被 AI 读懂，
才能被世界看见

基于 ChatGPT、Gemini、Google AI Mode、AI Overviews、Copilot
共 1,004.6 万条 AI 回答的实证洞察

姚金刚
www.laoyao.cn

AI 回答里的五个"入口"

- ChatGPT
- Gemini
- Google AI Mode
- Google AI Overviews
- Copilot

第二届GEO大会

深圳

活动举办城市
这次选择在了深圳

11.21 – 11.22

活动时间是11.21日到11.22日
两个全天

400+

预计超过400名会员参加大会，
同时超过40位嘉宾进行GEO与AI
方向的高密度干货分享

399

两天门票399，同时收获两本
签名书籍，以及相关的会员权
益

报名网址：hui.ailingdaoli.com

趋势 · 入口变了

AI 摘要减少点击，品牌影响发生在答案之中

8%

出现 AI 摘要时，用户点击搜索结果的比例；没有摘要时为 15%

Pew Research，2025

1%

用户直接点击 AI 摘要里来源链接的比例

Pew Research，2025

+693%

2025 美国假日季，生成式 AI 带到零售网站的流量同比增长

Adobe Analytics，2026

4.4×

AI 搜索访客的预计价值
相对自然搜索，按转化模型

Semrush，2025

AI 时代的品牌可见度，发生在被准确推荐和引用的答案里

来源：Pew Research Center（2025.7）；Adobe（2026.1）；Semrush（2025.6）

我们观察了什么

用一千万条真实 AI 回答，看 AI 怎么挑品牌

1,004.6 万条AI回答

5 个平台

ChatGPT · Gemini · AI Mode · AI Overviews · Copilot

5,756 万

条非官网来源链接，看 AI 从哪里取材

17 周

同一问题连续追踪，看答案怎么变

7,030 对

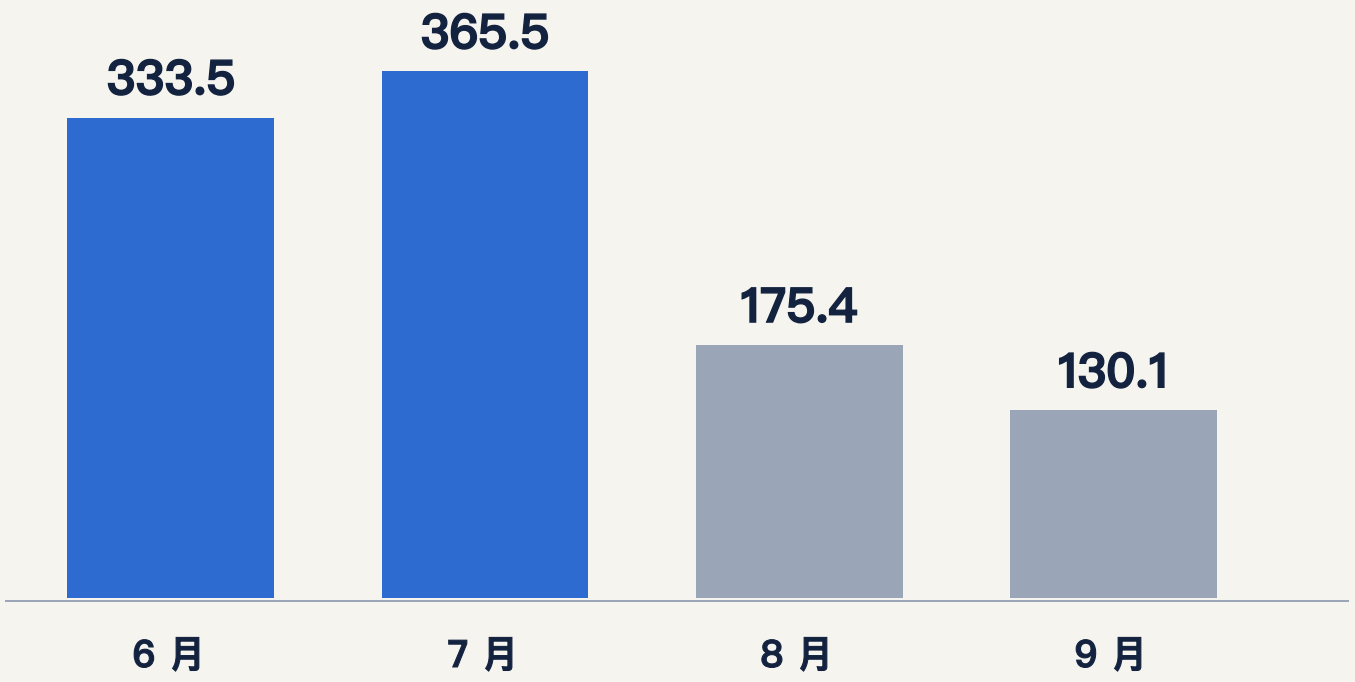
同日重测，看一次测量有多可信

数据：《GEO 白皮书 2026》实证专题，2026 年 6–9 月，美国、西班牙、印尼、新加坡、泰国等市场；观察性数据，不代表因果

研究边界 · 样本与口径

样本变化时，趋势判断优先使用比率

各月可分析回答数（万条）



44%

8–9 月 / 6–7 月样本量

对企业意味着什么

样本量变化会影响绝对次数
也可能改变题库构成

固定核心题库
用比率与重复测量判断变化

数据：白皮书第三部分图02；各月题库非固定，跨月总量不可直接比较

基础概念 · 指标定义

推荐、来源关联、引用与推荐分别计量

推荐

正文中检出品牌

正文提到商业品牌名
还需核对语境与准确性

来源关联

列表中出现链接

来源列表出现某个 URL
研究跨平台采用这一口径

引用

链接支撑具体表述

核对来源能否支撑事实
同时记录引用支持度

推荐

明确的正向选择

结合用户问题判断
品牌是否被作为适合选项

同一回答可能提到品牌、没有链接；指标分别计算，才能看清变化

口径：来源关联看来源列表，引用看正文支撑，推荐还需判断语义

第一性原理

AI 品牌可见度来自检索与既有认知

路径一：检索

先搜再答：你的页面被**找到**，再被**选中**写进答案，并附上出处。

主场：ChatGPT · AI Overviews · Copilot

商业品牌推荐比例

88.0%

有来源的回答

88.4%

无来源的回答

路径二：既有认知

品牌信息可来自参数知识与实体资料；没有链接，并不足以判断是否检索。

主场：Gemini · Google AI Mode

可见度的教学框架：页面可检索，事实可选用，品牌认知一致

数据：GEO 白皮书 T3，五平台 1,004.6 万条回答；有来源 785 万条、无来源 219 万条

第一性原理 · 生成链路

AI 如何选内容：七个环节逐项检查

01 进入候选

抓取与索引
确认页面能读取、能索引

检索触发
问题是否进入联网路径

查询扇出与召回
覆盖实际子问题

02 进入答案

重排与上下文分配
保留主题相关的事实与证据

生成与引用
提供可核验的结论与出处

吸收与保真
检查价格、名称和规格是否正确

03 进入行动

用户行为
用户是否继续搜索、访问或询盘

观察重点
将品牌露出与业务结果串联

排查顺序
先查丢在哪一环，再决定改什么

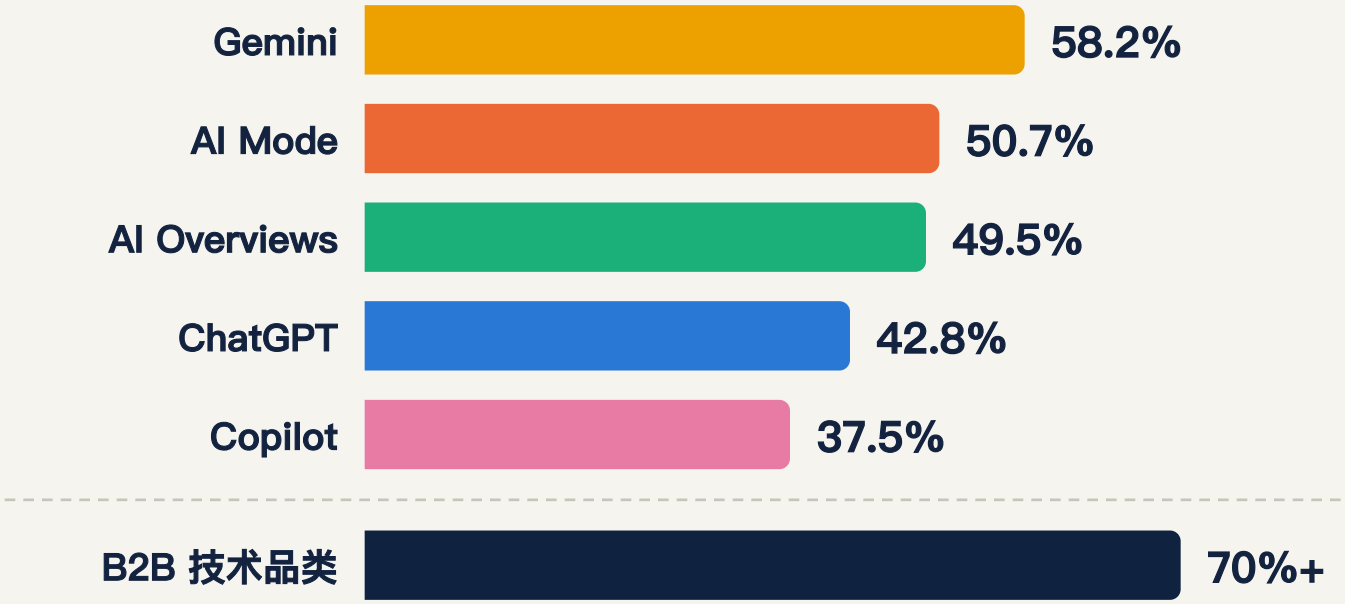
先检查页面是否进入候选，再检查模型选用了什么、如何表述

来源：白皮书第一部分七环节模型；分组仅用于课堂讲解

洞察 01 · 反常识

品牌官网是五个平台共同的第一大信源

企业/品牌官网在 AI 来源链接中的占比



对企业意味着什么

官网承载品牌事实底座：价格、规格、政策，应有准确的原始出处。

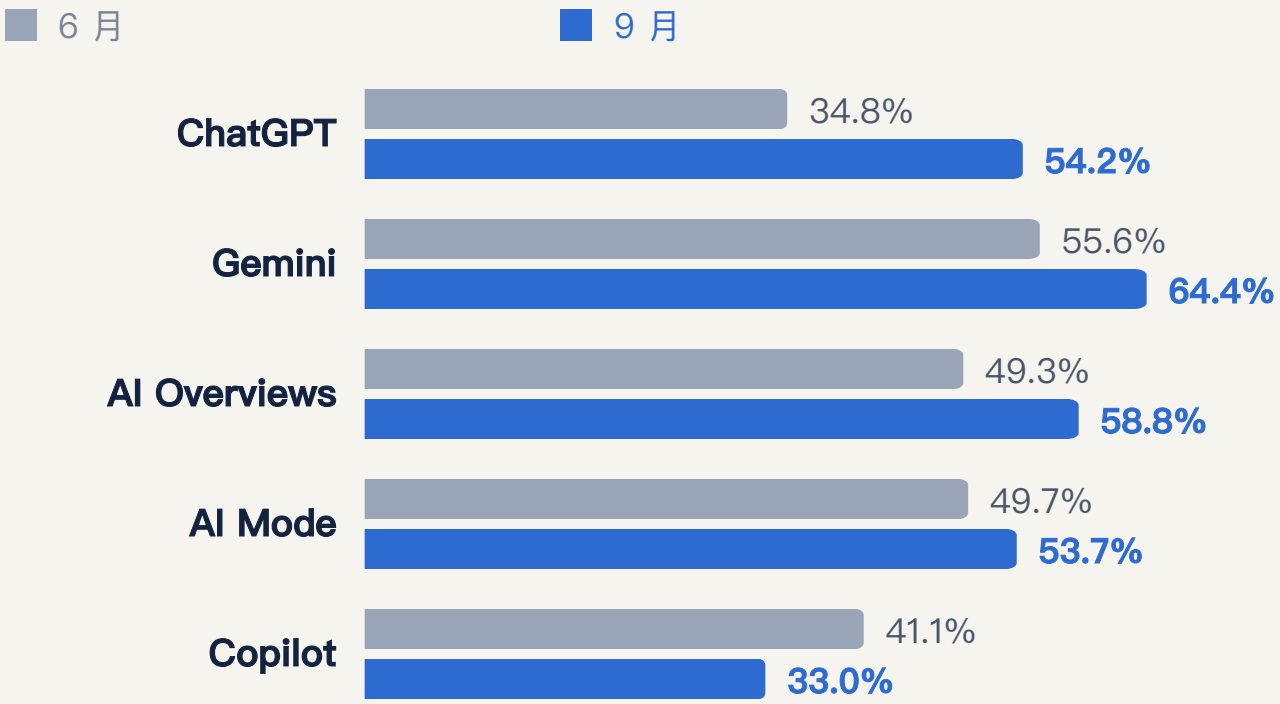
趋势还在加强：ChatGPT 的官网占比 4 个月从 **34.8% 升到 54.2%**。

动作：一个事实一个权威页面，写在 HTML 正文里，配齐结构化数据，确认 AI 爬虫能抓到。

数据：GEO 白皮书 T1/T2，2026 年 6–9 月；"官网"指品类内所有企业官网；B2B 为 AI 网关、网页采集 API、全渠道客户沟通三个品类

深入 01 · 官网趋势

官网引用占比上升



对企业意味着什么

ChatGPT 上升 19.4 个百分点
官网事实值得持续经营

用固定题库复核变化
保留月份和采集模式

优先完善价格、规格
以及兼容性与售后政策

数据：白皮书第三部分图06；6月与9月比较，样本构成可能影响比率

行动方法 · 官网建设

官网事实底座：五类页面优先补齐

一个事实有一个权威出处，官网正文、商品数据与站外表述保持一致

产品参数

规格与版本

尺寸、功率与型号
范围、单位与版本

价格套餐

金额与条件

币种、税费与周期
适用地区与生效日期

兼容边界

支持与限制

设备、系统与接口
地域限制和使用前提

交付政策

服务与售后

发货、退换与保修
取消规则和服务响应

案例证据

过程与结果

真实案例、测试方法
数据出处与更新记录

验收：正文可读取、事实可核对、URL稳定、更新有日期

方法：白皮书“五可”框架与T1/T2行动建议；效果需企业实测

洞察 02 · 反常识

五个平台呈现不同的信源偏好

≈ 20%

ChatGPT

偏好

机构学术

NIH、arXiv 进前十；前 10 域名只占 11%，长尾最分散

AI Mode

偏好

视频媒体

YouTube 一家占 17.6%，Reddit 7.8%

AI Overviews

偏好

用户社区

YouTube 14.4%，Reddit 9.6%；社区占比五平台最高

Gemini

偏好

品牌官网

每个回答平均只挂 1.8 个来源，惜字如金

Copilot

偏好

电商 ×3.8

Amazon 9.0%，Walmart 5.3%

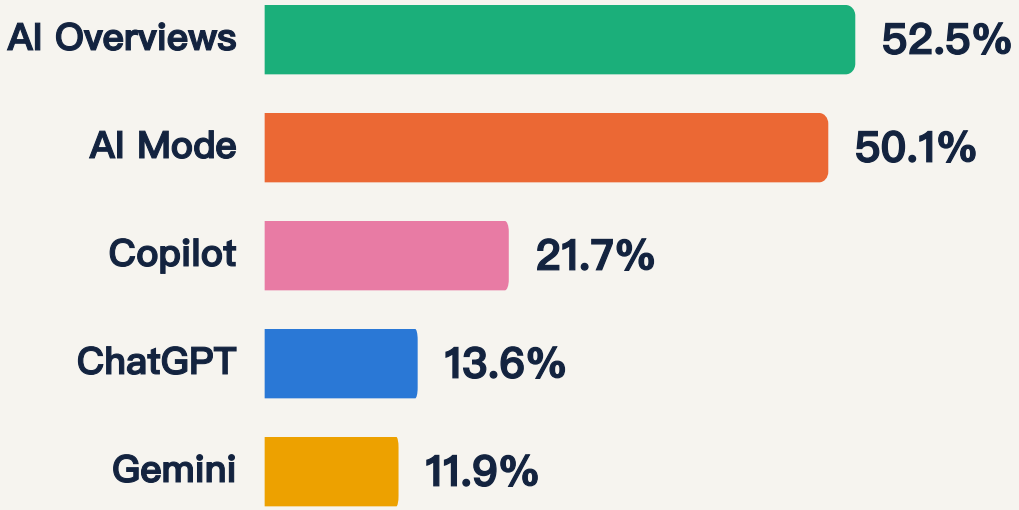
对企业：共享官网与品牌事实，按平台配置站外渠道；YouTube、Reddit、Amazon 是重要的跨平台信源。

数据：GEO 白皮书 T1，非官网来源域名；"×3.8"为该类信源占比相对五平台均值的倍数

深入 02 · 集中度

Google 搜索入口的头部来源更集中

前 10 大非官网域名占比（CR10）



对企业意味着什么

两个搜索入口中，约一半来源
流向排名前十的域名

ChatGPT 来源更分散
专业长尾内容有覆盖空间

分别配置头部合作
与垂直站点覆盖

数据：白皮书第三部分图07；2026年9月，CR10分母为非官网来源

行动方法 · 站外建设

按平台信源画像，安排站外内容与合作

01 视频与社区

适用观察
Google 入口的 YouTube、Reddit

建设动作
回答具体问题，呈现真实体验

验收
品牌被准确提及，事实有出处

02 专业与长尾

适用观察
ChatGPT 的机构、评测与垂直站

建设动作
提供研究、测试方法与可核验资料

验收
进入本品类实际被用到的来源

03 零售与商品

适用观察
Copilot 的电商与零售信息

建设动作
补齐商品参数、政策与真实问答

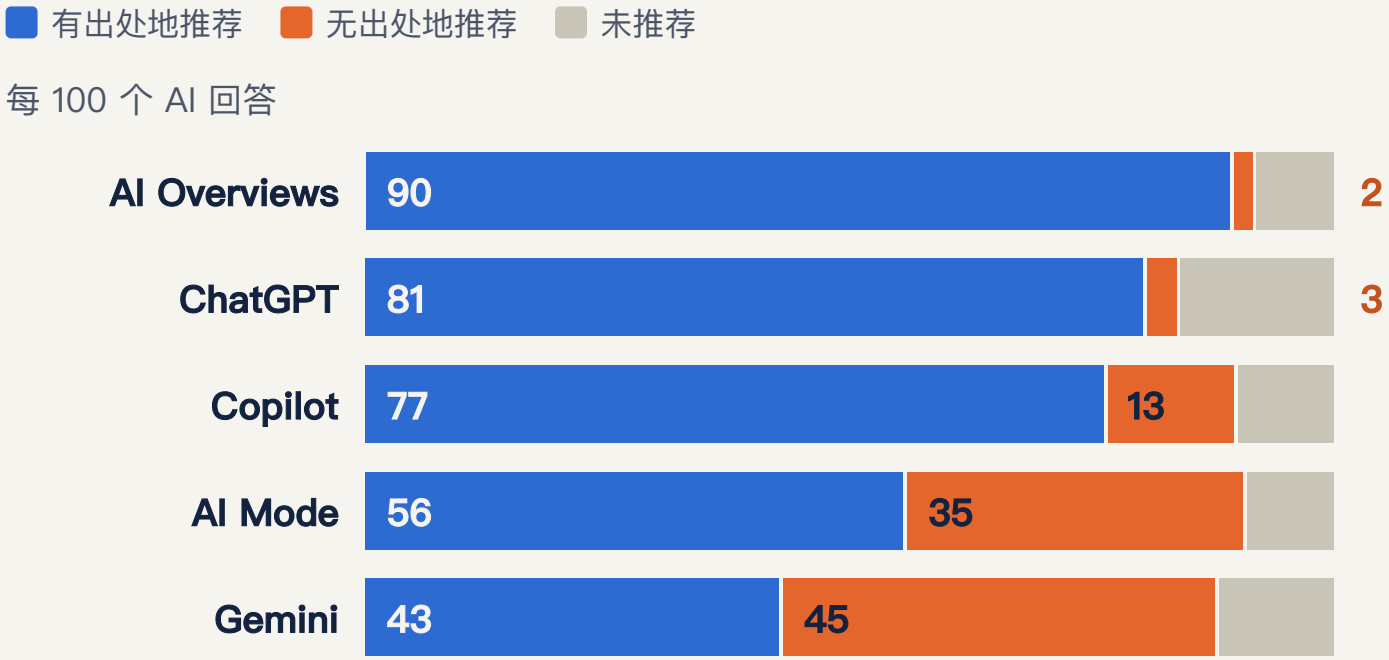
验收
型号、价格和库存表述一致

共同事实由官网承接，站外渠道按目标平台与实际品类数据选择

依据：白皮书T1平台来源分布；渠道策略属于可验证的行动建议

洞察 03 · 反常识

品牌推荐与引用分开看，Gemini 近半回答未附来源



对企业意味着什么

在 Gemini、AI Mode 上，无出处推荐较常见，需同时关注页面与品牌认知。

无链接属于输出特征；具体生成机制仍需要更多证据确认。

动作：统一品牌名与品类定位表述，维护知识图谱与百科条目，持续获得权威报道；"引用率"和"推荐率"分开考核。

数据：GEO 白皮书 T3，每 100 个回答的构成（最大余数法取整）；"无出处地推荐" = 提到商业品牌但没有任何来源链接

指标拆解 · 无来源品牌露出

45% 与 51.3%：分母不同，读法不同

全部回答作分母

≈45%

无来源且推荐的回答
÷ 全部回答

同一个输出特征

分母不同

两种比例分别报告

推荐回答作分母

51.3%

无来源且推荐的回答
÷ 所有推荐回答

报告写清分母；“无来源”仅描述输出，生成机制需更多证据确认

数据：白皮书T3图17/18；45%为每百回答取整，51.3%为精确分组比例

行动方法 · 品牌认知

让品牌名称、品类与事实形成一致的实体

01 身份一致

名称字段
法定名、品牌英文名、常用简称

品类字段
产品系列、核心用途、适用人群

边界
避免跨地区、跨版本混用事实

02 证据一致

站内事实
产品、价格、政策有权威页面

第三方证据
媒体、评测、目录真实可查

用户经验
保留体验条件与适用范围

03 更新一致

负责人
为关键事实指定维护人

变更记录
版本、日期、来源同步更新

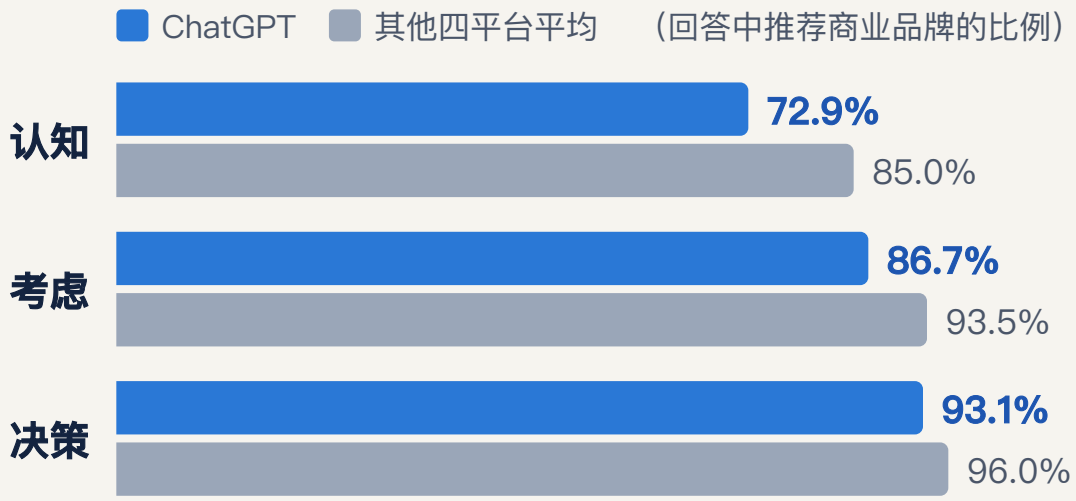
定期抽查
监测品牌描述是否准确

实体建设的交付物：一份可追溯、可维护、跨渠道一致的品牌事实表

方法：白皮书品牌实体与可推荐框架；实体建设效果需持续检验

洞察 04 · 商业机会

问题越接近决策，AI 越会推荐品牌



机构/政府/教育类信源

10.4% → 3.8%

从认知到决策，少了约三分之二

电商平台信源

2.6% → 4.8%

从认知到决策，近乎翻倍

对企业意味着什么

认知阶段

推荐率相对较低，可用原理内容和行业证据，验证新品牌的露出机会。

认知阶段：原理内容、行业报告、参与标准，让品牌在"顺带举例"时被想到。

决策阶段：管好官网定价页与电商商品页，价格、库存、政策要准。

数据：GEO 白皮书 T2，带有效漏斗标签的约 502 万条回答；信源变化为五平台合并口径

深入 03 · 行业差异

B2B 技术品类的官网来源占比达到 70%–73%

各行业来源链接中的官网占比



对企业意味着什么

三个技术品类更依赖官网
适合完善产品与开发文档

消费硬件同时重视媒体
专业评测与零售商品信息

先审计本行业
再决定官网与站外工作比例

数据：白皮书第三部分图14；行业平台覆盖不同，跨行业仅看方向

行动方法 · 内容架构

同一产品，准备认知、比较与决策三套内容

01 认知阶段

用户问题
快充协议怎样影响充电速度?

对应页面
原理说明、术语与测试方法

证据要求
定义准确，方法与数据可查

02 考虑阶段

用户问题
两种方案分别适合什么场景?

对应页面
条件清晰的对比与选型说明

证据要求
比较口径、适用前提一致

03 决策阶段

用户问题
这个型号是否兼容，怎样保修?

对应页面
产品、定价、兼容与政策页

证据要求
价格、版本、地区与日期明确

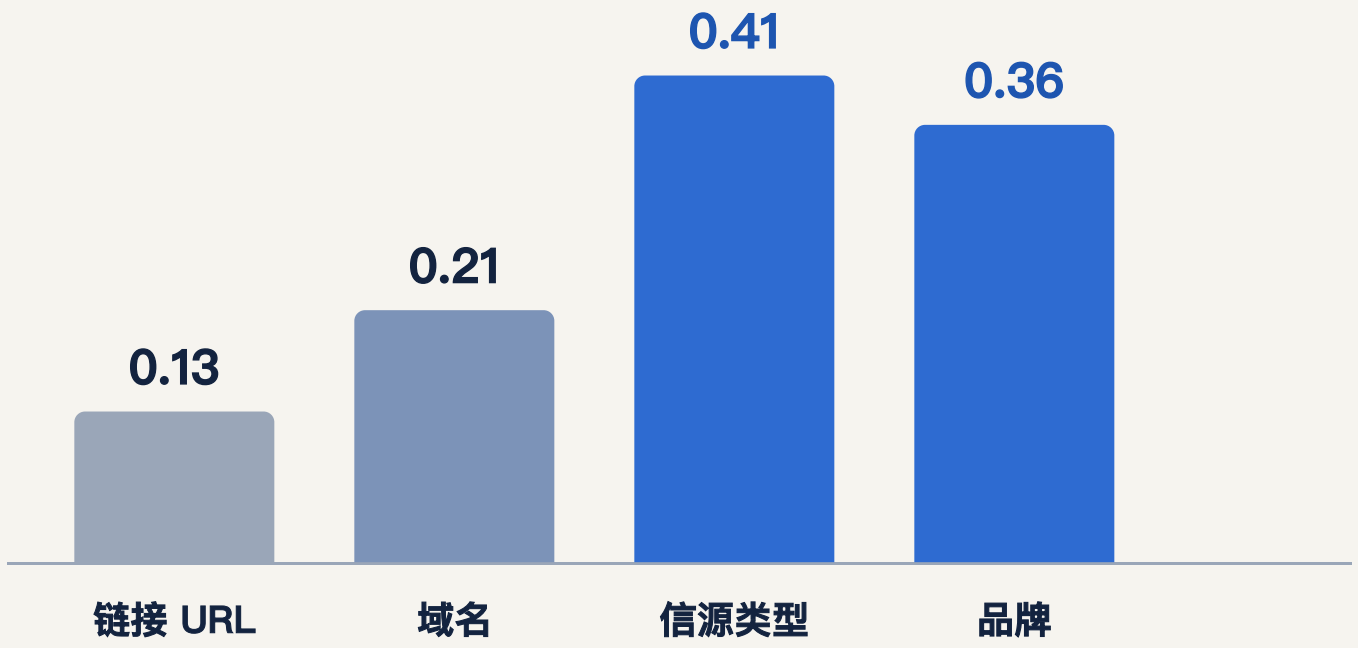
组织内容时保留共同事实，让不同页面分别回答真实的购买问题

依据：白皮书T2漏斗分层；问题示例用于教学，不代表实测效果

洞察 05 · 反常识

AI 的答案持续变化，品牌与信源类型相对稳定

相邻两周，同一问题答案的重合度（Jaccard，五平台平均；1 = 完全相同）



2%–13%

首次出现的 URL，次日仍连续保留的比例

对企业意味着什么

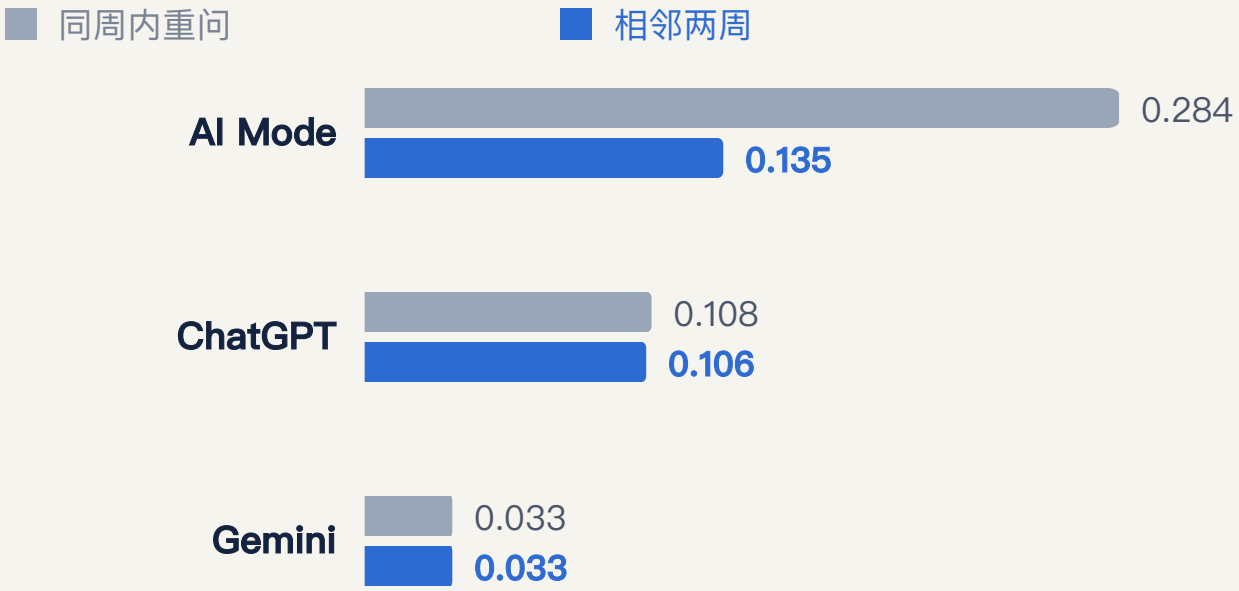
观察多个来源上的滚动出现率，避免单次观测影响资源判断。

在官网、YouTube、主流评测站、百科等多个稳定来源上建立**冗余存在**，考核 28 天滚动出现率。

数据：GEO 白皮书 T4，固定面板最长连续 17 周；次日留存为逐日复测的累计留存率

深入 04 · 随机性与时间

AI Mode 的跨周变化超出周内重复波动



对企业意味着什么

AI Mode 跨周相似度下降约53%
需要按时间追踪

ChatGPT、Gemini 的间隔差异小
重复波动更明显

固定题库、国家和模式
同时看题内波动与时间变化

数据：白皮书第三部分图22；URL Jaccard，周内与相邻周比较

指标拆解 · 引用漂移

连续留存与滚动出现率，回答不同的问题

01 周一：出现

品牌 / URL 出现

连续留存从这一刻开始
滚动窗口记为一次出现

02 周二：缺失

本次没有出现

连续留存首次终止
滚动窗口保留历史记录

03 周三：再次出现

品牌 / URL 重新出现

连续留存保持已终止
滚动窗口再记一次出现

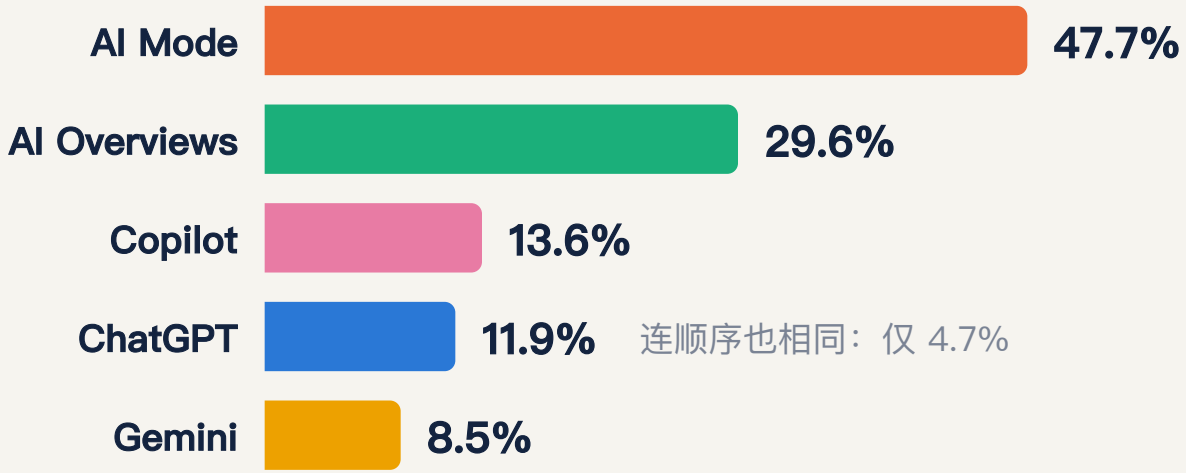
同一示意序列：连续留存在第 2 次观测终止，滚动出现率为 2/3

教学示意：三次观测只解释口径；平台实证的连续留存见T4

洞察 06 · 反常识

同题重测的品牌组合经常变化，单次截图信息有限

同一天把同一问题再问一遍，给出完全相同品牌组合的概率



±34 个点

仅重复 2 次时，95% 区间的平均宽度约为 68 个百分点

≈ 97 次

单个问题要把误差压到 ±10 个百分点，需要重复的次数

对企业：GEO 验收以问题数、重复次数和置信区间为依据。

数据：GEO 白皮书 T5，同日重测 7,030 对；区间为 Wilson 95% 区间；SparkToro 2026 研究：同一品牌清单重复出现的概率不到 1%

统计方法 · 样本规划

精度要求越高，单个问题需要的采样越多

±15 个百分点

≈43 次

用于核心问题的粗粒度追踪

±10 个百分点

≈97 次

用于更精细的单题变化比较

±5 个百分点

≈385 次

更高精度需要显著增加样本

规划前提：95% 水平、p≈0.5 且样本独立；相关性会影响实际精度

数据：白皮书T5图28；95%水平、p≈0.5、独立采样的保守近似

行动方法 · 监测协议

每份 GEO 周报，保留同一份测量记录

01 问题与权重

固定面板
问题、国家、语言与意图分层

扩展题库
新增问题单列，保留历史可比性

报告内容
题数、分层权重与有效覆盖

02 采样与环境

采集条件
平台、模式、日期与时间窗

有效样本
每题重复次数、缺失与重试

报告内容
口径变化、断点与异常

03 指标与业务

中间指标
推荐、来源、准确性分别统计

统计精度
样本量、置信区间与变化差异

业务结果
引荐、询盘、线索与转化

让每一条增长结论都能回答：测了什么、测了多少、变化有多确定

方法：白皮书T5与第四部分指标体系；按题聚类处理重复采样相关性

行动方法 · 五可框架

“五可”框架：每一层都配置验收项

把技术接入、内容质量和品牌证据接起来，按层排查建设缺口

可抓取

先进入候选

爬虫权限与返回状态
索引与正文可读取性

可理解

直接回答问题

标题与结论对应
步骤与单位清楚

可信任

证据有出处

作者与机构明确
原创方法与更正透明

可引用

事实可抽取

结论有条件与日期
引用能支撑具体表述

可推荐

品牌有证据

品类定位准确
第三方证据真实可查

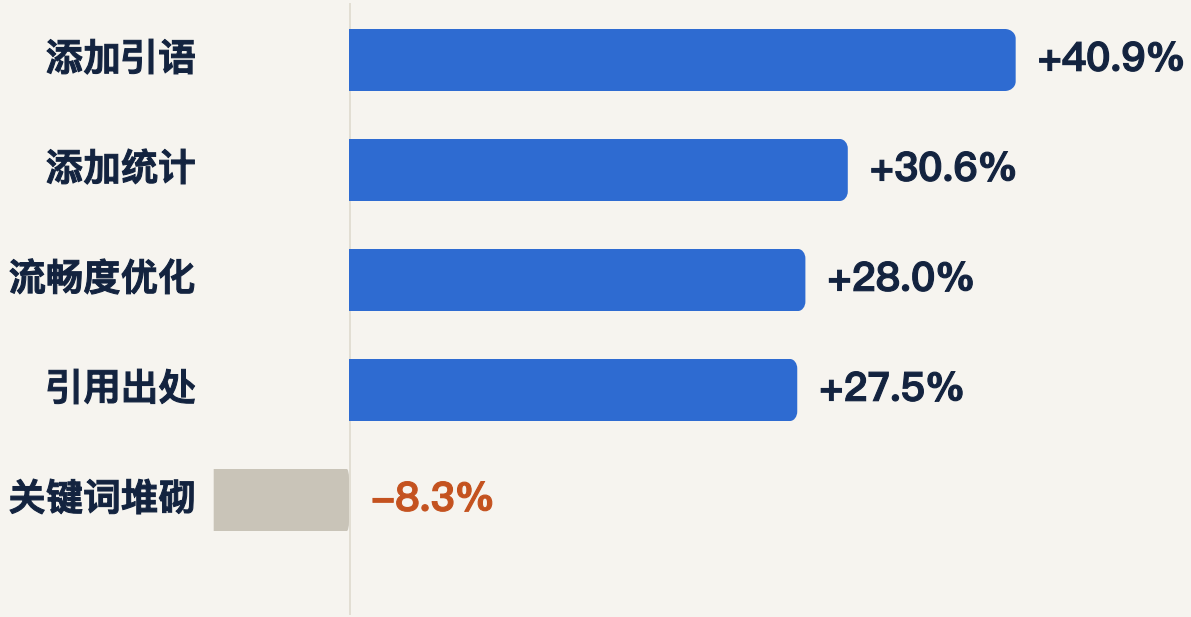
验收顺序：接入与事实先完成，再持续验证引用、品牌表述和业务结果

来源：白皮书第四部分方法论；证据强度和效果因环节、场景而异

证据边界 · 内容改写

“提升 40%”适用于已进入上下文的实验

固定上下文中的可见度变化



对企业意味着什么

这些结果反映固定上下文中
可见度的变化

线上页面先要被抓取与召回
才能进入这一实验范围

保持问题相关性
用对照验证完整链路效果

来源：GEO–bench / KDD 2024；GPT–3.5，固定候选上下文

行动方法 · 效果验证

用一轮有对照的实验，验证内容与信源改造

固定范围

先确定问题集

明确国家、语言和产品
选定平台与观察窗口

记录干预

一次改动可追溯

记录页面、事实与来源
保留上线时间及版本

保留对照

观察共同波动

选择同期未改动的题组
核对意图与基线是否可比

估计差异

按预定协议验收

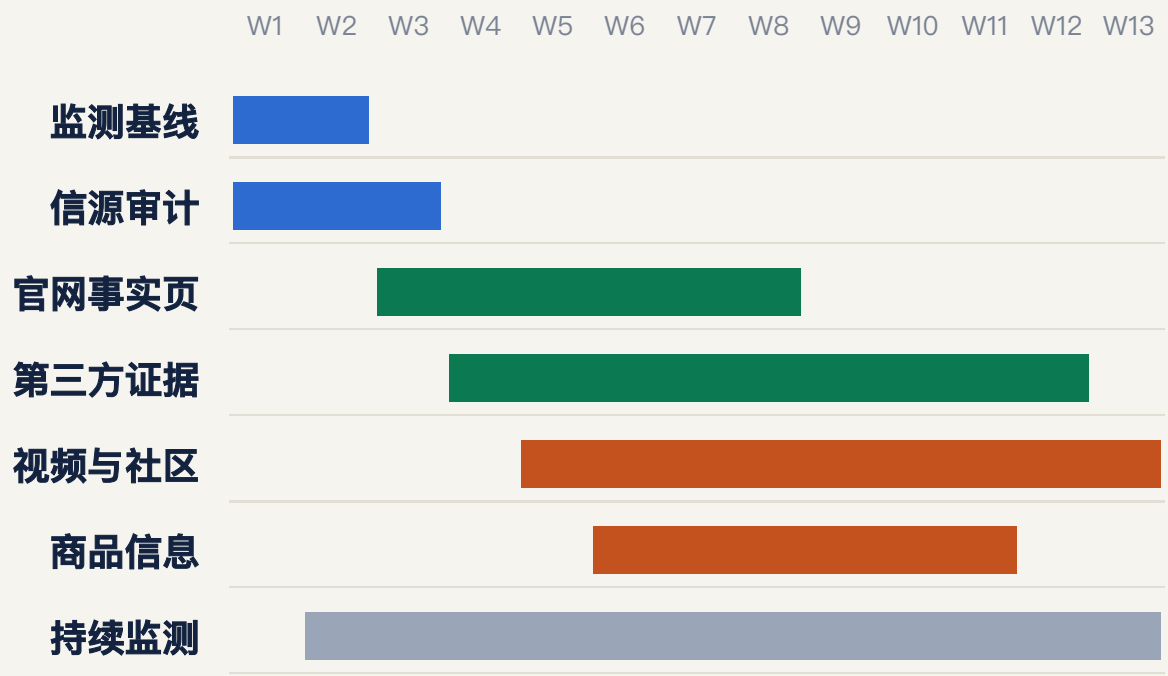
分别报告推荐与准确性
结合引用、业务指标验收

先定义怎样判断有效，再建设内容；把共同波动与项目影响分开估计

设计建议：预定义题库与指标，记录上线时间，保留对照问题或页面

实施路径 · 90天路线

90 天项目：先交付基线，再按同题复测



三个里程碑

第 2 周：交付监测基线
形成问题库与信源审计

第 8 周：核对上线进度
检查事实准确性与抓取

第 13 周：按同题复测
估计变化并安排下一阶段

排期建议：W1–W13约90天；核心题采样深度按精度目标另行规划

给出海企业决策者

三件事：建资产、换尺子、按节奏

01 建资产

官网事实底座：一个事实一个权威页面

品牌实体：统一品牌名与品类定位，经营知识图谱与权威报道

跨平台守门人：YouTube、Reddit、Amazon

02 换尺子

从"排名截图"换成**出现率 + 置信区间**

分平台考核，保留各平台结果明细

引用率与**推荐率**分开看，再看描述准不准

03 按节奏

第 1–2 周：分平台基线 + 信源审计

第 3–12 周：官网建设，布局媒体、评测、视频与社区

第 13 周：同一把尺子复测，再定预算

底线是白帽合规：不刷量、不伪造评价、不批量灌内容。被 AI 读懂，靠的是真实、准确、可核验的信息。

获得《海外GEO白皮书》完整版



扫描公众号，回复白皮书