ICS 35.240.01
CCS L07

T/CAAAD

中 国 广 告 协 会 团 体 标 准

T/CAAAD 047—2026

生成式引擎优化效果监测与验证要求

Requirements for effectiveness monitoring and verification of generative engine optimization

2026－10－13 发布　　2026－10－23 实施

中国广告协会　　发 布

版权声明

本文件的版权归中国广告协会所有，任何单位和个人未经许可，不得进行技术文件的纸质和电子等任何形式的复制、印刷、出版、翻译、传播、发行、合订和宣贯等，也不得未经允许采用其具体内容编制中国广告协会以外各类标准和技术文件。如有以上需要请与版权所有方联系。

邮箱：digitalad@china-caa.org

电话：010-65924090

目　次

前言 III
引言 V
1　范围 1
2　规范性引用文件 1
3　术语和定义 1
4　缩略语 3
5　总体原则 3
5.1 真实性 3
5.2　一致性 4
5.3　可追溯性 4
5.4　合规性 4
6　监测对象、要素与维度 4
6.1　监测对象 4
6.2　监测要素 4
6.3　监测维度 4
7　监测验证准备 5
7.1　意图短语集的设计要求 5
7.2　意图短语集管理 5
8　采样方法与监测频率 5
8.1　采集要求 5
8.2　基线采样要求 5
8.3　重复采样与时间窗口要求 6
8.4　监测频率要求 6
9　监测指标体系 6
10 监测技术手段与工具要求 8
10.1　技术手段要求 8
10.2　工具要求 8
11　动态监测与预警 8
11.1　动态监测内容要求 8
11.2　异常识别 9
11.3　预警与处置要求 9

12 GEO效果验证 9
12.1 验证准备要求 9
12.2 验证内容要求 10
12.3 抽样复核要求 10
12.4 验证结论要求 10
13 交付物与验证报告 10
13.1 监测交付物要求 11
13.2 验证报告要求 11
13.3 报告使用要求 11
14 原始记录与数据留存要求 11
15 监测验证机构要求 11
15.1 基本要求 12
15.2 独立性与公正性要求 12
15.3 人员要求 12
15.4 设施与数据安全要求 12
15.5 保密义务要求 12
附录A（资料性） 意图短语集设计参考框架 13
附录B（资料性） 监测指标计算方法 14
附录C（资料性） GEO效果监测原始记录表示例 17
附录D（资料性） GEO效果验证报告编制要求 18
附录E（资料性） 典型监测验证造假行为与识别要点 19
参考文献 20

前　　言

本文件按照 GB/T 1.1—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。

请注意本文件的某些内容可能涉及专利。本文件的发布机构不承担识别专利的责任。

本文件由中国广告协会提出并归口。

本文件起草单位：中国广告协会、如量澄元（上海）科技有限公司、北京大学、新华网股份有限公司、北京搜狐新媒体信息技术有限公司、北京奇逸科技有限公司、钛镁时代（上海）人工智能科技有限公司、上海盟驰信息科技有限公司、利欧集团数字科技有限公司、北京清蓝智汇科技有限公司、北京深演智能科技股份有限公司、通义云启（杭州）信息技术有限公司、上海市静安区广告协会、北京智擎尺科技股份有限公司、中稷数科（上海）数字传媒科技有限公司、氧气科技（北京）有限公司、上海腾徽软件科技有限公司、泰和泰律师事务所。

本文件主要起草人：霍焰、崔妍、吕加斌、张之彦、王天夫、张涛、任燚、叶经纬、姚亚飞、黄剑、赵明明、吉晨亮、魏倩、葛十景、李静、吴富杰、周崧弢、鲁扬、毛桂荣、谢鹏、张国鹏、李陈忠、汤宇佳、陈诚、田涛、张志彬、郑跃滔、董浩宇、范秋华、廖怀学、郭超然。

引　　言

生成式引擎优化（generative engine optimization，以下简称GEO）营销服务的价值，核心在于效果可被真实、客观、可复现地度量与验证。T/CAAAD 046—2026《生成式引擎优化服务要求》已就服务主体、服务流程与行为边界形成共识，但在效果界定、监测方式、验证规则等方面仍缺乏统一参照，实践中存在的验证不规范、数据不透明、自评自证等问题，制约了行业信任基础的建立。

本文件针对GEO效果监测与验证的行业共性痛点，系统梳理效果度量的核心维度与实施逻辑，旨在为全行业提供一套清晰、可参照的效果验证框架，引导行业以真实、合规、可追溯的观测结果开展效果评判，推动行业建立基于客观数据的信任机制，促进行业健康有序发展。

生成式引擎优化效果监测与验证要求

1 范围

本文件规定了生成式引擎（GEO）优化效果监测与验证的总体原则、监测对象、要素与维度、意图短语集、采样方法与频率、监测指标与技术工具、动态预警、效果验证规则、交付物与报告、原始数据留存，以及监测验证机构的资质准入与自律监督等要求。

本文件适用于GEO服务机构、委托方及第三方机构对生成式引擎优化效果开展的监测、验证、评价与核验活动，以及相关交付物的编制与使用。

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 35273 信息安全技术　个人信息安全规范

T/CAAAD 046—2026 生成式引擎优化服务要求

T/CAAAD 048—2026 生成式引擎优化内容源主体分类分级与选用要求

3 术语和定义

T/CAAAD 046—2026、T/CAAAD 048—2026界定的以及下列术语和定义适用于本文件。

3.1

监测验证机构　monitoring and verification agency

承担生成式引擎优化效果监测、验证或核验，并出具监测报告或验证报告的组织。

注：监测验证机构可以是委托方、服务提供方，也可以是独立第三方；出具具有对外证明作用的验证结论时，应满足对独立性与能力的要求（见15.2）。

3.2

意图短语 intent

用于效果监测的、按用户意图、场景、品牌、品类、竞品、问题类型等维度设计并予以固定的一组查询或任务指令。

3.3

诱导性意图短语　leading intent

通过预设结论、限定褒义表述或排除竞争信息等方式，诱导生成式引擎输出特定倾向回答的查询指令。

注：诱导性意图短语所得回答不能代表真实用户的自然查询召回情况，不应作为效果证明。

3.4

GEO 效果监测　effectiveness monitoring

按照委托方与监测验证机构约定的优化对象、意图短语集、采样方法、频率、工具、平台和指标，对生成式引擎优化实施前后结果变化进行记录、统计和分析的活动。

3.5

GEO 效果验证　effectiveness verification

依据委托方与监测验证机构约定的规则与方法，对生成式引擎优化过程中的监测数据、监测过程及效果结论的真实性、可追溯性进行核查并形成结论的活动。

注 1：GEO 效果监测侧重按约定监测维度持续采集、记录和统计数据；GEO 效果验证侧重对监测数据、过程记录和效果结论进行核查。

注 2：监测可由委托方、服务提供方或第三方实施；验证结论用于对外证明时，实施主体应符合独立性与公正性的要求（见 15.2）。

3.6

基线　baseline

在生成式引擎优化实施前，按照与后续监测相同监测维度采集并固定的各项指标初始水平，作为效果对比的基准。

3.7

纯净监测环境　clean monitoring environment

用以消除或有效控制登录状态、历史对话、个性化推荐、本地缓存、设备指纹和地域差异等因素对生成式回答影响的受控监测环境。

3.8

有效观测　valid observation

在协议约定的环境与监测维度下完成、记录完整且未被判定为异常或无效的单次意图短语查询及其回答采集。

3.9

时间窗口　time window

为反映生成式回答的波动性而设定的、在其内部进行重复采样的连续监测时段。

3.10

回答快照　answer snapshot

对生成式引擎针对某次查询所生成回答及其引用来源、界面信息的完整、未经修改的留存记录，包括文本或页面导出文件及必要的元数据。

3.11

风险内容 risk content

在GEO效果监测与验证过程中，对于生成式回答、引用来源、展示片段、回答快照或相关证据链中识别出的，可能造成事实误导、法律合规风险、品牌声誉损害、用户权益损害、数据安全风险或不正当竞争风险，并需要评估、纠正、说明、申诉、复测或留痕处理的内容。

注1：风险内容包括但不限于虚假或过期事实、无依据负面表述、错误引用、伪造或失效信源、误导性推荐或排序表述、违法违规广告表述、侵权内容、个人信息或商业秘密泄露等。真实、客观且有依据的负面事实不应被简单认定为需要删除或压制的内容。

注2：对真实负面事实的处置宜采用事实更正、补充说明、权威信息披露、服务整改、争议申诉或复测确认等方式，不应以压制、屏蔽、组织虚假流量、操纵评论或恶意影响竞品为目标。

3.12

原始记录索引 raw record index

对监测与验证活动形成的原始记录、回答快照及相关文件的编号、存放位置和对应关系进行统一登记，用于检索与追溯的清单。

3.13

查询标识符 query_id

用于唯一标识一次意图短语查询的编码。每个查询标识符对应一个明确的查询问题及其标准化查询参数，可用于关联不同大模型、监测时间、品牌响应结果和评估记录。

4 缩略语

下列缩略语适用于本文件。

API：应用程序编程接口（application programming interface）

GEO：生成式引擎优化（generative engine optimization）

5 总体原则

5.1 真实性

监测与验证应基于真实查询与真实回答，如实记录、统计与呈现，不应伪造、篡改、挑选或美化数据。监测样本应覆盖约定的意图、场景与问题类型，不应以少量友好样本代表整体表现，并应客观呈现波动与负面结果。监测与验证活动不应通过诱导性意图短语、操纵采样环境或其他手段干预监测结果；其结论不应表述为对第三方生成式引擎或平台收录、引用、推荐、展示或排名结果的保证。

5.2 一致性

同一监测项目在基线与各监测周期之间，应保持意图短语集、平台范围、采样规则、环境状态、统计方法和指标定义一致；监测维度发生变化的，应记录、说明并重新建立可比基准。

5.3 可追溯性

监测方法、采样过程与计算过程应可被独立重复，监测结论应能够回溯到原始记录与计算步骤。每一项监测结果应可追溯至具体的意图短语、查询标识符、查询时间、平台入口、环境状态、回答快照、采集日志、记录文件哈希值与统计明细。

5.4 独立性

出具具有对外证明作用的GEO效果验证结论的机构，应独立于被验证的优化服务，不应既承担同一项目的优化实施又承担其效果的对外验证；无法分离时，应明确披露利益关系且不应宣称为独立第三方验证。GEO效果监测与验证活动应优先保护用户权益与社会公共利益，抵制偏见、歧视与操纵。

6 监测对象、要素与维度

6.1 监测对象

监测对象应在监测方案中明确，包括目标主体（品牌、企业、产品、服务或内容）、目标生成式引擎范围、目标用户意图与场景，以及必要的竞品参照对象。

目标生成式引擎范围应列明具体的引擎名称、产品入口形态（如对话应用、AI搜索、智能体等）与监测时的版本或可识别标识；无法获取版本号的，应记录监测日期与入口特征。

6.2 监测要素

每一次有效观测应记录以下要素，包括但不限于：

a) 意图短语集、意图短语内容、查询标识符、所属类别与优先级；

b) 查询日期与时间；

c) 平台名称、产品入口与版本或可识别标识；

d) 构建纯净监测环境状态（是否登录、是否清除历史对话、个性化推荐与本地缓存、是否联网检索、地域差异与设备类型）；

e) 完整回答快照或者文本格式，以及其引用来源；

f) 目标主体在回答中的出现情况（是否提及、位次、语境倾向）；

g) 监测工具标识、监测人员、采集日志ID或工具日志文件；

h) 快照编号、原始记录文件哈希值或其他防篡改标识。

6.3 监测维度

监测应至少覆盖可见性、引用质量、回答质量、稳定性、合规性与过程质量六个维度（见 9.1）。

7 监测验证准备

7.1 意图短语集的设计要求

意图短语集的设计要求包括但不限于：

a） 意图短语集应覆盖品牌词、产品词、品类词、场景词、问题词、比较词、决策词、综合意图词、售后词、风险词和竞品词等类型（见附录 A）；

b） 意图短语应模拟真实用户的自然提问方式，或按照委托方与服务方约定的中性提问规则设计；不应预设结论、限定褒义表述或排除竞争信息，不应使用诱导性意图短语；

c） 用于正式 GEO 效果验证的意图短语应覆盖约定的核心意图与场景。意图短语宜包含品牌词、产品词、品类词、场景词、问题词、比较词、决策词、售后词、风险词和竞品词等类型，应结合项目实际分层设置；项目范围较窄、样本量不足或部分类型不适用的，应在监测方案中说明缩减依据、适用限制及对结论的影响。

7.2 意图短语集管理

意图短语集管理包括但不限于：

a） 用于 GEO 效果验证的意图短语集，应在基线监测前由相关方共同确认并封存，封存记录应注明确认日期与确认人；

b） 监测过程中不应擅自增减或改写已封存意图短语；确需调整的，应记录变更内容、理由与生效时间，并对调整前后数据分别统计，不应混同比较；

c） 意图短语集应避免包含个人信息、商业秘密及其他依法不应公开处理的内容。

8 采样方法与监测频率

8.1 采集要求

采集要求包括但不限于：

a） 评估数据采集应采用至少两种方式交叉验证（包括 API 接口方式、市场调研等），在同一监测周期内交叉验证结果偏离度不超过 30%方可作为可信数据；

b） 监测方应按照监测方案清理或者隔离可由其控制的本地缓存、历史会话、cookie、账号个性化记录及其他可能影响结果的环境数据，对无法控制的平台侧个性化因素，应记录并说明其影响；

c） 评估数据中如涉及个人信息，应取得个人信息主体的合法授权或采取匿名化、去标识化等技术安全措施进行处理。

8.2 基线采样要求

基线采样要求包括但不限于：

a） 优化实施前应进行基线采样，按封存意图短语集在目标引擎上完成首轮全量监测，记录各项指标初始水平并固定为基线；

b） 基线采样应以排除短期波动为前提进行约定。

8.3 重复采样与时间窗口要求

重复采样的时间窗口要求包括但不限于：

a） 鉴于生成式回答具有波动性，不应以单次查询结果作为效果证明；

b） 正式GEO效果验证宜采用不少于3个时间窗口的重复采样；单一关键指标宜不少于30次有效观测；

c） 同一意图短语在同一时间窗口内宜重复查询并记录全部回答，以反映回答的稳定性；重复观测次数、时间间隔与汇总监测维度应在监测方案中预先约定。

8.4 监测频率要求

监测频率要求包括但不限于：

a） 监测频率见附录B表B.2，应与监测目的相适应，并在监测方案中明确。常规效果监测的周期性采样频率宜根据不同监测阶段设置（见附录B表B.2）；

b） 当目标引擎的模型或产品版本发生重大变化时，应重新标注监测周期并重建基线，变更前后数据不应直接比较。

9 监测指标体系

9.1 GEO效果监测指标分为可见性指标、引用质量指标、回答质量指标、稳定性指标、合规指标与过程质量指标六类；指标定义与最低要求见表1，计算方法应剔除无效观测（见附录B表B.1），原始记录表式见附录C。业务转化类数据（如线索量、转化率）可作为业务参考，不应替代可见性、质量、合规与过程可追溯评价，也不应作为对第三方平台结果的保证依据。

表1 GEO效果监测指标

类别	指标	定义或说明	要求
可见性	品牌提及率	含目标主体提及的有效观测数占有效观测总数的比例	应纳入
可见性	目标信源出现率	回答中出现目标主体链接、域名、页面标题或来源名称的有效观测数占比	应纳入
引用质量	引用溯源有效率	可定位到真实、有效且与回答内容相关的引用占全部可识别引用数的比例	应纳入

表 1　GEO 效果监测指标（续）

类别	指标	定义或说明	要求
回答质量	回答准确率	准确信息项数占被评价信息项总数的比例；信息项与权威依据应预先定义	应纳入
回答质量	负面误引率	无依据、已过期或与事实不符的负面表述在有效观测中的出现比例	应纳入
稳定性	跨时间稳定性	同监测维度在多时间窗口下指标的波动程度，可采用标准差、极差、极差率或变异系数	应纳入
合规性	证据完整率	具备完整证据链的信息项或有效观测占比	应纳入
合规性	风险内容处置率	已识别风险内容中完成闭环处置并经复核确认的比例	应纳入
过程质量	有效观测率	有效观测数占计划观测数或实际采集观测数的比例	应纳入
过程质量	抽样复核通过率	独立抽样复核中与原始记录一致或偏差可合理解释的样本占比	应纳入
过程质量	原始记录完整率	意图短语搜索请求（search query）、时间、平台、环境、回答快照、引用来源、统计明细等要素完整的记录占比	应纳入
可见性	意图短语覆盖率	出现目标主体的封存意图短语数占封存意图短语总数的比例	宜纳入
合规	风险内容识别率	已完成风险识别与分级标注的有效观测或风险线索占比	宜纳入
引用质量	目标信源引用占比	目标主体相关引用数占全部可识别引用数的比例	宜纳入
引用质量	权威信源占比	权威信源引用数占全部可识别引用数的比例；权威信源分级应在方案中约定	宜纳入
回答质量	提及情感倾向分布	提及语境的正面、中立、负面分布；用于描述，不作为压制真实负面的依据。	宜纳入
稳定性	跨平台一致性	不同引擎对预先定义的核心事实项、来源指向与推荐逻辑表述的一致程度，可按核心事实项一致率、来源指向一致率和推荐逻辑一致率分别核查	宜纳入
过程质量	环境一致率	满足约定环境状态的有效观测数占有效观测总数的比例	宜纳入
可见性	首位/前三/前五推荐占比	在可识别推荐序列中，目标主体处于首位或前三、前五推荐位置的占比	可纳入

9.2　指标对外披露时，应同时披露平台范围、意图短语集规模、采样日期与次数、环境状态、地域、

设备、是否联网、统计方法与限制条件；不应脱离监测维度单独报出指标数值。

9.3 提及“情感倾向分布”“负面误引率”“风险内容处置率”等涉及负面或风险信息的指标，不应以消除真实负面信息为优化目标；对真实负面事实，应通过事实说明、补充信息、服务整改、公开说明、申诉核验或复测确认等合法方式处置，不应作为“压制”对象。风险内容处置应形成记录（见14.1）与可追溯证据要求（见12.2）。

9.4 指标统计应在报告中说明分母监测维度、有效观测数、无效观测剔除情况、人工复核比例、争议项处理规则与限制条件。涉及事实准确性、情感倾向、风险内容识别等需人工判断的指标，宜采用双人标注或独立复核；标注不一致的，应记录处理规则。监测方案应预先确定稳定性计算方法，不得在取得结果后更换计算方法。

10 监测技术手段与工具要求

10.1 技术手段要求

技术手段要求包括但不限于：

a) 监测宜优先采用符合目标平台规则、获得授权或公开允许的接口、导出能力或人工核验方式。采用人工监测等方式的，应评估并记录平台规则、账号状态、访问频次、数据安全、个人信息保护和反爬限制等合规风险，不应以规避访问控制、伪造身份、绕过反爬机制或污染生成式回答的方式获取数据；

b) 以API接口等方式监测的，应遵守目标平台的服务协议与访问规则，以及网络安全、数据安全、个人信息保护等法律法规要求；对于可能违反平台规则、引发数据合规风险或显著影响生成结果自然性的采集方式，不应作为正式效果证明依据。

10.2 工具要求

工具要求包括但不限于：

a) 用于GEO效果验证的监测工具应能够留存原始回答快照与可导出的原始数据流水。不应仅输出加工后的结论或静态截图；

b) 监测工具的指标定义（见9.1）与计算方法应公开、可解释（见附录B表B.1）；采用自定义指标的，应说明定义、计算方法与局限；

c) 监测数据精度、更新频率与覆盖平台应与监测目的相适应，并在交付物中说明。

本文件不指定、不推荐特定商业监测工具。使用任何工具均不免除监测验证主体对数据真实性、监测维度一致性与结论合理性的责任。

11 动态监测与预警

11.1 动态监测内容要求

动态监测内容要求包括但不限于：

a） 对重点领域、负面舆情敏感主体或竞争激烈场景，宜在周期性监测之外开展动态监测，持续跟踪可见性骤变、负面误引、引用信源失效、竞品异常提升与监测维度冲突等情况；

b） 动态监测应明确触发监测的事件类型、监测对象范围、采样加密规则与记录方式。

11.2 异常识别

动态监测宜对下列情形进行识别与标注：

a） 目标主体可见性在短期内显著下降，疑似出现非明示性可见度衰减；

b） 回答中出现无依据负面表述或事实性错误的扩散；

c） 引用信源大面积失效、被替换或指向不可信来源；

d） 竞品在相同意图短语下出现异常、非自然的集中提升；

e） 同一主体在不同平台或同一平台不同时间的监测维度出现严重冲突。

注：短期内大量同质低质页面出现、多站点复制同一错误事实、异常新域名集中指向负面结论、关键事实跨站点或跨语言统一失真、引用链中非官方来源异常升高等，可作为疑似语料投毒或异常操纵的识别线索；在形成行业共识前，不宜单独设置固定权重的“投毒风险指数”。

11.3 预警与处置要求

预警与处置要求包括但不限于：

a） 识别到异常时，监测方应及时形成预警记录，载明异常类型、发现时间、影响范围与初步研判，并按约定通知委托方；

b） 预警不应替代 GEO 效果验证结论；基于动态监测得出的趋势性判断，应说明其样本与监测维度限制；

c） 对不遵循 T/CAAAD 046—2026 第 10.2 条所列禁止行为（如虚假评价、误导性比较、信息投毒）导致的异常，监测方应如实记录并提示合规风险，不应参与、协助或掩盖相关行为。

12 GEO 效果验证

12.1 验证准备要求

验证准备要求包括但不限于：

a） GEO 效果验证应以完整、未经篡改的原始记录为依据，包括封存意图短语集、基线数据、各周期监测数据与回答快照。原始记录缺失或不可追溯的，不应出具肯定性验证结论；

b） GEO 效果验证应基于同监测维度的基线对比；无基线或基线监测维度不一致的，不应直接给出“提升”类结论；

c） GEO 效果验证以效果监测形成的封存意图短语集、基线数据、周期监测数据、回答快照和原始记录索引为核查对象，不重复监测的全部工作，但应包含独立抽样复核（见 12.3）。GEO 效果

验证可由委托方组织实施，也可委托具备相应能力的机构实施；验证结论用于对外证明的，实施主体应符合独立性与公正性的要求（见 15.2）。

12.2 验证内容要求

GEO 效果验证应至少核查下列事项：

a) 真实性：回答快照与监测数据是否真实、未经篡改，是否存在伪造截图、拼接或选择性删除。意图短语集规模、有效观测次数是否满足重复采样与时间窗口要求（见 8.3）；

b) 一致性：基线与各周期的意图短语集、平台范围、环境状态、统计方法、指标定义是否一致（见附录 B 表 B.1）；

c) 可追溯性：在可比条件下抽样重测，结果是否落在合理波动范围内；

d) 合规性：采样环境、技术手段（见 10.1）工具使用（见 10.2）是否符合要求，是否使用诱导性意图短语；效果结论是否与数据相符，是否存在夸大、绝对化或对第三方平台结果作出保证的表述。

12.3 抽样复核要求

抽样复核要求包括但不限于：验证应进行独立抽样复核，复核意图短语数量不宜低于封存意图短语集的 20%；涉及高风险、争议较大或对外证明用途的项目，宜提高复核比例。复核应在可比环境下重新监测并与原始记录比对；低于上述比例的，应说明原因、适用限制及对结论的影响。

复核结果与原始记录出现显著、无法合理解释的偏差的，应扩大复核范围；偏差源于数据造假的，应判定验证不通过并记录。典型监测验证造假行为与识别要点见附录 E。

12.4 验证结论要求

验证结论要求包括但不限于：

a) 验证结论应明确为“通过”“有条件通过”或“不通过”，并附判定依据与适用范围、监测维度及限制条件；

b) 存在下列情形之一的，不应判定为“通过”：

- 以单次查询、单张截图或诱导性意图短语所得回答作为主要效果证据；
- 伪造、篡改、选择性披露监测数据或回答快照；
- 基线缺失、监测维度不一致或样本量严重不足且无合理说明；
- 既承担优化实施又对外宣称独立验证，且未披露利益关系；
- 验证结论与原始数据明显不符或包含对第三方平台结果的保证性表述。

c) 验证结论应表述为规定监测维度与时间窗口下的观测事实，不应表述为对未来效果或第三方平台行为的保证。

13 交付物与验证报告

13.1 监测交付物要求

监测交付物要求包括但不限于：

a) 用于 GEO 效果验证或合同验收的监测交付物应包括监测方案、封存意图短语集、基线报告、周期监测报告与原始记录索引；监测报告应载明监测对象、平台范围、意图短语集规模、采样方法与频率、环境状态、指标结果、波动与异常说明及限制条件；

b) 仅用于服务过程跟踪、不作为 GEO 效果验证或合同验收依据的常规监测，其报告内容可由委托方与监测方约定从简；但封存意图短语集、原始记录和回答快照的形成与留存仍应符合要求（见 14.1）。

13.2 验证报告要求

验证报告要求包括但不限于：

a) 验证报告应按附录 D 的要求编制，至少包括：委托与被验证对象信息、验证依据、验证范围与监测维度、意图短语集与样本设置、基线与对比数据、验证内容核查结果、抽样复核情况、验证结论与限制条件、报告唯一性标识、出具日期与授权签字；

b) 验证报告应如实呈现不利结果与波动，不应只报有利数据；

c) 验证报告不应使用“保证”“控制”“永久”“唯一”或“官方内推”等误导性表述。

13.3 报告使用要求

报告使用要求包括但不限于：

a) 监测报告与验证报告应在其载明的监测维度、范围与时间窗口内使用，不应断章取义、超范围引用或用于与报告结论不符的宣传；

b) 对外引用验证结论的，应同时标明出具机构、报告编号、验证监测维度与时间窗口。

14 原始记录与数据留存要求

14.1 监测与验证的原始记录应真实、完整、即时生成，包括意图短语、查询标识符、查询时间、平台与版本标识、环境状态、回答快照、引用来源、统计明细、采集日志 ID、工具日志、记录文件哈希值与抽样复核记录（见附录 C）。

14.2 原始记录不应被覆盖、篡改或选择性删除；修改应留痕，标注修改人、时间与理由。

14.3 原始记录的留存期限自相应监测或验证活动完成之日起不应少于一年；用于合同验收、争议处理或对外披露的，留存期限不应少于相关事项了结后一年，合同或法律法规另有更长规定的从其规定。

14.4 涉及个人信息、商业秘密或委托方未公开经营信息的记录，应按 GB/T 35273 要求采取访问控制、加密存储与到期处置措施。

15 监测验证机构要求

15.1 基本要求

监测验证机构基本要求包括但不限于：

a) 监测验证机构应为依法登记、能够承担相应法律责任的法人或其他组织，其经营或业务范围应与所开展的监测验证活动相适应；

b) 机构应建立保证监测验证活动真实、客观、可复现、可追溯的管理体系，包括人员管理、方法管理、记录管理、报告管理与质量控制制度。

15.2 独立性与公正性要求

独立性与公正性要求包括但不限于：

a) 出具具有对外证明作用验证报告的机构，应保持独立、公正地位，机构及其人员与被验证项目的委托方、服务提供方、数据使用方不应存在影响公正性的利益关系；

b) 机构不应既承担同一项目的GEO优化实施，又对外出具该项目的独立验证报告；在同一法人组织内同时开展两类业务的，应建立隔离机制，识别并消除影响独立性与公正性的风险，并在报告中披露；

c) 机构不应参与语料投毒、组织虚假流量、操纵评论、伪造信源、诱导性监测等有损监测验证客观性的活动。

15.3 人员要求

人员要求包括但不限于：

a) 机构应配备与业务相适应的监测、分析与审核人员，关键岗位人员应熟悉生成式引擎机制、本文件及相关法律法规与方法；

b) 机构应设授权签字人，负责验证报告的审核与签发。授权签字人应熟悉本文件及相关技术规范，具备对监测验证结果作出评价与判断的能力，并经机构正式授权；

c) 机构应对人员开展培训与能力确认，并保存记录。

15.4 设施与数据安全要求

机构应具备开展监测验证所需的工具、环境与数据存储条件，对监测数据、原始记录和委托方信息采取保密和数据安全措施。

15.5 保密义务要求

机构及其人员对监测验证活动中知悉的商业秘密、个人信息及委托方未公开信息负有保密义务，未经授权或法律法规依据，不应向第三方披露。

附录 A
（资料性）
意图短语集设计参考框架

表 A.1 给出意图短语集设计的类型框架与示例方向，供设计封存意图短语集时参考。示例仅示意结构，应结合具体行业与场景替换为中性、自然的真实提问；在实际项目中可设置综合意图词，以反映真实用户将品牌、产品、场景、预算、问题等要素组合提问的情况。

表 A.1　意图短语类型参考框架

意图短语类型	意图说明	示例方向（中性表述）
品牌词	用户已知品牌，查询其信息	某品牌的主要产品与服务有哪些
产品词	查询具体产品特性	某类产品的常见功能与适用场景
品类词	未指定品牌的品类咨询	某品类有哪些值得了解的选择
场景词	结合使用场景求建议	在某场景下应如何选择某类产品
问题词	围绕具体问题求解答	使用某类产品时常见问题如何解决
比较词	多主体横向比较	某品类不同方案的差异与各自特点
决策词	强购买/选择意图	预算与需求给定时如何做出选择
售后词	售后、服务、口碑	某类产品的售后与使用反馈情况
风险词	风险、负面、注意事项	选择某类产品时需要注意哪些问题
竞品词	竞品视角的对照查询	某品类中有哪些主要参与者

注：风险词与竞品词用于客观反映真实表现，不应设计为诱导贬损竞品或预设褒贬的提问。

附 录 B
（资料性）
监测指标计算方法

各项监测指标应按表 B.1 规定的方法计算。统计前应剔除无效观测，并记录剔除规则、剔除数量及其对结论的影响。按回答或查询结果统计的指标，宜以有效观测数为分母；按意图短语覆盖统计的指标，以封存意图短语数为分母；按引用统计的指标，以可识别引用数为分母。品牌名、产品名以及“品牌名称与产品名称的组合”等组合在同一回答中同时出现的，宜按预先定义的统计对象分别计数；用于同一指标汇总时，不宜重复计入同一分母监测维度。

表 B.1 监测指标计算方法

指标	计算方法	说明
品牌提及率	含目标主体提及的有效观测数 ÷ 有效观测总数 × 100%	提及含品牌、产品、商标及可识别替代名称，按标注规则处理
目标信源出现率	含目标主体信源的有效观测数 ÷ 有效观测总数 × 100%	目标主体信源含目标主体官网、官方账号、权威页面、目标域名或约定来源名称
目标信源引用占比	目标主体相关引用数 ÷ 全部可识别引用数 × 100%	适用于带引用、来源或链接的回答；无可识别引用的观测应另行记录
首位/前三、前五推荐率	目标主体处于首位（或前三、前五）的可排序有效观测数 ÷ 可排序有效观测总数 × 100%	仅适用于推荐、比较、榜单等可识别排序场景
意图短语覆盖率	出现目标主体的封存意图短语数 ÷ 封存意图短语总数 × 100%	同一意图短语多次观测的，应按方案约定采用是否出现、均值或窗口汇总监测维度；未预先约定的，不得事后选择有利监测维度
引用溯源有效率	可定位真实有效且内容相关的引用数 ÷ 全部可识别引用数 × 100%	失效、伪造、无法定位、与回答内容无关或指向不可信来源的引用不计入有效
回答准确率	准确信息项数 ÷ 被评价信息项总数 × 100%	信息项、权威依据、时效范围与争议项处理规则应预先定义；宜采用复核机制
事实错误率	错误信息项数 ÷ 被评价信息项总数 × 100%	含虚假、过期、超范围、缺必要条件、与权威依据不一致等
负面误引率	出现无依据或错误负面表述的有效观测数 ÷ 有效观测总数 × 100%	区分真实负面与无依据负面，不应将真实负面列为删除或压制对象
权威信源占比	权威信源引用数 ÷ 全部可识别引用数 × 100%	权威信源分级按监测方案或 T/CAAAD 046 生成式引擎优化服务要求附录约定执行，并在报告中披露

表 B.1　监测指标计算方法（续）

指标	计算方法	说明
跨时间稳定性	同监测维度各时间窗口指标的标准差、极差、极差率或变异系数	应注明窗口数、每窗口样本量与计算式；不同监测维度或重大版本变更前后不直接比较
证据完整率	具备完整证据链的信息项数或有效观测数 ÷ 对应总数 × 100%	证据链含来源、采集、审核、记录、快照编号与必要元数据
风险内容识别率	已完成风险识别与分级标注的风险线索数 ÷ 应识别风险线索总数 × 100%	风险线索来源、识别规则、分级标准与复核情况应留痕
风险内容处置率	已完成闭环风险内容处置并经复核确认数 ÷ 已识别风险内容总数 × 100%	处置含事实更正、补充披露、服务整改、平台申诉、下架或更正、复测确认、记录为真实负面并完成说明等
有效观测率	有效观测数 ÷ 计划观测数或实际采集观测数 × 100%	应说明无效观测剔除规则、数量与原因
环境一致率	满足约定环境状态的有效观测数 ÷ 有效观测总数 × 100%	环境状态包括登录、缓存、历史对话、联网检索、地域、设备、时间段等
抽样复核通过率	复核通过样本数 ÷ 抽样复核样本数 × 100%	复核通过指原始记录真实、一致或偏差处于可解释范围内
原始记录完整率	要素完整的原始记录数 ÷ 原始记录总数 × 100%	完整要素至少包括意图短语、时间、平台入口、环境状态、回答快照、引用来源、统计明细、人员或工具标识

注：加权类综合指标（如对不同意图短语按意图热度加权的可见度得分）可作为可选指标，其权重设置依据应在监测方案中约定并在报告中披露。

表 B.2　监测频率参考

监测阶段	建议频率	说明
基线阶段	每日或隔日采集	建立稳定基线，排除短期波动
优化生效观察阶段	每周不少于 1 次，持续不少于 4 周	观察优化动作的收录与引用变化
常态监测阶段	每月不少于 1 次	跟踪长期可见性与稳定性
重点/高风险监测	每周不少于 2 次；出现负面舆情、平台重大波动或竞争异常时宜每日监测或启用动态监测	适用于重点领域、负面舆情或竞争激烈场景

表 B.2　监测频率参考（续）

监测阶段	建议频率	说明
平台重大变更时	变更后重新建立基线并加密采样	模型或产品版本重大更新后，数据不与变更前直接比较

附录 C
（资料性）
GEO 效果监测原始记录表示例

每次有效观测宜按表 C.1 逐条记录，记录宜即时生成并妥善留存。记录表宜包含查询标识符、采集日志 ID 和记录文件哈希值等可追溯字段，以及监测方式。

表 C.1　GEO 效果监测原始记录表

序号	查询标识符/意图短语/类别	查询时间	平台/入口/版本	环境状态	是否提及/位次	引用来源	快照编号/日志 ID/哈希值/记录人
1	查询标识符/意图短语/类别	（待填）	（待填）	登录/缓存/联网/地域	（待填）	（待填）	快照 ID/日志 ID/哈希值/记录人
2							
3							

注：环境状态栏宜记录是否登录、是否清除历史与缓存、是否联网检索、地域与设备类型；快照编号应能关联到完整回答快照文件；日志 ID 宜能关联采集日志或工具日志；哈希值用于校验原始记录、回答快照或导出数据包的一致性。

附录 D
（资料性）
GEO 效果验证报告编制

验证报告要素

a） 报告基本信息：报告名称、唯一性编号、每页标识与结束标识、出具机构、出具日期；

b） 委托对象信息：委托方、被验证服务提供方、被验证 GEO 项目及目标主体；

c） 验证依据：本文件以及约定的监测方案与封存意图短语集；

d） 验证范围与监测维度：目标引擎范围、意图短语集规模、时间窗口、采样次数、环境状态、统计方法、指标定义；

e） 基线与对比数据：基线水平、各周期数据及同监测维度对比结果；

f） 验证内容核查结果：对第 12.2 条各项的核查情况；

g） 抽样复核情况：复核意图短语数量、复核环境、比对结果与偏差说明；

h） 验证结论：通过/有条件通过/不通过及其判定依据；

i） 限制条件与声明：监测维度与时间窗口限制，以及不构成对第三方平台结果保证的声明；

j） 授权签字与机构信息：授权签字人签字、机构信息与联系方式；存在分包的，注明分包项目与承担机构。

附录 E
（资料性）
典型监测验证造假行为与识别要点

表 E.1 给出监测验证环节常见的造假与不规范行为及其识别要点，供监督、核查与自查使用。

表 E.1　典型造假行为与识别要点

行为类型	典型表现	识别要点
伪造截图	以单张或拼接截图冒充效果；裁剪、修改回答内容	要求提供完整回答快照与原始流水；核对时间戳与元数据；抽样重测
诱导性提问	以预设结论、限定褒义的意图短语生成赞美式回答作为凭证	审查意图短语中性度；用中性意图短语复测对比
样本操纵	仅监测友好问题、隐瞒样本量、剔除不利样本	核对封存意图短语集与实际监测覆盖；检查剔除规则与有效观测数
监测维度漂移	不同周期变更平台、环境、统计监测维度制造“提升”假象	比对各周期监测维度记录；核查基线一致性
数据篡改	修改指标数值、伪造监测日志、覆盖原始记录	检查记录留痕与修改日志；复算指标；抽样重测
环境操纵	利用登录态、历史对话、个性化缓存制造有利回答	核查环境状态记录；在纯净监测环境复测
独立性缺失	既做优化又自评自证，而宣称独立第三方验证	核查机构与项目的利益关系与隔离机制；查披露情况
结论夸大	数据与结论不符；使用保证性、绝对化表述	比对结论与原始数据；核查限制条件是否披露

参 考 文 献

[1] GB/T 27025 检测和校准实验室能力的通用要求

[2] T/CAAAD 046—2026 生成式引擎优化服务要求

[3] T/CAAAD 048—2026 生成式引擎优化内容源主体分类分级与选用要求

[4] 中华人民共和国数据安全法

[5] 中华人民共和国反不正当竞争法

[6] 中华人民共和国消费者权益保护法

[7] 中华人民共和国消费者权益保护法实施条例

[8] 中华人民共和国计量法

[9] 网络数据安全管理条例

[10] 中华人民共和国广告法

[11] 中华人民共和国认证认可条例

[12] 检验检测机构资质认定管理办法

[13] 检验检测机构资质认定评审准则

[14] 互联网广告管理办法

[15] 生成式人工智能服务管理暂行办法

[16] 人工智能生成合成内容标识办法

[17] Aggarwal P, et al. GEO: Generative Engine Optimization. Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD). 2024

中国广告协会

团体标准

生成式引擎优化

效果监测与验证要求

T/CAAAD 047—2026

※

中国广告协会 编印

北京市朝阳区光华路 9 号时尚大厦 8 层（100020）

电话：010-65924090

网址：http://www.china-caa.org

邮箱：digitalad@china-caa.org

版权专有　　侵权必究