm(s);
 - the number of sites, authors and comments, or posts included in the analysis;
- b) methodology for observation and/or measurement of behaviour by digital means:
 - the population of interest;
 - how individuals or units were selected and recruited for inclusion in the research;
 - definition of the units studied (e.g. devices, sites, individuals);
 - how the data were collected (e.g. cookies, web beacons, browsers, NFC);
 - the number of units included in the analysis.

When carrying out data cleaning, the requirements of [5.2.2](#) shall apply.

For all methods, the extent to which the findings can or cannot be generalized to larger populations shall be discussed.

6.2.2 Validation of data collection process

The details shall be documented on how the data collection process is monitored to ensure it is carried out as intended, and the resulting impact on accuracy and consistency of the data shall be discussed. Monitoring may include translation checks, consistency checks.

6.2.3 Participant safeguards

Where participants have been recruited and have agreed to participate, approaches to participants shall include a brief description of the principles of participant confidentiality, the general research purposes for which the data are to be used, and the name of the service provider, subcontracting agency and/or the client(s), as appropriate. The participants shall be informed that cooperation is voluntary and participants may withdraw at any time.

6.2.4 Weighting

If there is a need for weighting the results of digital analytics or web analyses to the target population researched, then the manner and means of the weighting factors shall be recorded.

6.3 Protection of individuals

In order to guarantee as far as possible the anonymity of the individual persons who are to be analysed, possible quotes shall be anonymized during the reporting in such a way that the identification is no longer reasonably possible at any point.

As far as possible, service providers and clients shall ensure that those whose statements or sentiments expressed via social media or those whose behaviour is registered while visiting websites will not experience adverse or direct personal consequences, e.g. non-consented individually targeted messaging, as a result of the data collection, irrespective of whether the persons concerned have been asked permission for the collection of the specified data.

Where a combination of data from different sources may lead to the potential identification of participants, all reasonable efforts shall be made to protect the anonymity of participants in line with the relevant professional codes of conduct.

6.4 Device monitoring

Service providers shall ensure that they have a fair and lawful basis for the monitoring of devices including the collecting and processing of personal data.

- a) Service providers shall specify types of data that can be collected by device monitoring and, where relevant, obtain consent. This can include:
 - website information such as ad impressions, advertising clicks, image clicks, link clicks, scrolls, downloads, page hits and visits, online transactions, keystroke logging, date, and time;
 - device information such as geo-location, apps downloaded, which apps are used and how often, contents of apps, websites visited and/or bookmarked, audio, images, videos played and/or stored on devices;
 - incoming, outgoing and missed calls, text messages/emails sent and received, audio and visuals taken, recording of calls;
 - ambient audio, images, and video when the device is and is not in use;
 - IP addresses, operating systems, browser user agents, cookie UIDs;
 - identifiable information that will be shared with third parties and for what purpose.
- b) Service providers shall inform participants of any likely consequences of device monitoring, such as substantial use of device memory, battery power, or impact on internet access, and obtain consent if devices, downloaded software, apps, or other programming can remove, revise, deactivate, or otherwise affect any other software or settings.
- c) Service providers shall obtain consent for every install of any device, software, app, or other programming, such as cookies or web beacons that can be used to capture behaviour; or where