AI 搜索可见度测量与同条件复测方法白皮书

灰犀AI公开说明品牌事实、受控问题、API 测量来源、任务分母、原始证据、质量复核、指标及同条件复测的适用范围。

生效 2026-09-21 发布 2026-09-21 更新 2026-09-21 复核 2026-09-21

修订 1.0.0-r3 SHA-256 5d68b39280abbdc7fee9f9fb4937bdea26bb150623ccc8ca51a9cd011c8e699d

适用范围：中国大陆、简体中文；当前测量来源简称为豆包与百度 AI 搜索，完整 API 名称及适用限制见测量来源章节。

测量来源：豆包[1]、百度 AI 搜索[2]。

已上线并经过生产验证：结构化建档、测量简报、受控问题方案和旧项目兼容已通过生产流程验证。

代码已上线，尚无真实 B1 测量案例：质量分析、人工复核与商户报告可解释方法，但不能宣称已有真实 B1 案例结果。

当前关闭：项目资料 AI 助手真实调用尚未开放；手工建档仍可使用。

合成教学案例：本白皮书中的逐任务样例专用于计算说明，不是客户或生产测量。

目录

01 摘要：测什么，不测什么

02 适用范围与限制

03 商户事实档案与测量简报

04 受控问题方案

05 测量来源与 API 接口

06 预期任务与完成状态

07 原始证据、分析与行动

08 指标字典

09 质量检查与人工确认

10 版本、可比性与新基线

11 同条件复测与非因果限制

12 合成教学案例

13 数据、隐私与方法治理

14 方法版本与变更记录

01 摘要：测什么，不测什么

先确定测量对象，再读单次回答证据。

灰犀AI记录明确品牌、问题、地区、语言、测量来源和版本条件下的回答、品牌与竞品提及及引用。

一次 API 回答不是所有消费者端 AI 产品的固定排名，也不保证未来返回相同答案。

Google 说明其搜索 AI 功能可能使用不同模型和方法，回答与相关链接会变化；满足收录条件也不保证页面被索引或展示。此事实仅说明外部系统的不确定性，不表示灰犀AI当前测量 Google。 来源：[google-ai-features]

02 适用范围与限制

只解释已明确披露的测量接口。

当前公开方法只覆盖中国大陆、简体中文，以及豆包和百度 AI 搜索两个指定 API 测量来源。

API 接口不等于豆包、文心等消费者端产品体验；不能从有限问题推算全网市场份额。

03 商户事实档案与测量简报

事实回答测量谁，简报回答为什么测。

事实档案区分已确认、建议和未知；只有已确认、公开、来源可解析且获准用于外部测量的事实能进入问题。

简报固定业务范围、受众、场景、目标及竞品状态；AI 助手建议不能自身成为商户事实依据。

04 受控问题方案

四个基础问题，确认竞品后增加一个比较问题。

新项目依次覆盖无品牌发现、场景推荐、品牌理解、事实与来源四类意图；确认竞品后增加竞品比较。最终确认的问题文字、题序和事实引用形成版本。

无品牌发现题不植入被测品牌；竞品未知或本次暂无竞品时，不伪造比较问题。

旧项目可能沿用历史问题方案；不能把新项目的四题规则反写为旧项目的历史运行事实。

05 测量来源与 API 接口

逐次披露提供商、模型、接口和时间。

每次运行保存测量来源、模型、API 接口、搜索配置、市场、语言和时间。正文简称为豆包[1]与百度 AI 搜索[2]，两个来源的回答不当作同一个回答。

来源升级或接口变化后须判断可比性；不能将 API 返回结果称为消费者端固定排名。

06 预期任务与完成状态

问题乘来源形成运行前的预期任务。

四个基础问题乘两个来源形成八个预期回答任务。每个任务有完成、失败或缺失状态；不能删除失败任务缩小分母。

[合成教学案例] 下方教学案例有八个预期任务：五个完成、两个失败、一个缺失，因此完成覆盖为 5/8。

07 原始证据、分析与行动

先保存回答，再解释判断，最后提出下一步。

事实层保存问题、来源、时间、状态、原始回答与引用；分析层处理相关性、实体与来源类别；行动层必须回链到已确认的证据。

没有已完成且可定位的证据时，不生成看似确定的优化建议。原始回答不因人工复核而改写。

08 指标字典

每个指标标明统计单位、分子、分母和未知项。

品牌发现和官网引用只把相关且能判断该维度的已完成回答放进分母；无法判断不是未命中。

竞品领先按问题聚合，不能把同一题的多个测量来源扩张为多个问题；来源结构按有效引用链接而非回答任务统计。

完成覆盖

业务问题：预期回答任务有多少完成？ 单位：回答任务

分子：已完成任务数。分母：测量前固定的预期任务数。

无法判断与排除：失败与缺失留在分母；不以已返回的证据条数缩小分母。

跨来源：每个问题与每个测量来源构成独立任务，来源之间不互相补齐。

跨版本：问题、来源或运行范围改变后建立新基线，不拼接完成覆盖趋势。

误读：完成不等于回答相关或可靠。

品牌发现

业务问题：可判断的回答里，品牌出现多少次？ 单位：回答任务

分子：相关且品牌确认命中的回答任务数。分母：完成、相关且品牌命中可判断的回答任务数。

无法判断与排除：另列完成总体、无法判断和排除数；unknown 不算未命中。

跨来源：逐来源判断并分别计入回答任务，不把两个来源合并成共识。

跨版本：实体档案或别名规则改变可重分析旧证据；问题或来源改变则不直接比较。

误读：不是所有消费者 AI 产品中的品牌排名。

官网引用覆盖

业务问题：可判断的回答里，已确认官网被引用多少次？ 单位：回答任务

分子：相关且确认包含官网引用的回答任务数。分母：完成、相关且官网引用可判断的回答任务数。

无法判断与排除：已确认主域和子域按同一规则；unknown 不算未引用。

跨来源：逐来源判断官网归属；同一 URL 在不同回答中仍属于不同回答任务。

跨版本：官网归属规则修正可重分析旧引用并保留旧版本；接口变化建立新基线。

误读：1/N 不能写成从未引用。

竞品领先问题

业务问题：适用的比较问题中竞品领先多少题？ 单位：问题

分子：已判定为竞品领先的问题数。分母：已判定方向的适用比较问题数。

无法判断与排除：方向包括品牌领先、竞品领先、无明确领先；无法判断单列，排除题单列。

跨来源：同一比较题先汇总来源证据，再由受控确认得到一个问题级方向；来源冲突时不得自动判领先。

跨版本：竞品或比较题变化会改变适用问题人口，必须建立新基线。

误读：不可把两来源的一题算作两题。

引用来源结构

业务问题：有效引用链接分别来自哪里？ 单位：引用链接

分子：按官网、权威媒体、行业媒体、用户内容、营销列表及 unknown 分类计数。分母：有效回答中的有效引用链接总数。

无法判断与排除：相关性 unknown 回答中的有效链接保留在 unknown 类别；无效或跑题回答的链接计入 excluded。

跨来源：按每条有效引用链接统计；不同来源返回相同 URL 时仍保留各自证据关系。

跨版本：来源分类规则变化须发布新分析版本；测量来源口径变化须建立新基线。

误读：链接数不等于引用它的回答任务数。

品牌描述观察

业务问题：回答如何描述品牌？ 单位：单次观察

分子：能定位到品牌命中位置的原句或受限上下文数。分母：不生成比例分母；每条观察独立展示。

无法判断与排除：无法定位原句时不生成；不同来源的描述不自动合并为共识。

跨来源：每个回答独立保存原句、问题与来源引用，不自动总结跨来源共识。

跨版本：重分析可产生新描述版本并保留旧观察；跨问题或来源版本不合并。

误读：观察文本不是平台对品牌的事实认证。

09 质量检查与人工确认

自动检查可报告性，歧义交由人工确认。

实体识别、标题证据一致、业务范围、竞品判断、品牌认知和来源结构都需要质量检查；结果区分可交付、需要确认与需要修正。

人工无法判断时保留 unknown；自动语义线索不能单独把回答判为错误，更不能把未知改写为否。

10 版本、可比性与新基线

方法版本不同于项目问题、来源和分析版本。

别名或官网归属修正可对旧证据重新分析并保留旧分析；实体、范围、问题文字或题序、模型或 API 接口实质改变时，建立新基线而不拼接连续趋势。

排版、错字和来源链接修正可更新 v1 并记录内容修订；指标公式、分母、适用范围、问题意图、质量规则或测量来源口径改变，必须发布 v2，不得覆盖 v1 方法含义。

现有历史报告尚未保存生成时的方法版本；其方法入口不代表报告快照已经绑定 v1。

11 同条件复测与非因果限制

比较可观察变化，不宣称效果归因。

复测逐字使用原问题与顺序、市场、语言及来源版本，分别比较完成覆盖、品牌提及、引用和竞品观察。

模型随机性、时间与外部环境均可能改变回答；一次前后变化不能证明某个行动造成变化。

12 合成教学案例

从商户事实、问题和逐任务证据走到质量确认、行动与同条件复测。

[合成教学案例] 这是虚构的教学数据，不来自客户、生产测量或真实 AI 调用。甲、乙仅代表合成测量来源。

[合成教学案例] 五个完成回答中，三个相关且品牌与官网维度均可判断，一个相关性无法判断，一个跑题。相关性未知回答的有效链接保留在来源 unknown 类别，只有跑题或无效链接被排除。

[合成教学案例] 本案例未配置竞品，因此只有四个基础问题；若后续确认竞品，新增比较题及其两个来源任务应形成新的问题方案和基线。

澄川智造：从建档到同条件复测

澄川智造、所有回答、URL 和结果均为合成教学数据，不对应真实公司或客户。

已确认商户事实

标准实体名：澄川智造

官网：https://chengchuan.example/

服务：工业设备状态分析与异常周报

受众与场景：中国大陆小批量制造团队；评估设备异常分析服务

测量目标

1. 判断无品牌发现问题中是否出现品牌
2. 核对出现品牌时如何描述及是否引用官网
3. 保存失败、缺失和无法判断，不缩小任务分母

实际问题方案

问题 1 · 无品牌发现：中国大陆的小批量制造团队评估设备异常分析服务时，可以关注哪些服务商？

设置原因：不植入被测品牌，观察自然发现。

问题 2 · 场景推荐：哪些服务商适合为小批量制造团队提供工业设备异常分析和周报？

设置原因：固定受众、服务和场景。

问题 3 · 品牌理解：澄川智造主要为哪些客户解决什么问题？

设置原因：核对品牌如何被描述。

问题 4 · 事实与来源：澄川智造提供哪些工业设备分析服务？请给出可核验来源。

设置原因：核对事实表达和引用结构。

四题 × 两来源：原始证据与独立判断

问题 1 · 来源甲 · 已完成 · 相关性：相关 · 品牌提及：是 · 官网引用：是 · 竞品判断：不适用

回答节选：可关注澄川智造，其面向小批量制造团队提供设备状态分析服务。

判断：回答提及合成品牌并引用其已确认官网。 品牌描述观察：澄川智造面向小批量制造团队提供设备状态分析服务。 来源：https://chengchuan.example/about (已确认官网)

问题 1 · 来源乙 · 已完成 · 相关性：相关 · 品牌提及：否 · 官网引用：否 · 竞品判断：不适用

回答节选：可先比较设备联网范围、告警响应和本地交付能力。

判断：回答未提及合成品牌；外部引用尚未分类。 来源：https://source.example/notes (未分类)

问题 2 · 来源甲 · 已完成 · 相关性：相关 · 品牌提及：是 · 官网引用：否 · 竞品判断：不适用

回答节选：澄川智造提供面向离散制造现场的设备异常分析与周报。

判断：回答提及合成品牌并引用行业内容，未引用官网。 品牌描述观察：澄川智造提供面向离散制造现场的设备异常分析与周报。 来源：https://industry.example/article (行业媒体)

问题 2 · 来源乙 · 已完成 · 相关性：无法判断 · 品牌提及：无法判断 · 官网引用：无法判断 · 竞品判断：不适用

回答节选：该讨论主要介绍通用设备联网经验，是否覆盖本题场景无法判断。

判断：是否回答了本题尚无法判断；其中有效链接保留为来源 unknown。 来源：https://community.example/thread (未分类)

问题 3 · 来源甲 · 已完成 · 相关性：跑题 · 品牌提及：无法判断 · 官网引用：无法判断 · 竞品判断：不适用

回答节选：文章讨论消费电子供应链，与本题的工业设备分析场景无关。

判断：跑题回答，其链接从有效来源统计中排除。 来源：https://unrelated.example/page (权威媒体)

问题 3 · 来源乙 · 失败 · 相关性：不适用 · 品牌提及：不适用 · 官网引用：不适用 · 竞品判断：不适用

判断：测量来源调用失败。

问题 4 · 来源甲 · 失败 · 相关性：不适用 · 品牌提及：不适用 · 官网引用：不适用 · 竞品判断：不适用

判断：测量来源调用失败。

问题 4 · 来源乙 · 缺失 · 相关性：不适用 · 品牌提及：不适用 · 官网引用：不适用 · 竞品判断：不适用

判断：预期任务未获得回答。

完成覆盖 5/8；失败 2，缺失 1。

品牌发现 2/3；无法判断 1，排除 1。

官网引用 1/3；无法判断 1，排除 1。

有效引用 4 条，其中官网 1 条、unknown 2 条；排除 1 条。品牌描述观察 2 条。竞品比较不适用。

质量确认与有限制交付

问题 2 · 来源乙：决定“无法判断”。相关性、品牌与官网判断继续保留 unknown，不改写为否。

source.example：决定“无法判断”。有效链接保留在来源 unknown 类别。

8 个预期任务中 3 个未完成：决定“有限制交付”。报告明确 5/8，仅对已完成且可判断人口给出指标。

证据型行动

补强官网对工业设备异常分析与周报的公开说明。证据任务：q2-甲。问题 2 来源甲提及品牌但没有官网引用；行动只回链这一条已完成证据。

同条件复测

· 逐字复用四个问题与题序

· 继续使用来源甲、来源乙及相同接口版本

· 地区、语言、品牌事实档案和官网归属规则不变

q1-甲 · 已完成 · 品牌 是 · 官网 是

q1-乙 · 已完成 · 品牌 否 · 官网 否

q2-甲 · 已完成 · 品牌 是 · 官网 是

q2-乙 · 已完成 · 品牌 无法判断 · 官网 无法判断

q3-甲 · 已完成 · 品牌 是 · 官网 否

q3-乙 · 已完成 · 品牌 否 · 官网 否

q4-甲 · 失败 · 品牌 不适用 · 官网 不适用

q4-乙 · 缺失 · 品牌 不适用 · 官网 不适用

复测完成覆盖从 5/8 变为 6/8；问题 2 来源甲新增官网引用。这里只描述同条件下观察到的变化，不声称行动造成变化。

13 数据、隐私与方法治理

公开方法不替代项目权限与隐私合同。

品牌与竞品使用同一判断规则；套餐、机构身份和客户关系不改变测量结果。公开方法不披露私密项目事实或跨租户数据。

不提供黑箱综合总分、机构认证、固定排名或推荐保证；具体数据保存与删除以现行隐私政策和产品合同为准。

14 方法版本与变更记录

公开修订内容和可比性影响。

v1 是当前公开方法版本；内容修订号及指纹使网页与 PDF 可以核对是否出自同一修订。方法含义变化必须新建 v2。

2026-09-21：发布 v1 内容修订 r3：增加测量来源简称、完整 API 名称及消费者端体验限制说明。 方法含义不变。

2026-09-21：发布 v1 内容修订 r2：补齐端到端合成案例、B1-5 指标解释和 PDF 分页门禁。 方法含义不变。

2026-09-21：发布 v1 内容修订 r1：公开范围、指标字典、逐任务合成案例与版本升级规则。 方法含义不变。

来源与审核

测量来源注释

[1] 豆包：火山方舟 API 联网搜索（豆包模型）。API 返回结果不等于豆包消费者端产品体验或固定排名。

[2] 百度 AI 搜索：百度千帆智能搜索生成 API。API 返回结果不等于百度消费者端产品体验或固定排名。

发布：灰犀AI。署名：灰犀AI 方法团队。内容责任：产品方法负责人。审核：事实与方法审核负责人。复核日期：2026-09-21。

[google-ai-features] Google Search Central：《AI features and your website》。https://developers.google.com/search/docs/appearance/ai-features 访问日期 2026-09-21。

[grayrhino-experiments] 灰犀AI：《灰犀AI 实验记录》。https://huixilab.com/zh/experiments 访问日期 2026-09-21。

本方法不代表消费者端 AI 产品的固定排名，不是认证或评级，不保证品牌被推荐、引用或获得商业结果。单次复测只能描述观察到的变化，不能自动证明因果。