

ICS: 03.100.20

A 10



中华人民共和国国家标准

GB/T ×××× - ××××

市场、民意和社会调查中数字分析与网络分析
术语和服务要求

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

**(ISO19731:2017, Digital analytics and web analyses for
purposes of market, opinion and social research
–Vocabulary and service requirements, IDT)**

××××-××-××发布

××××-××-××实施

国家市场监督管理总局 发布

国家标准化管理委员会

目 次

前言	III
引言	IV
1 范围	1
2 引用标准	1
3 术语和定义	2
4 调研项目管理要求	7
4.1 组织和职责	7
4.1.1 调研过程管理	7
4.1.2 项目管理人员职责	7
4.1.3 调研项目质量经理任命	8
4.2 信息的保密性	8
4.3 文件记录要求	8
4.3.1 基本要求	8
4.3.2 文件控制	8
4.3.3 调研项目文件的管理	8
4.3.4 记录管理	9
4.4 能力和培训	9
4.5 分包或外包	9
4.6 评审调研过程要求的有效性	10
4.6.1 项目管理	10
4.6.2 问题及投诉管理	10
4.6.3 内部审核	10
5 调研项目管理要求	10
5.1 服务提供商提交客户的建议书和投标书	10
5.2 由服务提供商确定的其他方面	11
5.2.1 基本要求	11

5.2.2 数据清洗和编辑 11

5.2.3 情感和/或文本分析..... 12

6 项目的执行 13

6.1 基本要求..... 13

6.2 数字分析与网站分析数据收集..... 13

6.2.1 数据收集方法 13

6.2.2 数据收集过程验证据 14

6.2.3 受访者保护..... 14

6.2.4 加权 14

6.3 对个人身份信息的保..... 14

6.4 对设备的监测..... 15

6.5 报告..... 16

6.6 数据保存..... 17

6.7 数据安全..... 17

6.8 调研结果的所有权与发布 18

附录A （参考性）本文件对应GB/T 26316相应条款的参考..... 19

参考文献..... 20

前 言

本文件按照GB/T1.1-2020《标准化工作导则 第1部分：标准化文件的结构和起草规则》的规定起草。

本文件等同采用ISO 19731:2017《Digital analytics and web analyses for purposes of market, opinion and social research – Vocabulary and service requirements》。

本文件对ISO 20252: 2019进行了如下编辑性修改：

1) 在各章节中对大部分ISO标准中表示为“—”的条款，改为“a)、b)、c) 或1)、2)、3)”序号，以便使用。

2) 增加了部分术语的“注”（3.2、3.41、3.70、3.78、3.86、3.92、3.104），提示与GB/T 26316中词条一致的术语，供使用者参考。

本文件由全国市场、民意和社会调查标准化技术委员会（SAC/TC320）提出并归口。

本文件起草单位：中国标准化研究院等

本文件主要起草人：

ISO引言

个人和公司对于数字行为及在线数字表现分析正变得愈发重要。例如：网站访问者行为的测量、通过库迹进行的行为测量，以及社交媒体用户的表现和情感的登记与测量。本国际标准旨在作为ISO 20252的补充并与之同时使用。

本国际标准提供了对服务提供商在数字分析和网站分析研究领域工作方法的洞察，从而为客户提供了其所提供服务的透明度。

用于市场、民意和社会调查目的的数字分析和网站分析，可以与为非研究目的而进行的等效相区分。在这两种情况下，被分析人员的隐私权保护都受到适用于特定项目的法律规定以及相关专业行为守则和其他道德准则的约束。

市场、民意和社会调查中数字分析与网络分析 术语和服务要求

1. 范围

本文件为进行数字分析和网站分析的组织和专业人员规定了术语、定义和服务要求。

适用于通过各种方法和技术，以市场、民意和社会调查为目的而收集、分析和报告数字数据。

本文件还制定了可评估和评价此类服务质量的标准。

适用于由服务提供商主动进行、由客户委托或由客户自己进行的数字分析和网站分析。

本文件适用于数字和网站分析的调研活动，例如：

- 通过使用库迹、网页展示和其他方式了解网站的使用情况，以及跨网站浏览，浏览者花费的时间及其行为；
 - 在线测量样本库，例如：通过安装在样本库成员台式电脑、移动设备或平板设备上的测量软件进行的网络访问的持续测量；
 - 通过基于标签的解决方案，在全域范围内测量在线使用情况，这种方式可以与在线样本库数据整合提供混合测量；
 - 社交媒体分析，收集、汇总和分析，诸如由“微博”和“微信”等提供的在线评论，以及用户生成的内容，例如：博客、论坛和对新闻网站或其他网站的评论。
- 注：全域也可理解为全人群。

本文件应解释为涵盖所有形式的数字数据收集，包括台式电脑、平板电脑、移动设备和OTT设备，以及IoT（物联网）设备（如适用）。

2. 规范性引用文件

下列文件中的内容通过文中的规范性引用而构成本文件必不可少的条款。其中，注日期的引用文件，仅该日期对应的版本适用于本文件；不注日期的引用文件，其最新版本(包括所有的修改单)适用于本文件。

GB/T 26315 市场、民意和社会调查 术语

GB/T 26316 市场、民意和社会调查（包括洞察与数据分析） 术语和服务要求

3. 术语和定义

下列术语和定义，以及GB/T 26315术语和定义适用于本文件。

3.1

广告展示 ad impression

在设备上展示广告。

[参见GB/T26316 中3.4]

3.2

算法 algorithm

在计算或求解其他问题过程中遵循的流程和规则。

[参见GB/T26316 中3.5]

3.3

匿名化 anonymisation

删除、模糊、整合或更改标识符的过程，目的是防止原始数据相关个人身份信息泄露。

3.4

应用程序编程接口API application programming interfaceAPI

是一组定义，在其基础上计算机程序可以与另一程序或组件通信，并且还可以支持内部或外部的访问/数据交换。

3.5

审核 audit

系统、独立并有书面记录的过程，目的是为了获取并客观评估审核证据，以确定符合标准要求的程度。

[参见GB/T26316 中3.9]

3.6

审核标准 audit criteria

用作判断所依据的政策、程序或要求。

3.7

自动程序 bot

作为用户或程序的代理，运行或模拟人类行为的自动软件。

[参见GB/T26316 中3.13]

3.8

客户 client

需要或委托调研项目的个体、组织、机构或部门，可以是内部或外部服务提供商

[参见GB/T26316 中3.14]

3.9

编码 code

与各类回应相关的字符和组合，包括数字、字母或是其他类型字符。

[参见GB/T26316 中3.15]

3.10**编码框 code frame**

将观察或表述的回答进行归类所需的分类编码（3.8）目录。

[参见GB/T26316 中3.16]

3.11**同意 consent**

在收集受访者（3.27）数据之前，受访者根据所获得的充分信息自愿给予的同意。

[参见GB/T26316 中3.20]

注：当研究涉及儿童时，应适用额外的同意要求，例如：负责的监护人的许可。

3.12**库迹 cookie**

一小段存储在浏览器中的信息（即程序代码），以识别浏览器的访问活动和时间。

[参见GB/T26316 中3.21]

3.13**可视化界面（仪表盘） dashboard**

可以评审、管理和展示一些小型应用的软件。

例如：小型应用包括用于导出数据或允许API访问的应用程序。

[参见GB/T26316 中3.22]

3.14**数据收集 data collection**

通过不同手段收集信息的过程。

注：包括诸如网页监测等活动。

3.15**设备标识 device identification****设备ID device ID**

与可被识别的特定数字设备相关联的特殊数字或编码。

[参见GB/T26316 中3.32]

注：这样的设备可以具有多个设备ID，每个设备ID用于不同的目的，例如：用于启用WiFi或蓝牙的服务，或在移动运营商网络上特别识别特定设备的运行。

3.16**数字分析 digital analytics**

以测量和了解人群及其行为为目的的电子数据的收集、分析和报告。

[参见GB/T26316 中3.33]

3. 17

数字指纹 **digital fingerprint**

一种部署**算法**（3. 2）的技术，可分析设备上的大量技术特性和设置，以便生成唯一的识别符，该唯一识别符可以使特定计算设备生成**机器ID**，并且可被个人识别。

[参见GB/T26316 中3.35]

注1：此ID可能会因设备上执行的操作而随时间变化。

注2：也称为设备、机器或浏览器指纹。

3. 18

软件狗 **dongle**

小型硬件设备，通常是USB设备，可以连接到其他电子设备，还可提供附加服务，例如访问互联网。

[参见GB/T26316 中3. 37]

3. 19

酬谢 **incentive**

为了提高受访者（3. 27）的配合度而给予潜在受访者的礼品、现金或其它报酬。

[参见GB/T26316 中3. 47]

3. 20

内部审核 **internal audit**

由公司内部经过培训的员工进行的定期检查，确定公司内部执行项目是否按照流程执行。

[参见GB/T26316 中3. 51]

3. 21

物联网 **internet of things**

IoT

把相关联的物品、人、系统和信息资源与智能服务连接在一起的基础设施，允许他们处理来自现实和虚拟世界的信息，并作出反应。

3. 22

机器学习 **machine learning**

不需要明确的编程即可让计算机产生反应的技术。例如：语音识别、高效率的网页搜索。

[参见GB/T26316 中3. 52]

3. 23

恶意软件 **malware**

各种类型的攻击性或入侵性的软件。包括计算机病毒、蠕虫、木马、勒索软件、间谍软件、广告软件、恐吓软件等其他恶意程序。

3. 24

移动设备 mobile device

一种便携式电子设备，可收集数据并实时或利用适当的软件同步上传给第三方。

[参见GB/T26316 中3.34]

注：移动设备包括手机、智能手表、健身或健康跟踪设备、平板电脑及地理定位设备。

3.25

近场通信 nearfield communication

NFC

使设备之间能够进行短距离通信的一种无线技术。

[参见GB/T26316 中3.55]

3.26

互联网电视 Over the Top

OTT

通过互联网而不是通过服务提供商专用、受监管的互联网协议电视（IPTV）网络所提供的数字服务。

注：提供数字服务的示例包括：小米电视、CNTV、华为烽火、乐视等。

3.27

受访者 participant

提供数据的个人。

[参见GB/T26316 中3.62]

3.28

质量经理 quality manager

在组织内部负责质量要求的制定以及实施质量控制的程序和测量的人员。

3.29

机器人指令文件 robot instruction files

用于定义互联网搜索引擎应如何与网站页面和文件交互的文件，通常用于定义自动化系统中不允许访问的部分。

[参见GB/T26316 中3.84]

3.30

样本 sample

从目标总体（3.99）中抽取的用于采集数据的子集。

[参见GB/T26316 中3.86]

3.31

情感 sentiment

与声音、图像或语言有关的情绪，通常是从积极到中性再到消极的连续过程。

注：该信息是为了方便本文件的用户，并不构成我们对这些产品的认可。

[参见GB/T26316 中3.93]

3.32

服务提供商 service provider

调查服务提供商 research service provider

使用统计数据和/或社会科学方法和技术，执行全部或部分市场、民意和社会调查项目的组织。

[参见GB/T26316 中3.92]

3.33

社交媒体 social media

人们利用网络在线技术和服务，分享彼此意见、见解、经验和观点，由传统的一对多的交互方式转变为多对多的互动方式的媒介。

[参见GB/T26316 中3.95]

3.34

间谍软件 spyware

未经受访者（3.27）同意（3.20）而获取受访者信息或行为的设备或软件。

[参见GB/T26316 中3.96]

3.35

分包 subcontracting

外包 outsourcing

调查服务提供商将项目的部分工作委托第三方服务机构的行为。

[参见GB/T26316 中3.97]

3.36

目标总体 target population

调研项目中可推及总体的人群。

[参见GB/T26316 中3.99]

3.37

文本分析 text analysis

用于描述信息特征的方法。

[参见GB/T26316 中3.100]

3.38

独立访客 unique visitor

独立用户 user

在指定时间内，无论访问频率如何，推断唯一的、不重复的请求网页访问的个人。

[参见GB/T26316 中3.102]

3.39

验证 validation

使用客观证据来确认是否满足已定义的预期用途、应用或结果的要求的过程。

[参见GB/T26316 中3.103]

注：验证可以在真实条件下或在模拟环境中进行。

3.40

网络分析 **web analyses**

对网络平台用户/受访者(3.27)的行为、语言及情绪(3.93) 进行分析和报告。

[参见GB/T26316 中3.107]

3.41

网络信标 **web beacon**

web bug

在网站上用于跟踪网页活动的一段代码(3.15)，通常为1×1像素。

注：与信标不同，信标是一个低功率的发射机，通知附近设备表示自身存在，并触发相关动作。

3.42

加权 **weighting**

一种计算过程，通过赋予数值对不同单元或子集进行重新计算，改善或调整样本（3.86）估计值的代表性。

[参见GB/T26316 中3.109]

4. 调研项目管理要求

4.1 组织和职责

4.1.1 调研过程管理

数字分析和网络分析的服务提供商应为其调研项目采用满足本文件所有要求的调研过程。

调研过程应形成文件。由于以下因素，每个组织的调研过程可能不同：

- a) 组织的规模和调研活动的类型；
- b) 调研项目及其相互作用的复杂性和相关风险；
- c) 人员的能力。

按照本文件的要求为完成不同任务所需的程序、指南和方法都应形成文件，并加以实施、监控、维护，且可被审核。

文件可针对一个或多个程序的要求。一项形成文件的程序的要求可以包含在多个文件中。文件可以是任何形式或类型的媒介形式。

4.1.2 项目管理职责

作为项目管理的一部分，服务提供商应承担以下责任：

- a) 承诺提供适合于调研项目目的的客户服务质量（包括质量政策声明）；

- b) 将服务提供商的组织架构形成文件，包括参与提供调研服务的人员职责；
- c) 审核和改进调研项目；
- d) 确保为该调研过程提供充分和适当的资源和信息，包括任命调研质量经理；
- e) 确保参与提供调研服务的每个人都熟悉适用的国家和国际的伦理和职业准则、相关法律要求以及文件化的程序和方法。

4.1.3 调研项目质量经理任命

应任命一名调研项目质量经理，向其授权，使其足以负责整个调研过程管理并负责组织内部审核，确保本文件的实施。

注：在某些情况下，调研项目质量经理可以是兼职，而在其他情况下，任命多名调研项目质量经理可能更有效（即分担职责）

4.2 信息保密

客户为开展调研项目而提供给服务提供商的所有信息应得到最严格的保密处理。所有信息只能用于该调研项目，未经客户授权不得向第三方提供。保密信息应被安全储存。参阅第6.7。

服务提供商获得的与特定客户有关的研究结果，未经原客户授权不得用于其他客户，且应严格保密。

受访者（3.27）提供的数据及其个人身份信息都应被严格保密。直接或间接向受访者作出的所有保证都应得到履行。例如社交媒体供应商对其用户作出的保证。

4.3 文件要求

4.3.1 基本要求

服务提供商应建立和维护调研记录 and 文件（可以是纸质版和/或电子版），以提供可追溯性证据。应说明调研项目在多大程度上涉及特定的隐私法规。

应使用最新的杀毒软件对所有电子文件进行恶意软件检查。

4.3.2 文件管理（不包括调研项目相关文件）

服务提供商应对调研过程所需的文件进行管理。应对所有文件进行版本控制，以明确识别当前的版本。

4.3.3 调研项目文件的管理

每一个调研项目应有项目文件（可以是电子版形式），该文件包含项目基本要求的记录或该记录的出处说明。

除了4.3.2的要求和唯一的项目标识外，文件应具有唯一标识，以便追溯和确保

被检索到。

4.3.4 记录管理

为提供符合本文件和项目要求的证据而建立的记录应安全保管，防止未经授权的使用，且可使用并在规定的期限内保存（例如备份）。

4.4 能力和培训

服务提供商应：

- a) 详细说明并以文件记录在调研过程中每一职位所需的技能和能力；
- b) 为所有参与收集和处理调研数据的人员提供适当的培训，包括让他们容易获得适当的培训材料；
- c) 确保培训充分涵盖处理和收集个人数据所需的所有技术，以及伦理和法律方面的考量；
- d) 确保所有参与处理和收集个人数据的工作人员和分包商都受适当的管理监督。

4.5 分包或外包

服务提供商应对所有与研究项目有关的服务负全部责任。研究项目的所有服务，包括与本文件有关的任何分包和/或外包的工作，除非分包商的选择超出了服务提供商的控制范围。

服务提供商应明确制定在调研领域里选择分包商的程序，与分包商建立合同关系并控制其所提供服务的质量。

服务提供商应确保分包商了解本文件中与分包过程有关的要求，并提供有关这些要求的书面说明。

注：被认为合适的书面说明详细程度可以取决于分包商是否曾为服务提供商开展过类似的工作，以及分包商是否能够在必要时提供符合本文件的必要证据。

服务提供商应从分包商处获得确认，即调研相关工作的开展符合本文件的规定。适当的确认应包括以下一项或多项内容：

- a) 按照本文件的要求开展工作的书面合同协议；
- b) 对本文件的评估和认证；
- c) 分包商提供的对发包方所要求的适当检查和验证的文件证据；
- d) 由服务提供商进行适当的检查和验证。

服务提供商应保留相关记录，包括从分包商处收到的任何对服务不满意以及所采取措施的记录。这些记录均应可追溯到项目。

4.6 评审调研过程要求的有效性

4.6.1 项目管理评审

服务提供商应通过以下方式监督和管理商定的研究过程及其对调研项目的管理：

- a) 根据政策监测和评审调研过程管理的运作情况（如内部审核）；
- b) 根据商定的客户要求评审项目和成果；
- c) 定期（至少每年一次）监测和评审客户满意度，以便提高服务质量。

4.6.2 问题和投诉管理

当问题和投诉被识别时（无论是在调研过程中或结果中），应予以纠正，并采取措施防止再次发生。

应保留处理记录，包括投诉、问题、原因、措施和解决方案的详细信息。

4.6.3 内部审核

服务提供商应按预先计划的时间周期进行内部审核，以确定符合以下内容：

- a) 调研过程的要求；
- b) 与客户商定的项目要求，包括在整个项目期间发生变更的书面记录。

项目审核体系应包括审核计划、范围和方法。调研质量经理负责组织内部审核体系（见4.1.3）。所有的内部审核人员都应接受培训。在组织规模允许的情况下，内部审核人员不可对自己的工作进行审核。审核记录、后续措施和结果应予以记录，并可酌情提及客户的项目。

5. 建议书和投标书

5.1 服务提供商提交客户的建议书和投标书

建议书和投标书应以书面形式，线上或线下提交。

如果项目需求书中存在不确定事项，服务提供商应就这些事项联系客户。

建议书和投标书应明确说明客户和服务提供商在工具使用和结果实施方面的责任。

在建议书和投标书中，至少应涵盖以下内容：

- a) 客户的信息要求；

注：某些信息需求可能只能在项目执行期间具体确定。

- b) 服务提供商推荐的数字分析或网络分析方法或技术；
- c) 要分析的数据源，以及在选择数据源时采用的标准；如果服务提供商和分包商对数据源的定义不同，建议书和投标书应阐明所使用的定义；
- d) 明确说明在数据采集过程中所使用的会决定数据质量的技术与抽样维度；
- e) 数字分析或网络分析将包含的确切时间期限；
- f) 调研结果的范围、类型和交付频率；
- g) 收集数据方存储数据的方式，以及保存时间都应与客户达成一致；
- h) 项目完成后，调研结果的所有权；
- i) 项目执行成本。阐明任何可能发生的项目成本，包括只有在完成分析后才可以确定的成本；应说明开具发票的方式和时间；
- j) 将要或可能转包给其他外部服务提供商的环节，根据要求，应告知客户相关分包商的身份。在接受客户委托之后，如果确实发生了某些服务必须从分包商处购买，而这一点并未在建议书或投标书中写明，则服务提供商应立即告知客户；
- k) 相关法律法规，以及适用的专业行为准则；
- l) 关于在执行数字分析和网络分析项目时服务提供商遵守本文件的声明；

注：认证是提供遵守本文件证据的一种方式。

向潜在客户提交的建议书或投标书应被列为“保密”的活动，即使提交建议书或投标书之后并未收到客户委托。

5.2 服务提供商需确定的其他方面

5.2.1 基本要求

为促进对项目及其自身能力和/或限制的透明度和理解，服务提供商及限制的理解，服务提供商应准备并随时与客户讨论（如有需要）章节5.2.2、5.2.3和5.2.4中的信息。

5.2.2 数据清洗和编辑

服务提供商应记录并随时提供与研究目的相关的数据清理过程。

- a) 根据目标的不同，要清洗的数据可包括：
 - 1) 同形同音异义词（例如：小米（谷物）与小米（品牌））；
 - 2) 来自客户及其竞争对手的数据，而不是来自消费者的数据（例如客户的推文）；

- 3) 第三方广告、优惠券、宣传材料等（例如：一家鞋店宣传他们自己的鞋品牌）；
- 4) 垃圾数据（例如：博客评论可能会说“购买运动鞋、购买运动鞋、购买运动鞋、购买运动鞋”），虚假的社交媒体帐户；
- 5) 源自新闻渠道的数据（例如，链接到杂志文章的推文）。

b) 在适当的情况下，服务提供商应明确其流程是否可以识别和删除无效或非人为流量，例如：

- 1) 机器人和爬虫；
- 2) 无效浏览器；
- 3) 内部流量；
- 4) 为欺骗广告商/出版商而激励浏览；
- 5) 其他无效流量来源。

c) 服务提供商应准备并随时与客户讨论（如有需要）以下信息：

- 1) 使用手动、自动算法或两者的组合进行数据清理的程度和细节；
- 2) 解释即将应用的数据清理策略的预期效果，包括可以识别和清理不适当记录的程度和过程，以及谁可以执行此类任务；
- 3) 从有效数据集中删除数据的标准；
- 4) 如何处理、填充缺失值数据的标准（例如“0”填充，“数据中和”，推及）。

5.2.3 情感和/或文本分析

服务提供商应准备并随时与客户讨论（如果要求）以下信息：

a) 情感和/或文本分析是手动（包括由软件辅助编码），自动使用算法（不必对每条数据进行人工编码），或某种组合（例如机器学习）；

b) 对情绪分析的类型和方法的解释；

c) 对文本分析的类型和方法的解释；

d) 用于分析的数据类型（例如短句，长句，非句子，表情符号，俚语，褒贬语，讽刺，首字母缩略词，拼写错误的单词）；

e) 评估分析的可靠性，准确性和有效性的频率和过程（例如，每周或每年，验证的记录的数量和类型，是否使用盲编码），包括前期验证；

f) 分析是否在逐个任务的基础上进行微调；

g) 是否可以识别可能的编码错误和/或谁将进行数据的重新编码；

h) 包含或排除的语言，多语言信息；

a) 关于情感或文本分析的任何质量测量的结果，例如编码器间可靠性打分或指数（在手动编码的情况下）以及原始自动编码的手动重新编码。

5.2.4 网站使用和测量分析

服务提供商应准备并随时与客户讨论（如果需要）以下信息：

a) 收集的数据类型（例如IP地址，操作系统，浏览器用户代理，库迹唯一标识符，地理位置，记录事件的数据或时间戳）；

b) 数据是否反映：

1) 唯一的库迹（非重复的库迹，可以表示在测量期间访问互联网内容，并且可以包括使用同一帐户的同一计算机上的多个用户）；

2) 唯一的浏览器（调整库迹删除后的唯一库迹数量）；

3) 唯一的设备（使用一台计算机上多个浏览器后，计算唯一的浏览器数量）；

4) 独立访客（3.38）；

c) 报告的时间基准（例如每日，每周，每月），适用的时区以及访问或所花费时间的定义方式；

d) 在进行编码时，使用什么规则来解决以下问题，例如，放弃，处理多个标签和不正确的标签，确定使用持续时间，以及数据调整的支撑、规则；

e) 所收集的数据是来自可用的全域还是来自对样本的推测，数据的构成，可用的人口统计数据水平，以及是否提供数据匹配或浓缩；

f) 是否能处理库迹共享，删除，或拒绝，如果是，则如何处理；

g) 对提供的数据和/或验证或审核能力的任何限制，例如对数据保留和/或积累期限的限制。

6. 项目的执行

6.1 基本要求

当客户接受建议书或投标书之后，服务提供商应任命一名负责的项目经理。并告知客户这一任命，因为项目经理应是客户的第一联系人。项目应根据数字信息源的相关条款和条件执行。如果项目发生变化或偏离预期。

有标准程序的变化或偏离，服务提供商应按照标准流程，尽可能提前告知客户这种变化的预期结果，或在此之后尽快告知客户。

6.2 数字分析与网络分析数据收集

6.2.1 数据收集方法

服务提供商应当记录所使用方法的详细信息。包括以下内容：

a) 从网站上收集在线评论的方法，包括但不限于社交媒体，例如；

1) 包括各种网站及其他社交媒体论坛的总体；

2) 如何选取调查中包括的网站、个人或其中的作者及其评论或发帖；

- 3) 如何收集评论和/或帖子的方式（例如从问卷或用户生成的内容）；
- 4) 使用的搜索算法，如适用；
- 5) 纳入分析的网站、作者、评论或包含在分析中的帖子数量。

b) 通过数字方式进行的行为观察和/或方法：

- 1) 目标总体；
- 2) 选取和招募调研的个体或单元的方式；
- 3) 所研究单元的定义（例如：设备、网站、个体）；
- 4) 进行数据收集的方式（例如：库迹、网络信标、浏览器、NFC）；
- 5) 纳入分析的单元数量。

当进行数据清洗时，应适用5.2.2的要求。

对于所有研究方法，应对研究结果是否能够推广到更广泛群体进行论述。

6.2.2 数据收集过程验证

对如何进行数据收集过程的监控应予以文件记录，以确保其执行符合预期，并且对数据准确性和一致性的结果影响也应进行论述。监控可能包括翻译检查和一致性检查。

6.2.3 受访者保护

在受访者同意参与调查项目的情况下，服务提供商应向受访者简要介绍个人信息进行保密的原则、数据将用于何种一般研究目的、服务提供商的名称，也可根据情况告知受访者分包机构和/或客户的名称。应告知受访者合作是自愿的，受访者可随时退出。

6.2.4 加权

当需要对研究目标总体的数字分析或网站分析的结果进行加权时，应记录加权的方式以及所采用的加权变量。

6.3 个人身份信息保护

为了保障参与调研分析的个人信息安全，在报告中对个人信息的引用应进行匿名化处理。不再出现个人标识。

对个人通过社交媒体发表的言论和情感或网站浏览行为，服务提供商和客户应尽可能确保对其不产生不利影响或直接后果。例如，个人未同意情况下的消息推送，导致直接后果的数据采集，以及不考虑个人是否同意情况下的特定数据的采集。

多种来源的数据混合可能导致受访者身份被识别，应尽可能参照相关专业准则保证受访者信息的匿名性。

6.4 设备监测

服务提供商应确保其对设备的监测是公正和合法的，包括对个人数据的收集和处理。

a) 服务提供商应指定可通过设备监测所收集的数据类型，并在相关时获得同意。这些数据类型包括：

- 1) 网站信息，例如：广告展示、广告点击、图片点击、链接点击、屏幕滚动、下载、页面点击和访问、在线交易的次数，按键记录、日期和时间；
- 2) 设备信息，例如：地理位置、下载的应用、应用的使用和频率、应用的内容、访问和/或加入标签的网站、音频、图像、在设备上播放和/或存储的视频；
- 3) 呼入、呼出和未接电话、发送和接收的短信或电子邮件、拍摄的音频和视频、电话录音；
- 4) 当设备正在使用和未使用时的环境音频、图像和视频；
- 5) IP地址、操作系统、浏览器用户代理、用户标识、库迹；
- 6) 将与第三方共享的可识别个人身份的信息以及分享目的。

b) 服务提供商应告知受访者对设备的监测的任何可能后果，例如：大量占用设备内存、电池电量或对互联网访问的影响，并获得设备、下载的软件、应用程序或其他程序是否可以删除、修改、停用，或以其他方式影响任何其他软件或设置的许可。

c) 服务提供商应在每次安装任何设备、软件、应用程序或其他程序（例如：可用于捕获行为的库迹或网络信标），或者通过诸如U盘或软件狗等硬件进行设备监测，均应取得同意。

服务提供商应根据项目及相关要求，定期且以合理的时间间隔（但不得低于每年一次）获得持续数据收集的同意。

d) 服务提供商应告知受访者以下事项（如适用）：

- 1) 提前告知任何设备、软件或应用程序的每次更新、修订、停用和卸载；
- 2) 监测是否以及何时将过期的具体时间范围，以及监测是在特定时间点还

是连续进行；

3) 随时撤销许可、移除和停用服务和设备以及完全且合理方便执行的过程（视情况而定）。

e) 服务提供商应遵守网站上关于允许收集哪些类型的信息的机器人指令文件。

f) 服务提供商不得使用间谍软件。

6.5 报告

服务提供商应向客户提供足够的信息，使其能够理解数据分析、任何潜在的误差和偏差来源，以及偏差控制措施。

对项目调查成果的诠释应切中调查主题，应有项目实证研究发现的支持，如适用，还应有其它数据的支持。基于实证研究成果的诠释和建议，应当与个人的观点和看法区别开来。

a) 项目完成后，对客户的报告应包括（或以其他方式提供）：

- 1) 客户名称；
- 2) 服务提供商的名称；
- 3) 任何全部或部分分包的过程的详细信息，包括诸如算法类型及其目的等内容；
- 4) 项目的目标；
- 5) 项目的目标总体。

b) 对于从网站（例如但不限于社交媒体）收集的在线评论，报告应包括：

- 1) 所包括的各种网站及其它社交媒体；
- 2) 所分析的网站、作者、评论的数量，和/或包含在分析中的帖子数量；
- 3) 选取网站、个体或其中的作者以及他们的评论或帖子纳入研究的方式；
- 4) 收集评论和/或帖子的方式，以及所采用的搜索算法（如适用）。

c) 对于通过数字手段进行的行为观察和/或测量，报告应包括：

- 1) 纳入分析的单元数量；
- 2) 选取和招募参与研究的个体或单元的方式；
- 3) 若给予了酬谢，说明酬谢的类型；
- 4) 所研究单元的定义（例如：库迹、浏览器、个体）；
- 5) 数据的收集方式。

d) 对于所有研究方法，报告应包括：

- 1) 分析覆盖的时间段；
- 2) 如果进行了加权，说明加权程序和加权变量；
- 3) 如果进行了估算和插值，说明估算及插补的程序；
- 4) 对潜在的抽样及其他与调研相关的问题如何进行处理的说 明，以及调研结果是否能够推广到更广泛的群体；
- 5) 标准规程的任何更改或偏差，以及这些更改的预期结果；
- 6) 详细说明监控数据收集过程的方式，以确保数据收集按预期执行，以及由此产生的对数据准确性和一致性的影响；
- 7) 声明：项目的执行符合本文件。

注：认证是提供遵守本国际标准的证据的一种方式。

如果初始的项目建议书与最终报告之间存在偏差，应提及最终报告偏离原建议书中所述研究目的的程度、导致这种偏差的原因，以及为尽可能把这种偏差控制在有限的范围内而采取的措施。

6.6 数据保存

服务提供商应保存包括报告在内的数据，以便在项目完成后的指定时间期限内能够进行审核、识别趋势、调查错误或答复客户的问题（根据客户可能的要求，以及与数据量有关的技术限制）。

服务提供商应告知客户有关保存期限的情况，包括对于不同类型的数据和/或技术的任何变化。

6.7 数据安全

服务提供商应为员工提供充分的数据中心、处理和报告服务器以及一般系统的访问控制技术和协议条款的访问，以及加密和密码策略。服务提供商应保证安全保障措施足以确保只有获得授权的人方可访问系统和数据。

如果客户和受访者提出要求，服务提供商应提供与安全问题相关的信息。这些信息应包括组织和技术保障措施，例如：对个人数据的访问控制、云安全、灾难恢复要求、维护与主要计算机操作系统和浏览器连续版本保持兼容性的程序，以及供应链保护。

6.8 调研结果的所有权与发布

未经相关客户许可，服务提供商不得发布代表客户执行的项目的全部或部分结

果。

注1： 项目拥有者，有权将其发布或提供给第三方。

如果发布，服务提供商应提醒客户调研结果及必要说明的呈现存在局限性。

有关原始数据可用性的状态，通常应在项目的建议书阶段由服务提供商向客户明确说明。

注2： 一般而言，客户及其服务提供商本身并没有收集到可供其使用的原始数据。但是，在很多情况下，分包商拥有收集到的可供使用的原始数据。

如适用，服务提供商应在合同条款和条件、建议书、投标书、报告和项目结果的任何出版物中明确提及遵守本文件。

注3： 认证是提供符合本文件的证据的一种方式。

附录A

本文件中对应GB/T26316-××××条款的参考

表A.1列出了本文件对应GB/T26316-××××的条款。

表A.1

本文件条款	对应GB/T26316-××××的条款
3.5 审核	3.9
3.8 客户	3.14
3.9 编码	3.15
3.10 编码框	3.16
3.19 酬谢	3.47
3.32 服务提供商/调查服务提供商	3.92
3.30 样本	3.86
3.33 社交媒体	3.95
3.35 分包/外包	3.97
3.36 目标总体	3.99
3.42 加权	3.109
4.1.2 项目管理职责	4.2
4.2 信息保密	4.1.2
4.5 分包或外包	4.4
4.6.1 项目管理评审	4.6
4.6.2 问题和投诉管理	4.6.1
4.6.3 内部审核	4.7
5.1 服务提供商提交客户的建议书和投标书	D.2
6.2.2 数据收集过程验证	D.3.1.2
6.2.3 受访者保护	D.3.1.3

参考文献

- [1] ISO 26362 市场、民意和社会调查中可访问样本库 术语和服务要求
-