

完整研究报告 · v7.0

超级入口

AI 搜索与 GEO

深度研究报告

机制、证据、商业价值、平台差异、风险治理与企业实施体系

个人署名研究 吴兆华 Codex 协作研究与制图 · 非 e签宝正式立场 · 证据截止 2026-07-12

先区分事实、观察、框架、预测与建议，再讨论企业是否投入。

已验证事实

观察性关联

作者框架

趋势预测

管理建议

本报告不承诺固定排名，不把单次模型回答当作稳定规律，也不把厂商自报、观察相关性或趋势预测包装成普遍因果结论。所有核心数字应能回到证据台账核查样本、时间、地区、单位和分母。

AI 搜索与 GEO 深度研究报告

从检索、证据与推荐，到 Agent 可执行的企业信息基础设施

版本： v7.0

证据截止： 2026 年 7 月 12 日

作者： 吴兆华个人研究

研究协作： Codex 协助资料整理、交叉核验、结构化写作与制图

报告属性： 独立行业与企业战略研究，不代表 e签宝或任何任职机构的正式观点；不构成证券投资、法律或采购建议

本报告按“机制—证据—商业—平台—风险—实施”组织。企业实施手册已合并进入正文和附录，不再单独重复论述。管理层若只需要投入判断、90 天议程与停止条件，应先阅读配套《管理层决策简报 v7.0》。

摘要

【已验证事实】AI 正在改变信息发现的界面和任务链，但尚未让传统搜索失去基础作用。Google 把生成式回答嵌入搜索结果并公开 query fan-out；ChatGPT、Claude、Perplexity 等助手获得实时检索能力；商品发现、购物和工作型 Agent 开始把搜索、数据调用与行动连接起来。与此同时，平台仍依赖搜索索引、网页、商品源、合作数据和结构化信息，经典 SEO、准确事实和真实业务能力并未失效。

【作者框架】本报告将 GEO 定义为 (CLM-031)：

企业在 AI 介导的发现、理解、比较、推荐和行动过程中，系统提高自身信息被正确发现、被作为合适证据、被准确表述、被纳入合适候选以及被机器顺利执行的能力。

这一广义定义不是行业唯一共识，也不是某个平台的官方术语。它的价值在于把内容优化、技术可访问性、品牌事实、第三方证据、结构化数据、交易接口、测量和治理放进同一经营体系。

【管理建议】企业应先完成事实防御，再验证增长机会：建立唯一事实源、主张与证据台账、平台基线和纠错机制；随后围绕一个品类、一个地区和一组高价值问题进行受控试点；只有在准确性、候选进入、商业信号和风险成本同时改善时才扩大投资。

【趋势预测】Agentic Commerce 可能把 GEO 从营销能力推进到产品、数据与交易基础设施，但协议采用速度、最终赢家和收入贡献仍高度不确定。企业应优先建设可跨渠道复用的商品、服务、权限、库存、价格、履约与审计数据，不应把预算押在单一聊天入口或协议上。(CLM-032)

结论标签与证据规则

本报告给重要结论标注五类性质：

标签	定义	可否直接当作因果或标准
【已验证事实】	有可追溯的官方、政府、学术或可复核来源	仍须受样本、时间、地域和方法限制
【观察性关联】	数据中存在共现或相关，不能排除选择偏差和混杂因素	否
【作者框架】	作者用于组织问题和经营决策的模型	否，不是平台算法
【趋势预测】	根据现有产品、协议、行为与资本信号推演	否，必须管理情景和时点风险
【管理建议】	面向企业的实施起点、阈值或治理要求	否，需按企业校准

配套证据台账为关键主张分配 CLM-001 至 CLM-022 等稳定 ID。所有数据图应从同一结构化证据记录生成，并显示样本、时间、地区、单位、分母、证据等级和因果边界。AI 生成概念图仅用于解释机制，不作为事实证据。

第一部分 机制：AI 搜索究竟改变了什么

1. 搜索没有消失，而是从前台动作变成后台能力

【已验证事实】传统搜索结果页正在 AI 化。Google 公开说明 AI Overview 和 AI Mode 依赖 Google Search 的索引、核心排名与质量系统；复杂问题可通过 query fan-out 拆成多个相关查询。Google 还在 Search Console 为部分站点提供生成式 AI 专项表现视图，用于观察 AI Overviews、AI Mode 等入口的表现，但覆盖范围和指标可见性仍由平台控制。(CLM-022、CLM-002) [Google AI 优化指南](#) · [Search Console 更新](#)

【已验证事实】聊天助手也在获得检索能力。OpenAI 说明 ChatGPT Search 可以把自然语言问题改写成一个或多个定向查询；Anthropic 说明 Claude 会判断是否需要搜索并可连续调查；Perplexity 将搜索抓取器与用户触发访问者分开说明。不同平台的索引、合作来源、抓取方式、重排和引用政策并不相同。(CLM-027、CLM-024、CLM-025) [OpenAI](#) · [Anthropic](#) · [Perplexity](#)

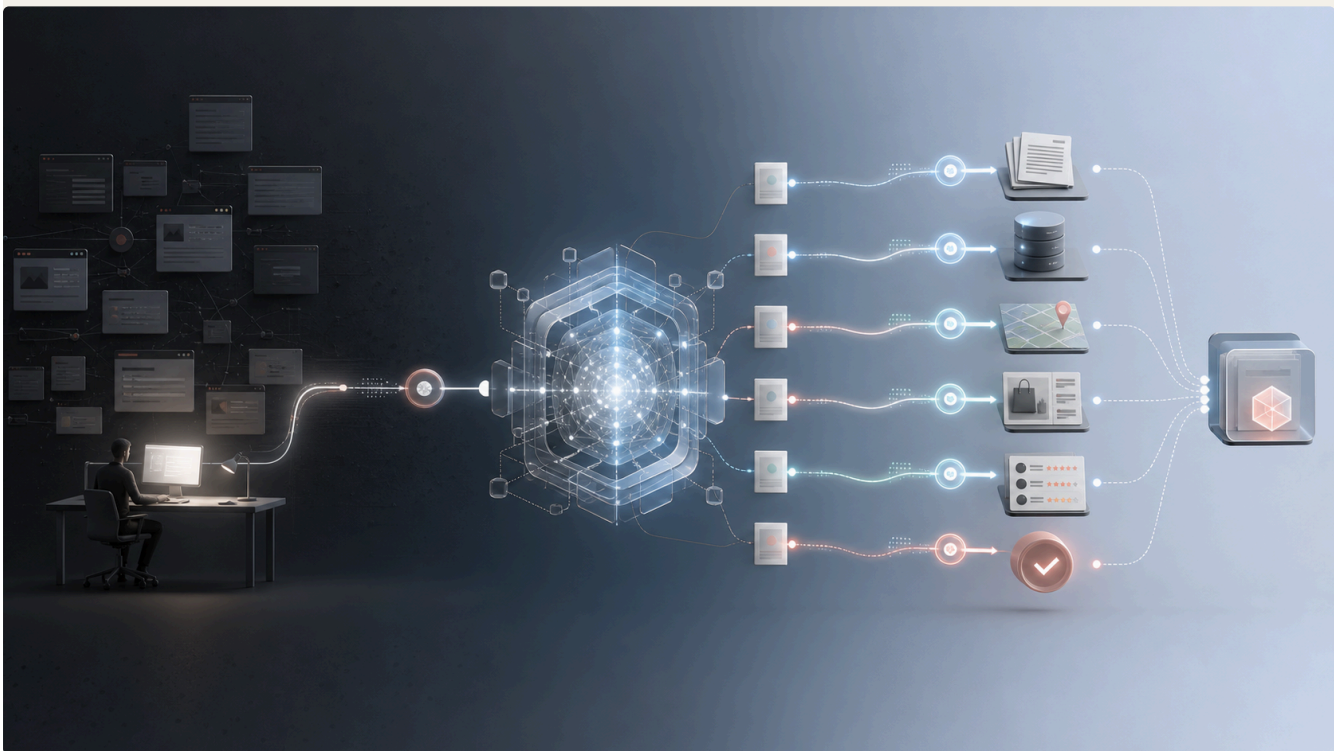
【作者框架】更准确的变化不是“搜索被某个 AI 产品取代”，而是：

人类少做一部分显式关键词搜索；

搜索引擎在结果页内部生成和综合答案；

AI 助手在后台改写、扩展和连续执行查询；

Agent 把检索、比较、数据调用和行动串成 workflow。



从人类显式搜索到 Agent 后台检索

AI 生成概念图，仅用于机制与场景表达，不作为事实证据。

【已验证事实】公开资料足以证明查询拆解和多轮检索存在，却不足以证明“每次用户意图都固定触发几十或几百次搜索”。Cloudflare 的“网页爬取与人类回流”比率反映机器读取与回流的不对称，不是每个用户问题的查询次数。[Cloudflare](#)

2. 五类知识来源不能混为一谈

来源	主要作用	时效性	企业可控性	主要治理对象
模型参数记忆	通用知识、语言模式、历史品牌印象	低至中	低、反馈慢	长期公开共识、品牌实体一致性
搜索索引	从已抓取网页中召回候选	中至高	中	可抓取、可索引、站内结构、经典 SEO
实时网页抓取	读取当前网页或用户指定页面	高	中至高	WAF/CDN、正文可读性、版本和更新时间
结构化数据库/商品源	价格、库存、地点、规格、商户、航班等	很高	高	Feed、Schema、PIM、Merchant Center、API
私有和用户上下文	企业知识库、文件、记忆、位置、偏好和权限	视数据而定	企业内高	权限、数据质量、来源、审计与撤销

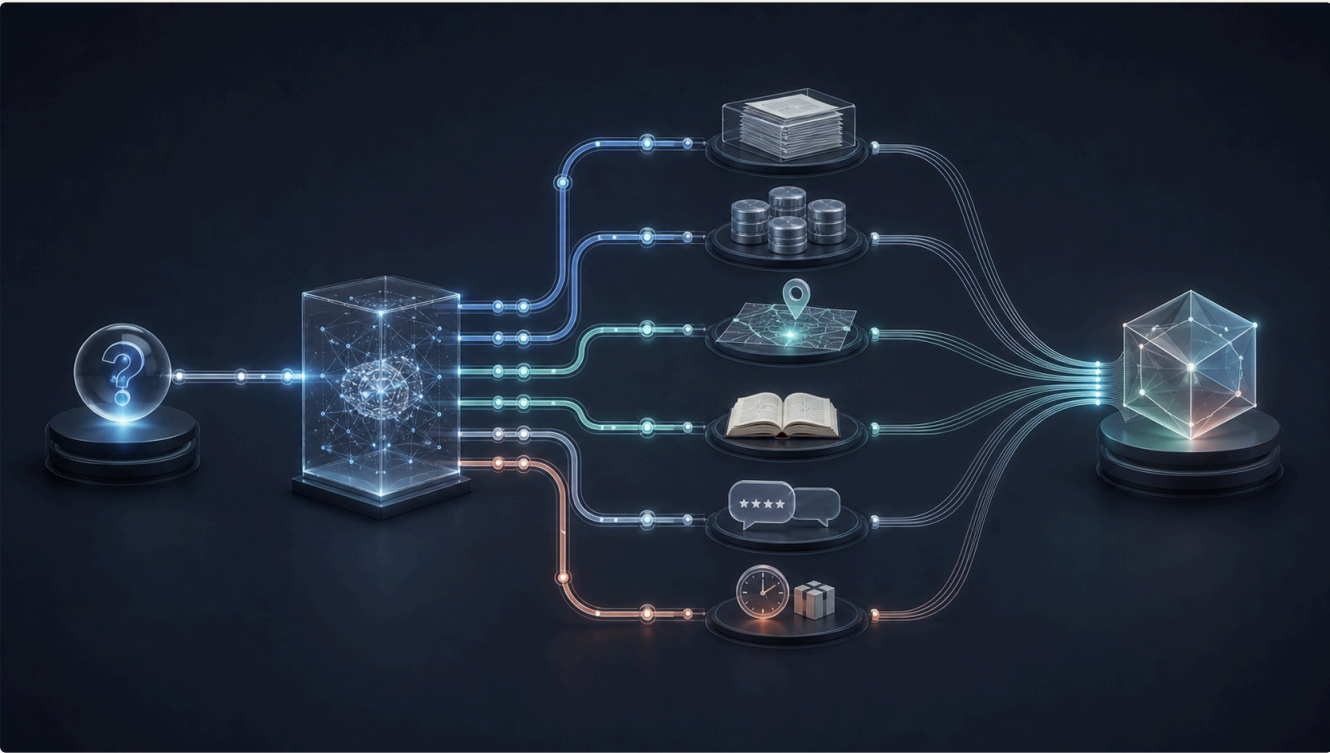
【已验证事实】发布一篇新文章通常不等于立刻改变模型参数；但新页面有机会通过搜索索引、实时抓取或 RAG 进入当次回答。央视 3·15 案例反映的主要是检索证据污染，而不是已经证明基础模型参数被永久训练。

【管理建议】企业在讨论“AI 是否知道某件事”时，必须先说明是哪一种知识来源。训练许可、搜索发现、用户触发访问和 Agent 行动也应分别制定 crawler、权限与审计策略。

3. 一条典型 AI 搜索链路

【作者框架】可把一次回答拆成八个决策点：

环节	系统问题	企业可能失去机会的原因
用户与上下文	用户是谁、场景和约束是什么	企业只覆盖宽泛关键词，没有角色和约束内容
搜索判断	是否需要调用搜索或工具	系统直接用旧参数记忆回答
查询拆解	要生成哪些子问题	企业没有覆盖价格、风险、比较、售后等子问题
数据路由	搜索网页、商品源、地图还是私有数据	企业信息只存在于不可访问渠道
候选与重排	哪些来源进入有限上下文	内容相关但证据弱、过期、重复或不适配地区
证据提取	哪些片段足以支撑主张	页面没有直接答案、条件、方法和可追溯来源
生成与引用	如何处理冲突、展示何种来源	企业被采用但未引用，或被用于负面/限制语义
比较与行动	是否列入候选并允许下一步	商品、库存、价格、预约、权限或履约数据缺失



AI 查询拆解与多路检索机制

AI 生成概念图，仅用于机制与场景表达，不作为事实证据。

【已验证事实】“被检索、被采用、被引用”是不同事件。研究者对真实搜索型回答的观察也发现，检索日志和引用披露并不完整，因此引用率只是可见代理变量，不等同于全部影响。[Cambridge Data & Policy](#)

【管理建议】监测至少要区分：未检索；检索未采用；采用未展示引用；采用并引用；正面推荐；负面或限制性引用；发生点击/行动。不能用“出现次数”替代这条漏斗。

4. AI 为什么采纳一条信息

【作者框架】不存在跨平台统一、固定的“AI 认可分”。在特定问题、时间、地区、账号和上下文中，系统会综合判断：

任务相关性： 是否直接回答用户问题和约束；

证据密度： 是否包含数字、条件、时间、方法和来源；

原始性： 是否拥有第一手数据、实验、产品文档、现场经验；

可追溯性： 能否回到法规、论文、公告或原始记录；

时效与版本： 是否写明发布日期、更新日、地域和适用版本；

来源适配： 规格偏向厂商，监管偏向官方，效果评价偏向独立证据；

独立性： 多个来源是否真正独立，而非同一稿件转载；

可解析性： 正文、表格、标题、链接和图片说明是否稳定可读；

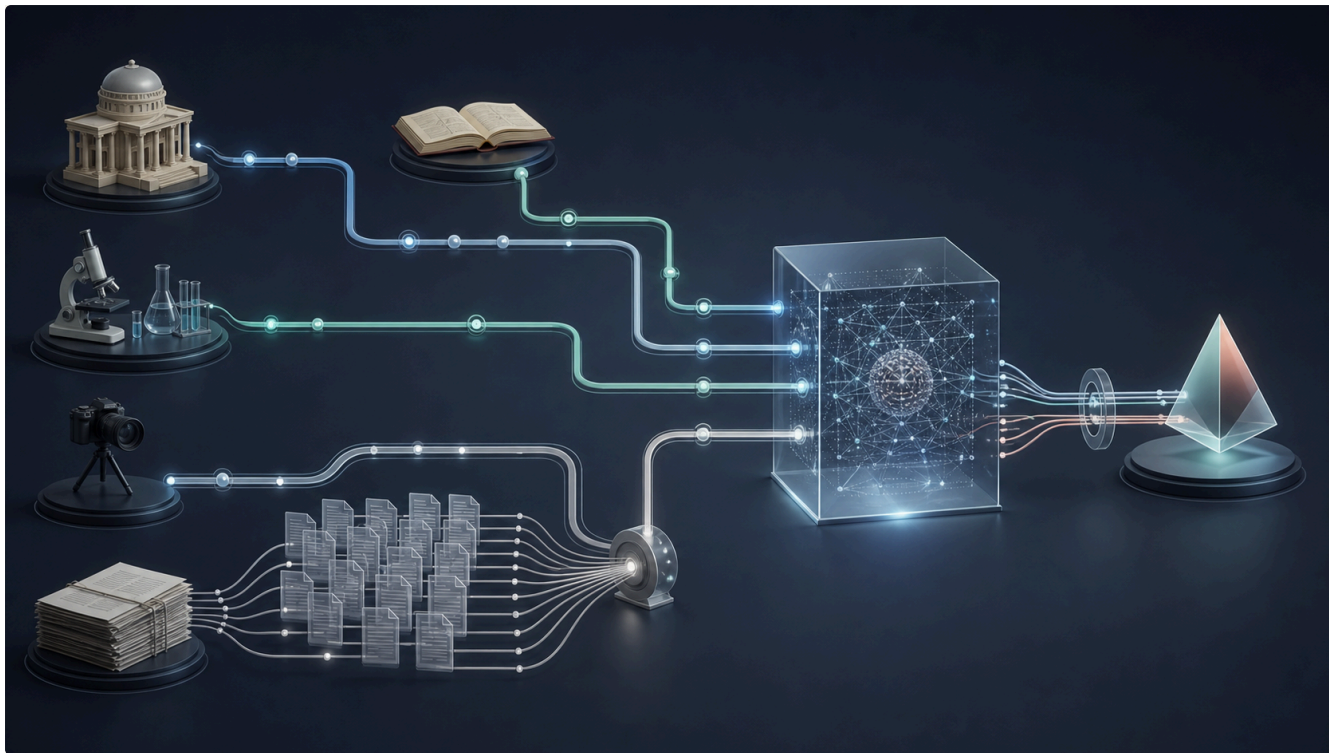
风险与安全： 高风险领域会有更严格来源和表述边界；

用户适配： 预算、地区、角色、偏好和已有工具会改变候选。

【已验证事实】Google 的 AI 搜索指南说明，传统 SEO 基础仍适用，网站不需要特殊 AI Schema 或新的 AI 专用文件；独特、面向用户且可索引的内容仍是基础。(CLM-022) [Google](#)

【管理建议】企业不应为迎合 AI 批量制造近似页面、虚假外部提及或机械切块；当前也没有充分证据把 `llms.txt` 包装成普遍有效的独立排名捷径。这是基于公开机制与风险证据形成的治理建议，不是 Google 对每一种做法的逐项官方判定。

【作者框架】官网和第三方承担不同信任职能：官网负责企业可承担责任的标准事实；政府、标准、客户、伙伴、专业媒体、真实评论和独立研究负责外部佐证；结构化数据负责可计算；真实业务结果负责最终说服。第三方来源只有在来源、所有权和商业关系可识别时，才构成真正独立证据。



从企业主张到独立证据的来源血缘

AI 生成概念图，仅用于机制与场景表达，不作为事实证据。

第二部分 证据：市场信号与研究边界

5. 用户规模和入口变化

5.1 全球平台信号

【已验证事实】Google 官方在 2026 年披露 AI Overviews 的月度用户规模达到约 25 亿。该数字说明生成式回答已进入大规模搜索体验，但它是产品触达规模，不等于独立用户主动选择 AI 搜索，也不等于网站流量或商业收入。(CLM-001)

【已验证事实】Google Search Console 于 2026 年推出生成式 AI 专项表现报告，向部分站点展示 AI Overviews、AI Mode、Discover 等体验的页面、国家、设备和日期数据。这表明平台正在把 AI 搜索纳入站长测量，但数据可见范围仍有限，且不能替代企业自己的答案语义和引用监测。(CLM-002)

【观察性关联】Similarweb 的重叠数据应按单向口径解读：大多数 ChatGPT 用户也访问 Google，并不等于 Google 与 ChatGPT 双方各有 95% 用户重叠，也不能证明两个入口可相互替代。(CLM-004) [Similarweb](#)

AI 搜索的覆盖信号已经进入十亿级，但口径并不相同

全球产品月活与中国过去半年使用人口分栏呈现，不做规模排名

不同口径，不可横比



样本：Google 内部活跃数据；CNNIC 全国互联网统计调查
时间：Google 截至 2026-05；CNNIC 截至 2025-12 | 地区：全球 / 中国大陆
单位：人；Google 为月活，CNNIC 为过去半年使用人口 | 分母：各产品月活 / 中国整体人口；三项不可直接比较或相加
证据等级：B | 因果性：不适用；描述性覆盖事实
边界：Google 未披露去重方法；CNNIC 与产品 MAU 不是同一统计口径。
来源：CLM-001 Google；CLM-019 CNNIC

AI 搜索的覆盖信号已经进入十亿级，但口径并不相同

图表口径：各平台公开规模快照，不构成同口径市场份额排名；来源、分母与限制见证据台账。

95% 是单向条件概率，不是 Google 与 ChatGPT 的对称重叠

非数值比较图

Similarweb：在 ChatGPT 用户中，约 95% 同时使用 Google



CLM-004

样本：Similarweb Q4 2025 audience overlap 面板；样本规模未公开

时间：2025 Q4 | 地区：来源未披露完整地域范围

单位：单向用户重叠率 | 分母：ChatGPT 用户

证据等级：C | 因果性：否 | 描述性用户重叠

边界：不能反推 Google→ChatGPT 也是95%，也不能把两个用户群画成95%对称交集。

来源：CLM-004 Similarweb

95% 是单向条件概率，不是 Google 与 ChatGPT 的对称重叠

图表口径：以 ChatGPT 用户为分母的单向重叠，不可反向解释。

5.2 中国用户规模

【已验证事实】CNNIC 第 57 次《中国互联网络发展状况统计报告》显示，截至 2025 年 12 月，中国生成式人工智能用户规模为 6.02 亿，普及率 42.8%；同一报告图表保留了 2025 年 6 月 36.5% 的历史点。该数据支持“生成式 AI 已是大规模信息入口”，但不能证明每个行业的采购决策已经同等迁移。（CLM-019）
[CNNIC 第 57 次报告 PDF](#)

【观察性关联】QuestMobile 2025 年 6 月口径下，AI 搜索引擎月活跃用户约 6.85 亿，AI 助手约 6.12 亿；统计包含 In-App AI，类别之间存在重叠，不可简单相加，也不能与独立 App 榜单混为一谈。（CLM-020）
[QuestMobile](#)

【管理建议】用户规模决定“值得审计”，不直接决定“应该投入多少”。企业应补充自己的客户证据：销售访谈、站内来源、品牌搜索、采购调研和 Prompt 基线，判断 AI 是否已经进入真实决策链。

中国 AI 使用规模：全国人口调查与移动赛道月活必须分开看

仅同面板内可比较

不展示缺少可追溯来源与统一月份口径的单品数字



样本：CNNIC 全国互联网统计调查；QuestMobile 移动行为面板
时间：CNNIC 截至 2025-12；QuestMobile 2025-06 | 地区：中国大陆 / 中国移动互联网
单位：过去半年使用人口 / 赛道月活用户 | 分母：CNNIC 为整体人口；QuestMobile 为各移动 AI 赛道且可重叠
证据等级：C | 因果性：不适用；描述性规模数据
边界：三项不得相加或做单品排名；QuestMobile 赛道口径也不能替代独立 App MAU。
来源：CLM-019 CNNIC；CLM-020 QuestMobile

中国 AI 使用规模：全国人口调查与移动赛道月活必须分开看

图表口径：CNNIC 公开统计快照；不代表所有行业采购决策已同比例迁移。

6. 来源覆盖与平台差异

【观察性关联】一项研究分析 55,936 个查询、6 个 LLM 搜索引擎和 2 个传统搜索引擎，发现约 37% 的域名只出现在 LLM 搜索结果中。该研究没有证明这些独有来源质量更高，反而提醒企业：AI 来源覆盖与传统搜索不完全相同，且安全、可信度不一定更好。(CLM-018) [Source Coverage and Citation Bias](#)

【观察性关联】Ahrefs 对 7 个 AI 平台和约 1,700 万条引用的工具数据观察到，AI 引用内容整体更“新”，并对近期更新内容呈现一定偏好。该结果受平台、查询和厂商定义影响，不能转化为“频繁修改日期即可提高推荐”的规则。(CLM-009) [Ahrefs](#)

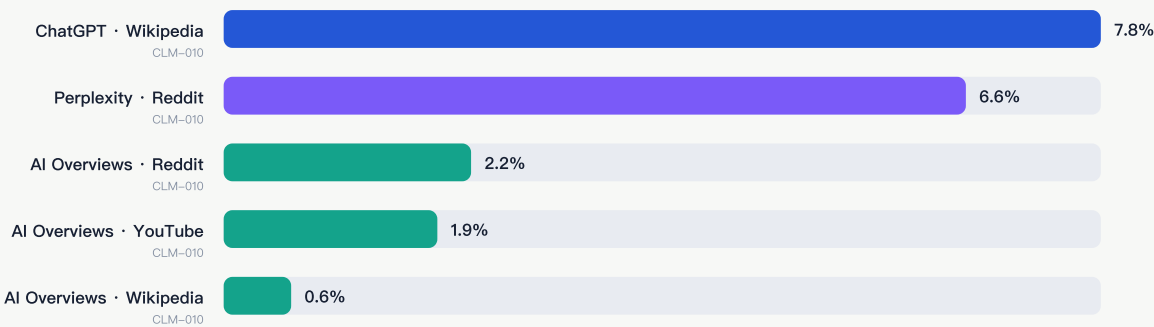
【观察性关联】商业监测也显示不同平台的常见来源域名名单存在较大差异。平台会快速调整来源偏好，因此任何“某媒体永久是权威源”“一个发布渠道覆盖所有模型”的承诺都缺乏依据。

【观察性关联】Profound 对约 6.8 亿条引用、2024 年 8 月至 2025 年 6 月窗口的商业监测显示，按“各平台全部引用”统一分母：Wikipedia 约占 ChatGPT 引用的 7.8%，Reddit 约占 Perplexity 引用的 6.6%；Google AI Overviews 中 Reddit、YouTube、Wikipedia 约占 2.2%、1.9%、0.6%。这些数字不能与该研究另一张“Top 10 来源内部份额”图直接混用，也不是平台排名权重或长期规则。(CLM-010) [Profound](#)

平台引用来源差异：全图统一使用“占该平台全部引用”分母

同分母，可作描述性比较

不再混用 Top 10 来源内部份额；数值仅代表 Profound 监测样本



样本：Profound 约 6.8 亿条引用

时间：2024-08 至 2025-06 | 地区：Profound 监测查询覆盖市场；地域权重未完整披露

单位：百分比 | 分母：各平台监测样本中的全部引用

证据等级：C | 因果性：否；描述性监测数据

边界：样本查询组合可能偏离总体用户；不代表固定平台偏好、排名权重或未来表现。

来源：CLM-010 Profound

平台引用来源差异：全图统一使用“占该平台全部引用”分母

图表统一使用“占平台全部引用”分母；不可与 Top 10 来源内部份额混用。

【管理建议】按“平台 × 问题 × 地区 × 语言 × 时间”建立重复测试，不用一次截图推断平台长期偏好。平台差异越不透明，越需要保留完整回答、引用和测试条件。

7. KDD GEO 论文：重要起点，也最容易被误读

【已验证事实】KDD 论文《GEO: Generative Engine Optimization》在受控环境中发现，加入引用、引语和统计性表达等策略，可使来源可见性指标获得不同程度提升，最高相对提升约 40%；关键词堆积帮助很小。（CLM-007）[GEO 论文](#)

这项研究的边界必须完整保留：

主要实验从 Google 结果中取固定候选，再由 GPT-3.5 基于候选生成，不等于当前所有开放式商业 AI 搜索后台；

“40%”是实验中的相对可见性指标，不是点击、线索、市场份额或销售额；

对 Perplexity 的验证包含文件上传并要求基于材料回答，不等于开放 Web 的自然发现；

论文中部分策略提示允许虚构引用或生成假设性统计，所以实验说明“权威信号可影响可见性”，不说明系统验证了这些证据的真实性；

第五名来源提升 115.1%、第一名下降 30.3% 等数字来自同时优化多个候选的特定条件，不能解释为弱品牌普遍逆袭。

【作者框架】这篇论文最有价值的启示是：生成式回答会在有限答案空间中重新分配来源可见度；局部证据更适合问题的来源可能突破传统搜索座次。最危险的误读是：批量伪造统计、引语和权威口吻可以成为商业方法。

8. 点击、零点击与商业效果

8.1 零点击是压力，不是统一因果定律

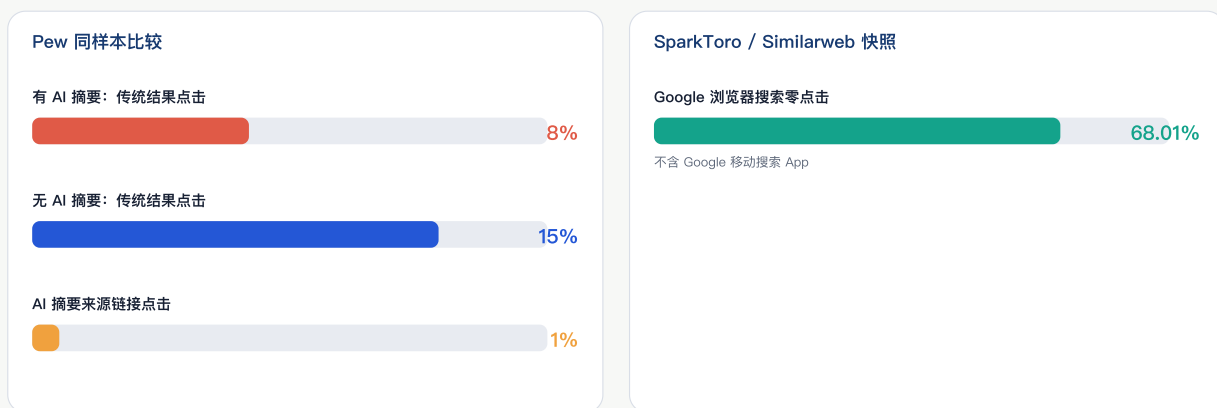
【观察性关联】Pew 对美国 900 名成人、68,879 次 Google 搜索的观察显示：出现 AI Overview 时，点击传统结果的搜索占 8%，未出现时为 15%；点击 AI 摘要引用来源约为 1%。这是观察性数据，触发 AI Overview 的查询本身更长、更偏信息型，不能把差异全部归因于摘要。（CLM-005）[Pew](#)

【观察性关联】SparkToro 与 Similarweb 的 2026 年美国浏览器面板估算，Google 搜索中约 68.01% 没有产生向开放 Web 的点击。该面板不含 Google 移动 App，研究方法也与往年不同，不能制作连续同比趋势或声称“零点击增长了固定百分点”。（CLM-006）[SparkToro](#)

点击减少的两个研究快照：不能拼成一条趋势线

仅同面板内可比较

左侧比较同一 Pew 样本内有无 AI 摘要；右侧是另一面板的零点击快照



样本：Pew：约900人、68,879次搜索；Similarweb：美国桌面与移动网页面板
时间：Pew 行为 2025-03；Similarweb 2026-01-04 | 地区：美国
单位：搜索访问 / 会话占比 | 分母：Pew 为有/无 AIO 的搜索访问；Similarweb 为浏览器内 Google 搜索
证据等级：C | 因果性：否；均为观察性研究
边界：两项研究样本、设备、会话定义与时间不同；不得跨研究计算增幅或归因给 AI。
来源：CLM-005 Pew；CLM-006 SparkToro/Similarweb

点击减少的两个研究快照：不能拼成一条趋势线

图表把 Pew 行为观察与 SparkToro/Similarweb 面板分开呈现，不连接成时间趋势。

【作者框架】零点击并不等于零商业价值。AI 可能在回答内建立认知、把企业纳入候选、影响后续品牌词搜索和直接访问、带来更少但意图更高的访问，或直接完成预约与交易。真正的困难是传统末次点击归因无法识别这些影响。

8.2 商业生态观察不能互相横比

【观察性关联】Shopify 的 2026 年早期生态数据称，AI 引荐购物者的商品页转化率比自然搜索近 50% 更高，平均客单价高 14%。这是 Shopify 自身生态的早期观察，可能受渠道选择、品类和样本结构影响。（CLM-011）[Shopify](#)

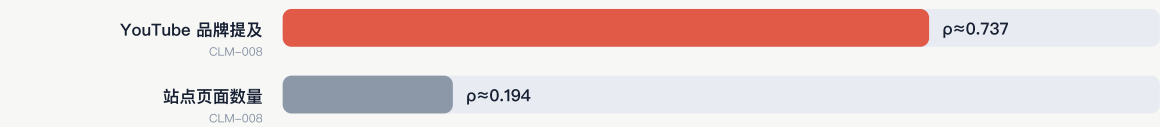
【观察性关联】Adobe Analytics 曾观察到生成式 AI 到美国零售和旅游网站的访问从较小基数出现显著相对增长。它反映 AI 引荐正在增长，但样本、时期和指标与 Shopify 不同，不能并列成统一转化结论。(CLM-012) [Adobe](#)

【观察性关联】Ahrefs 对约 7.5 万个成熟品牌的筛选样本中，YouTube 提及与 AI 可见度的 Spearman 相关系数约 0.737，站点页面数约 0.194。样本设置包括较高域名评级和品牌搜索量门槛；相关系数不是平台权重，也不能证明“反向链接失效”或“发视频导致推荐”。(CLM-008) [Ahrefs](#)

成熟品牌样本中的相关性观察，不是平台权重

同分母，可作描述性比较

YouTube 提及相关性较高；页面数量相关性较低，均不能推出因果



样本：75,000 个成熟品牌；DR>40 且主品牌词月搜索量≥800
时间：Ahrefs 研究发布于 2025-12 | 地区：Ahrefs 数据库覆盖市场
单位：Spearman 相关系数 ρ | 分母：筛选后的成熟品牌样本
证据等级：C | 因果性：否；相关不等于因果
边界：不能写成“权重体系”或“反向链接失灵”；样本选择限制对新品牌与中小企业的外推。
来源：CLM-008 Ahrefs

成熟品牌样本中的相关性观察，不是平台权重

图表展示 Spearman 相关性，不是平台权重，不构成因果结论。

AI 引荐的三组商业观察：方向可参考，数值不可横向比较

不同口径，不可横比

流量增幅、商品页转化和访客价值是三种不同指标



样本：Adobe 超1万亿零售访问；Shopify Q1生态；Semrush 自有监测（样本未完整公开）
时间：Adobe 2024-07—2025-02；Shopify 2026 Q1；Semrush 截至 2025-06 | 地区：美国 / Shopify 生态 / Semrush 覆盖市场
单位：相对访问增幅 / 相对转化率 / 相对访客价值 | 分母：三项分母不同，禁止排序、平均或相互印证固定倍数
证据等级：C | 因果性：否：平台生态与厂商观察
边界：不采用来源不清的总流量占比说法；三项只支持“部分样本出现高意图信号”的方向判断。
来源：CLM-012 Adobe；CLM-011 Shopify；CLM-026 Semrush

AI 引荐的三组商业观察：方向可参考，数值不可横向比较

Adobe、Shopify 与 Semrush 的样本、时间和指标不同，只能分栏理解。

【管理建议】Adobe、Shopify、Ahrefs、Semrush、Profound 等商业数据只用于发现假设。企业自己的投资决策必须回到本地基线、受控变化、可重复测试和销售/交易结果，不能把不同分母、时期和生态的数据拼成一个增长漏斗。

9. 对现有证据的总评

命题	当前证据支持程度	更准确的表述
AI 已成为重要信息入口	高	已进入大规模用户与主流搜索体验，但行业使用差异大
传统搜索即将消失	低	搜索结果页 AI 化、助手检索和传统搜索叠加存在
SEO 对 AI 搜索不再重要	低	索引、技术可访问和质量系统仍是重要底座
AI 答案会影响点击	中	多个观察显示相关性，但查询构成和方法影响结果
AI 引荐一定更高转化	低至中	部分商业生态观察到高意图，不能普遍外推
第三方提及有价值	中	成熟品牌样本中相关性较强，尚非因果或平台权重
内容表达能影响可见度	中	受控实验支持，但不等于开放平台长期商业效果
Agent 会改变交易	方向证据高、时间证据低	协议与产品正在形成，采用速度和收入贡献未知

第三部分 商业：GEO 的真实经营价值

10. 从网页排名到六种经营权

【作者框架】企业在 AI 入口中争夺的不是单一“排名”，而是六种经营权：

候选权： AI 是否把企业纳入可选集合；

定义权： AI 如何解释企业、产品和所属品类；

证据权： 企业信息是否支持回答中的关键主张；

比较权： AI 用哪些维度、与谁比较企业；

推荐权： 在何种场景、用户和约束下建议企业；

交易权： Agent 是否能查询、预约、购买、支付和完成服务。

【作者框架】可用一条概率链诊断商业结果：

AI 商业结果 = 可访问 × 可检索 × 可理解 × 可采纳 × 合适推荐 × 可行动 × 用户信任

该公式只用于定位瓶颈，不代表任何平台排名函数或可直接计算的财务模型。网站不可访问、内容无法进入候选、事实与场景不匹配、商品数据错误或用户不信任，都可能让最终价值接近零。

11. SEO 与 GEO：相同资产，不同目标

维度	SEO 主要关注	GEO 增加的经营问题
入口	搜索结果和网页排名	搜索内生成式回答、助手、Agent 和垂直入口
查询	关键词与搜索意图	多轮对话、角色、场景、约束和子问题
内容目标	进入索引并获得排名/点击	被准确采用、合适引用、纳入候选和推荐
数据对象	网页、链接、结构化数据	事实、证据血缘、商品/服务 Feed、API、权限
成果	曝光、点击、自然转化	答案份额、推荐语义、辅助转化、机器行动
风险	垃圾内容、处罚、流量波动	错误放大、投毒、虚假共识、自动行动损失

【管理建议】SEO 负责让企业进入可用信息池；GEO 在此基础上治理 AI 的表述、证据、候选、推荐和行动。组织上宜采用跨职能机制，不应停止 SEO 团队或另建一个只负责批量文章的孤岛。

12. 内容从关键词工厂转向证据资产组合

【作者框架】 AI 场景的内容规划单位应从“关键词”升级为“决策问题”：

决策阶段	典型问题	核心资产
认知	这是什么、为何需要	定义、机制、边界、行业地图
诊断	为什么出现问题	诊断清单、风险、适用条件
方案研究	有哪些路径	架构、方法比较、成本模型
品牌发现	哪些公司能做	客观候选标准、资质、覆盖与限制
比较	A 与 B 如何取舍	维度、方法、TCO、适用与不适用
证明	是否真实有效	原始数据、案例、实验、认证和独立验证
采购	如何招标和验收	RFP、POC、合规、SLA 与退出
交易	能否立即行动	SKU、价格、库存、预约、支付与售后

【管理建议】高可用资产至少应交代：直接结论；适用条件；原始来源；样本和方法；实体与版本；反例和限制；更新时间；作者责任；结构化字段；纠错入口。不是每篇文章都要机械套模板，但关键主张必须可追溯。

【作者框架】企业应建设六类互补资产：类别定义、场景解决方案、原创证据、比较决策、实体事实和机器行动。真正稀缺的是别人无法复制的事实、经验、数据和履约能力，而不是同义改写的文章数量。

13. 从引用测量升级为影响测量

【管理建议】指标应分三层：

领先指标用于确认能否进入信息链：抓取成功、索引覆盖、实体一致性、证据完整性、结构化数据有效率。

中间指标用于观察答案位置：主张采用率、可见引用率、事实准确率、候选进入率、推荐语义、风险提示完整率、来源独立性。

结果指标用于判断经营价值：品牌搜索、直接访问、AI 引荐、询价/预约、受影响机会金额、辅助转化、Agent 行动成功率、错误纠正成本和合规事件。

【管理建议】不把引用率作为单一北极星。引用可能是负面、限制或反例；采用也可能不展示引用。企业要分析“引用的语义角色”，并用品牌搜索、销售问卷、CRM 自报来源和时间序列补充传统点击归因。

14. Agentic Commerce：内容之后是数据与授权

【已验证事实】OpenAI 已把 Agentic Commerce Protocol 扩展到商品发现，允许商户共享产品源和促销等信息；Google 已发布 UCP 与 AP2，分别覆盖 Agent 商业流程互操作以及用户授权、交易意图和支付证明。（CLM-003、CLM-028、CLM-021）[OpenAI ACP](#) · [Google UCP](#) · [Google AP2](#)

从答案可见性到商品与支付：平台能力正在连接，但终局尚未确定

只记录已发布能力，不把协议存在写成规模化采用

非数值比较图



样本：平台与监管机构官方发布
时间：2025—2026 | 地区：全球平台能力 / 中国监管
单位：发布事件 | 分母：已公开发布的能力与政策，不含采用率
证据等级：A | 因果性：不适用；事件时间线
边界：协议、报告与监管行动的存在不等于统一标准、普遍采用或确定商业回报。
来源：CLM-002/003/013/021/022

从答案可见性到商品与支付：平台能力正在连接，但终局尚未确定

图表只表示公开发布与产品里程碑，不表示大规模采用或协议终局。

【趋势预测】购物和服务型 Agent 会把竞争分成两层：认知层由场景、比较、品牌、评论和可信度决定；交易层由产品唯一 ID、规格、价格、库存、配送、退换、权限、支付与订单接口决定。协议方向可信，规模化采用时点不确定。

【管理建议】企业应按四级准备机器行动数据：

- 可发现：名称、唯一 ID、类别、描述、图片和落地页；
- 可比较：规格、价格、币种、认证、评价、适用场景和差异；
- 可购买/预约：实时库存、交期、服务范围、税费、退换和商户身份；
- 可审计：用户授权、价格锁定、操作权限、订单状态、取消、退款和日志。

接入协议不是排名捷径。数据错误、履约不稳、评论失真或权限过宽时，Agent 只会更快放大问题。



Agentic Commerce 从发现到授权交易

AI 生成概念图，仅用于机制与场景表达，不作为事实证据。

第四部分 平台：中国与海外不能套用一套方法

15. 海外平台的公开机制与治理重点

入口	可确认机制	企业首要动作	关键边界
Google AI Overviews / AI Mode	搜索索引、质量系统、RAG、query fan-out	经典 SEO、可索引文本、独特内容、结构化数据、Merchant Center	不需要特殊 AI Schema；符合资格也不保证被采用
ChatGPT Search	查询改写、搜索合作来源、公开网页与抓取	放行 OAI-SearchBot、检查 WAF、维护事实页、跟踪引用	OAI-SearchBot 与 GPTBot 用途不同
ChatGPT 商品发现	商户 ACP 数据、公开商品与零售来源	完整 Feed、价格、库存、评价、履约和退货	接入数据不等于优先推荐
Claude Search / Research	定向查询、多来源读取、多轮调查	区分 Claude-SearchBot、ClaudeBot、Claude-User	搜索发现、训练和用户访问不能用一个总开关治理
Perplexity	搜索索引与用户触发抓取并存	区分 PerplexityBot/Perplexity-User、验证正文读取	来源偏好和展示会变化
Copilot / Bing	Microsoft 365 Copilot 可向 Bing 发送网页查询；声明式 Agent 可用 Bing 索引、限定公开站点或企业知识源	维护 Bing 可索引事实，并治理企业知识授权与管理策略	企业 Copilot 与 Agent 配置不能外推消费者排名（CLM-034）

【已验证事实】Microsoft 官方文档说明，Microsoft 365 Copilot Chat 可用发送至 Bing 的网页查询补充公开信息；声明式 Agent 可使用 Bing 索引，也可把搜索限定到最多四个公开网站，并通过 Copilot connectors 等接入组织数据。管理员可以关闭网页搜索，这些机制不证明消费者 Copilot 的固定来源偏好或企业商业效果。（CLM-034）[Microsoft 365 Copilot Chat · Agent 知识源](#)

【管理建议】robots 策略至少区分搜索索引、用户触发访问、模型训练和 Agent 行动。允许搜索而拒绝训练可能是合理组合；文本规则必须与 CDN、WAF、Bot Management、速率限制和访问日志一起验证。

16. 中国市场：多生态、低透明与平台内数据

【作者框架】中国不是一个统一索引，而是六类信息生态并存：

- 公开 Web 与通用搜索索引；
- 微信等超级 App 与私域内容；
- 短视频和内容推荐平台；
- 电商、本地生活、地图和履约数据；
- 政府、协会、标准、论文、专利与企业数据库；
- 模型自有搜索、合作来源和平台内部内容。

【作者框架】基于本报告截至证据截止日的公开资料梳理，多数中国 AI 产品对检索、重排和来源选择的技术披露少于 Google Search Central 等海外平台；这不是对全部产品的穷尽统计。公开机制不足不等于可以用经验传言代替证据，反而要求企业增加时间、地区、语言、账号和设备维度的重复实测。

【管理建议】中国官网应成为企业愿意承担责任的唯一事实源，至少包含法定主体和品牌关系、产品/版本/地域、功能与限制、价格口径、资质有效期、案例方法、更新时间、历史版本和纠错入口。平台内容负责触达和场景解释，不应与官网事实冲突。

16.1 中国渠道按证据职能分工

生态	主要职能	推荐资产	主要风险
官网和公开文档	标准事实、产品与原始证据	产品页、白皮书、政策、案例、FAQ、更新日志	自我声明冒充独立证明
公开搜索	网页发现问题候选	高价值场景、比较、专家研究和数据资产	堆词、镜像、批量同质页
微信生态	专业读者与客户关系	深度解读、客户教育、服务说明	假设所有内容都能被外部 AI 索引
短视频/内容平台	场景、演示和现场证据	实测、演示、专家解释、工厂/门店现场	虚构演绎和营销关系不披露
知乎/行业社区	复杂问题和观点竞争	选型方法、失败案例、专家答疑	伪装用户、自问自答网络
电商/地图/本地生活	商品、门店和履约事实	价格、库存、配送、地址、预约和售后	宣传文案覆盖实时字段
政府/协会/客户/伙伴	外部现实锚点	备案、标准、认证、联合案例、采购记录	控制或伪造“独立媒体”

16.2 中国平台实测

【管理建议】问题集至少覆盖品牌、品类、场景、比较、风险、地域、价格和售后八类。小型试点可从 40—80 个问题起步，大型企业可扩展到 100—300 个；这些是建议值，不是行业标准。（CLM-035）

每次测试记录：平台与模型/产品、账号状态、地区、设备、时间、是否联网、原始问题、完整回答、引用链接、截图或导出。相同问题跨日期重复，不把一次成功回答当成固定排名。

【管理建议】黑盒实验采用四步：建立发布前基线；一次只改变一类主要资产；等待合理抓取和索引窗口；复测品牌、品类、比较和风险问题，核对来源变化和失败样本。

17. 多语言、多国家和多版本

【作者框架】海外 GEO 的基本单位是“国家 × 语言 × 产品 × 用户任务”，不是把总部内容翻译成多种语言。每个市场的价格、货币、税费、库存、配送、认证、案例、法律声明和搜索习惯都可能不同。

【管理建议】使用稳定 URL、正确 hreflang 与 canonical；建立品牌、法人、型号、单位和认证的实体映射；为每个市场提供当地联系方式、履约和投诉路径；用当地语言原生问题测试；分别记录国家、语言、设备和账号条件下的答案与来源。

【管理建议】版本事实必须显式写出时间、地区和产品版本。若企业无法保证全球同步更新，应宁可显示“未知/需确认”，不要让 AI 把一个国家的旧价格、案例或资质外推到其他市场。

第五部分 风险：可信 GEO、投毒与监管

18.3 · 15 案例：检索污染，不是商业成功证明

【已验证事实】央视 2026 年 3·15 调查中，测试者虚构“阿波罗 9 号智能手环”及荒诞卖点，先发布两篇文章；约两小时后观察到一款 AI 能复述相关卖点并引用新文章。后续报道又记载三天内发布 11 篇内容，并在泛化推荐问题中观察到两款 AI 推荐该虚构产品。报道没有明确 11 篇是否包含最初两篇，因此不能简单写成 13 篇；本报告也不采用来源不清的多模型数量说法。（CLM-015）[央视视频](#) · [文字记录](#)

【已验证事实】截至 2026 年 3 月 20 日，公开报道为“公安部门已经启动调查”，不能升级表述为“立案侦查”。（CLM-016）[人民网](#)

【作者框架】该案例可以证明：特定新内容能在特定时间、平台和查询中影响检索增强回答；它不能证明所有模型都能以相同成本被操纵，不能证明基础模型永久训练，更不能证明正规 GEO 有稳定、可归因且 ROI 为正的商业效果。



可信 GEO 与检索投毒的边界

AI 生成概念图，仅用于机制与场景表达，不作为事实证据。

18.1 污染链路

【作者框架】一次虚构可以通过以下链路被放大：

虚构主张 → 多文体改写 → 多站点发布 → 搜索收录 → Top-K 召回 → 模型综合 → 展示引用 → 用户产生信任

五个薄弱点是：检索先优化相关而非完全验证真实；上下文容量有限；转载可能伪装独立共识；生成模型在缺乏现实锚点时会接受检索前提；引用本身会增强用户主观信任。

18.2 不同攻击层级

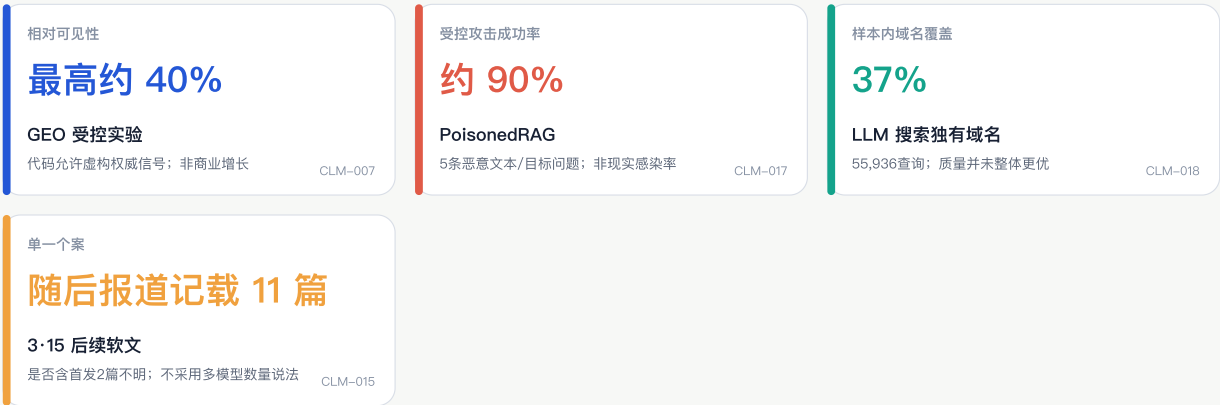
攻击层级	污染对象	典型影响	主要防线
训练数据投毒	预训练或微调语料与模型参数	生效慢、可能长期持续	数据供应链、清洗、训练安全评估
搜索/索引污染	公开网页和搜索结果	小时或天级，随索引变化	来源去重、反证检索、异常发布检测
RAG 知识库投毒	私有库、向量库、检索数据库	对特定问答持续生效	写入权限、文档签名、来源与版本治理
提示注入	网页或工具中的恶意指令	改变摘要、选择或 Agent 行动	内容隔离、指令层级、工具权限和沙箱
记忆投毒	用户或企业的长期记忆	跨会话影响未来判断	来源、写入批准、审计、过期和撤销

【已验证事实】PoisonedRAG 在包含数百万文档的受控知识库中，对每个目标问题注入 5 篇恶意文本，在其实验条件下达到约 90% 攻击成功率；该结果不能直接等同开放互联网成功率，但证明检索层是独立攻击面。
(CLM-017) [USENIX](#)

学术与案例证据回答的是不同问题，不能互相替代

不同口径，不可横比

可见性实验、投毒攻击、来源覆盖与新闻个案分别标注适用边界



样本：GEO-bench；PoisonedRAG实验；55,936查询；单一3·15新闻案例
时间：2024—2026 的不同研究与事件窗口 | 地区：实验环境 / 多平台 / 中国个案
单位：相对可见性 / 攻击成功率 / 域名占比 / 文章数 | 分母：四项分母完全不同，禁止横向排序或合成结论
证据等级：B | 因果性：仅受控实验可谈实验因果；观察研究与个案不构成普遍因果
边界：央视个案能证明特定外部内容可能影响特定检索增强回答，不能证明正规 GEO 的稳定 ROI。
来源：CLM-007 KDD；CLM-017 USENIX；CLM-018 arXiv；CLM-015 3·15

学术与案例证据回答的是不同问题，不能互相替代

图中数值仅对应各自实验条件，不外推为开放互联网或全部平台成功率。

19. “什么是真的”必须按命题类型判断

命题	例子	首选证据	AI 与企业应如何表达
客观事实	公司是否存在、价格、许可	工商/监管、官方产品、实时交易数据	可明确判断，并标注日期和地域
检测事实	性能、材料、安全、准确率	合格实验室、方法、样本和原始记录	说明测试条件与误差
市场统计	份额、用户量、增长率	可追溯数据、样本、口径和分母	说明估计与不可比口径
因果主张	使用后提升效率	对照、前后测、混杂控制	不用相关性替代因果
体验评价	易用、服务好	可验证用户与独立样本	表达分布和人群差异
预测判断	将成为第一、三年普及	假设、情景和敏感性分析	明确预测和失效条件

【管理建议】主张台账应给每项外部主张记录：原文、责任人、证据来源、方法、适用条件、禁用表述、有效期和高风险审批人。没有证据的主张不得因“AI 更喜欢明确答案”而写得更确定。

20. 监管与合规边界

【已验证事实】中央网信办 2026 年 4 月专项行动把训练语料管理、AI 数据投毒、利用 GEO 技术恶意营销、虚假信息和合成内容标识等纳入治理范围。(CLM-013) [中央网信办](#)

【已验证事实】中央网信办 2026 年 7 月公布的第一阶段精确口径为：处置违规网站、应用程序、智能体等 AI 产品 1.4 万余款；清理违法违规信息 600 余万条；处置账号 2.6 万余个；下架违规 AI 商品 1,300 余个；下架违规开源数据集 9 个。“AI 产品”与“AI 商品”不能混写。(CLM-014) [中央网信办](#)



中央网信办专项行动第一阶段：五类处置对象必须保持原口径

图表严格区分 AI 产品、AI 商品、账号、信息与数据集，不将专项行动总量误写成 GEO 单项执法。

【管理建议】中国企业至少建立以下控制：

按适用规则对 AI 生成或合成内容做显式/隐式标识；

不虚构产品、用户体验、专家身份、销量、榜单、认证或媒体报道；

不使用群控、刷量、控评或伪装独立来源制造共识；

高风险行业主张由合格人员与法务复核；

外包商交付账号、素材授权、数据来源、发布记录和删除证明；

建立纠错响应、升级和复盘时限。任何具体 24 小时、72 小时或 30 天时限均为企业建议值，不是法定期限。(CLM-037)

【已验证事实】美国 FTC 规则禁止制作、购买或销售虚假评论，并关注受控网站伪装独立评论来源等行为；欧盟 AI Act 第 50 条透明度义务涉及机器可读标记、部分深度伪造和公共利益文本披露，并有适用日期与例外。具体业务必须结合最新指南、司法辖区和法律意见，不能用本报告替代法律审查。(CLM-029、CLM-030) [FTC · EU AI Act 正式文本](#)

21. 可信 GEO 的治理原则

【作者框架】可信 GEO 的目标应从“让 AI 尽可能推荐”升级为：有充分真实证据时正确推荐；证据不足时表达不确定；信息冲突时展示边界；出现错误时可追溯、可纠正、可撤回。

【管理建议】建立七项控制：

来源血缘： 识别原始来源、转载和商业关系；

独立性去重： 十个 URL 若同源，只计一个证据簇；

原始证据： 高风险主张回到法规、实验、交易或认证记录；

反证检索： 主动寻找冲突、失败和不适用证据；

风险分级： 医疗、金融、法律、儿童、安全和交易提高门槛；

现实验证： 对产品存在、库存、价格、资质和门店连接现实系统；

责任审计： 保留审批、版本、发布、模型测试、纠错和删除记录。



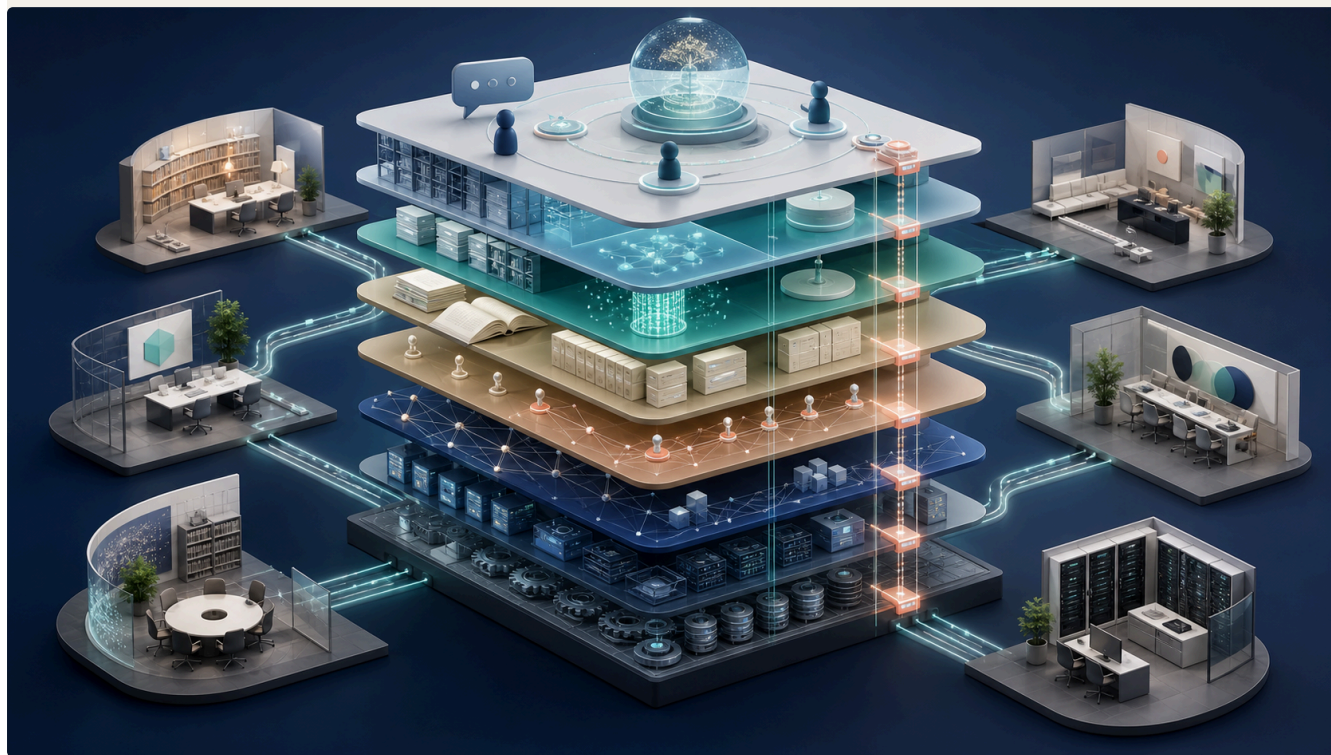
错误发现、复核、纠正与复盘闭环

AI 生成概念图，仅用于机制与场景表达，不作为事实证据。

第六部分 实施：企业级 GEO 操作系统

22. 八层能力模型

【作者框架】企业级 GEO 不是单项内容工作，而是八层相互依赖的能力。该模型用于识别短板，不是行业成熟度标准。



企业级 GEO 八层操作系统

AI 生成概念图，仅用于机制与场景表达，不作为事实证据。

22.1 第一层：需求与 Prompt 地图

【管理建议】从真实客户决策出发，而不是从企业想宣传的关键词出发。问题地图至少包含：

角色： 使用者、业务负责人、IT、安全、法务、采购、财务、管理层；

阶段： 认知、诊断、研究、发现、比较、证明、采购、交易、售后；

约束： 预算、规模、地域、行业、技术栈、风险、交期和监管；

意图： 信息型、比较型、风险型、品牌型、行动型；

价值： 是否影响候选、招标、预约、下单、续费或风险处置。

【管理建议】小型试点可从 40—80 个 Prompt 起步，大型企业可用 100—300 个建立分层基线。这些数量是操作建议，不是统计充分性标准。问题应来自销售录音、客服工单、站内搜索、招标文件、竞品比较、监管要求和客户访谈；每季度复核是否仍代表真实决策。（CLM-035）

22.2 第二层：技术可访问性

【已验证事实】内容不能被搜索索引、目标 crawler 或用户触发访问读取时，后续采纳概率会受限；但 crawler 访问成功也不代表已经进入索引、答案或推荐。

【管理建议】技术审计覆盖：

- robots、meta robots、canonical、站点地图和内链；
- HTTP 状态、重定向、JavaScript 渲染、正文文本层和 PDF 可解析性；
- CDN/WAF/Bot Management 的 User-Agent、IP、速率和地域策略；
- 移动端、桌面端和不同语言版本的一致性；
- 商品 Feed、结构化数据和页面可见内容的一致性；
- 搜索 crawler、训练 crawler、用户触发访问和 Agent 行动的分离策略；
- 访问日志中的成功率、正文大小、错误码、频率和关键页面覆盖。

22.3 第三层：企业实体与事实底座

【管理建议】建立企业愿意承担责任的唯一事实源，不以宣传首页代替事实数据库。最低实体包括公司、品牌、产品、版本、型号、功能、价格、服务、地域、资质、人员、案例、政策、门店和联系方式。

每个实体至少有：唯一 ID、标准名称、别名、主体关系、当前状态、适用地区、版本、生效/失效时间、事实责任人、原始来源和纠错入口。并购、改名、停售、价格变化、认证到期和地区差异必须保留历史版本。

【管理建议】为高价值主张建立“主张台账”：

字段	示例
主张	“产品支持某项能力”
性质	客观事实 / 检测结果 / 体验评价 / 预测
证据	产品文档、测试记录、客户授权、认证原文
条件	版本、套餐、地区、配置、日期
禁用表述	“所有场景”“永久”“零风险”等无法证明的绝对词
有效期	下一次产品发布或认证到期日
责任	产品负责人批准，法务复核高风险表述

22.4 第四层：证据资产

【作者框架】事实回答“是什么”，证据回答“为什么相信”。企业应把证据分为：

现实锚点： 工商、许可、备案、认证、标准和交易记录；

- 产品证据：** 原始技术文档、测试方法、性能数据和更新日志；
- 客户证据：** 有授权、有样本、有时间、有边界的案例与证言；
- 独立证据：** 专业测评、行业研究、伙伴验证和学术/机构资料；
- 反面证据：** 失败案例、限制、不适用场景和纠错记录。

【管理建议】证据不以数量计分，而以权威性、方法质量、外部适用性、时效性和独立性评估。转载同一新闻稿的十个链接只计一个证据簇；付费关系和共同所有权必须披露。

22.5 第五层：决策内容组合

【管理建议】围绕问题地图建设可复用资产，而不是“一问一页”批量生成。优先级通常为：

- 品类定义和适用边界；
- 角色/行业场景与问题诊断；
- 方案架构、实施路径和成本；
- 真实案例、方法和原始数据；
- 客观比较、替代方案与不适用条件；
- 采购、合规、安全、迁移和退出；
- 商品、门店、价格、库存、预约和售后。

一份内容应能同时回答直接问题、解释条件、展示证据和引导下一步。不得为覆盖长尾问法生成大量语义近似页面，也不得把机器可读性误解为短句、问答和关键词的固定模板。

22.6 第六层：第三方共识

【管理建议】第三方工作的目标是让真实业务事实在独立环境中得到验证，不是制造“全网都这么说”。企业应优先选择客户、行业机构、标准组织、政府、专业媒体、伙伴、真实用户和有方法的评测。

第三方计划至少记录：来源所有者、商业关系、原始证据、内容编辑权、披露方式、发布日期、可撤回性、纠错联系人和转载血缘。任何要求伪装普通用户、隐藏付费关系或批量复刻同一稿件的方案应被拒绝。

22.7 第七层：结构化数据与 Agent 行动

【管理建议】对可交易企业，内容治理必须与 PIM、商品 Feed、CRM、库存、订单、预约、权限和售后连接。关键控制包括：

- 唯一产品/服务 ID 与版本映射；
- 实时价格、币种、税费、库存、交期和服务半径；
- 规格、兼容性、认证、人群限制和退换政策；
- 用户授权、最小权限、价格锁定、操作确认和审计日志；

取消、退款、异常、人工接管与争议处理；

网页、Feed、API 和线下系统的一致性检查。

【趋势预测】随着 Agent 从阅读走向行动，“能否被调用”会成为与“能否被引用”同等重要的竞争层，但短期内不同协议、入口和地区仍会并存。

22.8 第八层：监测、实验与治理

【管理建议】建立跨平台版本化测试、来源血缘、异常告警、实验日志和纠错机制。监测不只抓成功结果，还要保留未出现、错误、负面推荐、无引用和行动失败样本。

治理会议至少回答：哪些关键事实错了；哪些证据过期；哪些来源同源；哪些问题进入/退出候选；业务是否受到影响；是否出现合规风险；哪些变化可归因；是否满足扩大或停止条件。

23. 企业是否需要做：八维诊断与路径选择

【管理建议】正确问题不是“做不做”，而是“先做哪一层、投入到什么程度”。以下八维诊断可按 0—3 分内部评估，但总分只是排序工具，不是统计验证的行业标准。(CLM-033)

维度	0 分	1 分	2 分	3 分
客户 AI 研究程度	几乎不用线上研究	偶尔问品牌	经常比较方案	AI 已进入采购/消费决策
决策链长度	即时低考虑	少量比较	多角色、多轮	数周至数月、多人审批
产品比较强度	同质且地域固定	少量参数	价格、功能、服务综合	需证据、合规和 TCO
线上行动比例	完全线下	线上获客	可预约/询价	可下单、续费或调用
错误信息风险	影响有限	增加客服成本	影响入围和品牌	可能产生监管/人身/财务损失
跨市场复杂度	单市场单语言	两种语言	多国家多语言	多法域、多货币、多产品线
证据成熟度	几乎没有	有官网资料	有案例、数据和第三方	事实、证据、接口可治理
监管敏感度	普通消费	一般广告约束	专业效果主张	医疗、金融、教育、法律等高风险

建议解释：0—7 分做事实防御；8—14 分做边界清晰的试点；15—19 分建立系统；20—24 分进入持续经营。高监管企业即使总分低，也应优先治理准确性、风险提示和纠错，而非追求推荐率。

【管理建议】扩大投入前同时确认：有明确业务问题；有可批准事实源；有至少一类独立证据；有连接销售、交易或风险结果的测量链。缺失其中一项时，优先补底座。

24. 90 天路线图

90 天、Prompt 数量、阶段交付和响应时限均为建议起点。企业应按抓取周期、审批、监管、国家数量和產品复杂度调整。(CLM-035、CLM-037)

24.1 第 0—14 天：建立真实基线

目标：回答“AI 现在如何描述我们”。

任务：选定一个品类和市场；建立 40—80 个高价值 Prompt；至少覆盖两个有业务意义的 AI 入口；重复测试；记录完整条件；访谈销售、客服和客户；扫描错误、缺席、负面和过期来源。(CLM-035)

交付：问题地图、平台基线、来源图谱、关键事实错误、商业事件定义、风险初筛。

退出门：如果客户几乎不使用相关入口、基线不存在明显风险或机会，维持低成本季度审计，不进入规模化内容阶段。

24.2 第 15—35 天：建设事实与证据底座

目标：让企业能够回答“我们愿意为哪些话负责”。

任务：建立实体字典、主张台账、唯一事实源；修复 robots、WAF、索引和正文可读性；核对第三方错误；标记过期内容；确定高风险主张审批；建立纠错入口。

交付：事实数据集、证据包、技术审计、权限矩阵、纠错和版本流程。

退出门：若产品和法务无法批准关键事实，停止扩大发布，先解决内部事实冲突。

24.3 第 30—60 天：发布决策资产

目标：用真实证据回答高价值问题。

任务：优先发布原创数据、方法透明的案例、场景与比较、成本模型、限制/不适用条件、产品文档和结构化字段。一次只改变一类主要资产；建立发布前后版本。

交付：少量高价值内容、原始证据下载、Schema/Feed 改进、实验记录。

退出门：如果内容只能提高品牌词表现、不能进入目标品类和场景问题，复核问题选择、证据相关性和平台可访问性。

24.4 第 45—75 天：形成独立佐证

目标：将自我声明转化为可验证的外部证据。

任务：争取真实客户、伙伴、标准、认证、行业机构和专业评测；披露商业关系；修正已有第三方错误；禁止同稿多站伪装共识。

交付：证据血缘图、独立来源清单、披露与纠错联系人、过期复核计划。

退出门：若所谓第三方证据主要由企业控制或无法披露关系，不把它计为独立证据。

24.5 第 60—90 天：复测、归因与投资决策

目标：回答“改变是否可重复、是否有价值、是否值得扩大”。

任务：按同一条件重复基线；比较准确、采用、引用、候选、推荐语义和行动；检查品牌搜索、直接访问、销售自报和 CRM；核算人力、工具、供应商和风险成本；复盘失败。

交付：前后对照、因果限制、成本与业务信号、扩大/维持/停止建议。

扩大门：至少两个核心指标跨日期和条件稳定改善，同时无重大事实或合规恶化。停止门：结果不可重复、只在单一截图成立、风险上升、边际成本失控或无法连接业务结果。

25. 六类行业实施路径

25.1 B2B 软件与专业服务

【管理建议】按业务、IT、安全、法务、采购和财务角色拆解问题。资产覆盖业务价值、架构、集成、安全、合规、实施、迁移、SLA、TCO、合同和退出；每项能力写清版本和套餐。

核心指标：非品牌品类问题入围率、关键事实准确率、独立证据覆盖、销售中客户主动提及 AI 研究的机会占比、受影响机会金额和纠错时长。

25.2 工业制造与出海

【管理建议】以型号、参数、公差、材料、认证、产能、交期、服务地域、适用工况和工程案例为主。治理经销商旧页面、跨国家型号、计量单位和认证有效期；PDF 图纸和目录应有文本层与版本。

核心指标：型号检索准确率、国家站一致性、技术文档覆盖、RFQ 辅助率、错误型号推荐和过期认证数量。

25.3 消费品与电商

【管理建议】内容、商品数据和履约同等重要。围绕预算、场景、尺寸、材料、兼容性、人群限制、评价、售后和替代品组织信息；同步价格、库存、配送、退货和商户身份。

核心指标：品类候选率、商品字段完整率、价格/库存错误率、AI 引荐加购、辅助转化和退货原因。

25.4 本地服务与 O2O

【管理建议】地址、营业时间、服务半径、预约空档、门店主体、资质和真实评价优先于全国性软文。地图、平台、官网和线下系统保持一致。

核心指标：地域问题可见率、错误门店信息、电话/预约成功率、无效线索和纠错时长。

25.5 医疗、金融、教育、法律等高监管行业

【管理建议】目标顺序是准确、证据、边界、风险提示，再考虑可见度。效果、收益、疗效、适用人群和资格必须由专业人员与法务审核；不得用虚构病例、收益或学员结果影响推荐。

核心指标：高风险主张审定覆盖、错误答案率、风险提示完整率、投诉、监管事件和撤回时长。AI 回答不能替代个体化专业意见。

25.6 媒体、研究与知识机构

【管理建议】同时治理引用价值和内容许可。区分搜索发现、用户访问、训练许可和商业授权；探索摘要、全文、API、数据集和订阅的不同产品层。

核心指标：署名引用率、正确归因率、机器访问成本、引荐/订阅、授权收入和未授权使用处置时长。

26. 组织、预算、TCO 与 ROI

26.1 企业内部不能外包的权力

【管理建议】企业内部必须保留：事实批准权；合规批准权；数据、账号和日志所有权。服务商可以提供研究、内容、技术、监测和执行，但不能替企业决定产品是否真的具备某项能力，也不能垄断提示词、原始答案和发布账号。

26.2 建议 RACI

工作	负责 R	最终批准 A	协作 C	知会 I
问题地图与优先级	GEO/市场研究	业务或 CMO	销售、客服、产品	管理层
公司与产品事实	产品运营/知识管理	产品负责人	销售、客服、法务	市场
高风险主张	业务/产品	法务或合规	专业人员、市场	管理层
技术可访问性	SEO/工程	数字负责人	安全、数据、法务	市场
内容与证据	内容/研究	业务负责人	产品、客户、法务	销售
商品/服务数据	电商/PIM/产品	业务负责人	工程、财务、履约	市场
监测与实验	数据/GEO	业务负责人	SEO、销售、产品	管理层
纠错与事件	GEO/客服	法务/品牌负责人	产品、平台、供应商	管理层

26.3 人员模型

【管理建议】试点可由一名跨职能负责人协调研究/内容、SEO/工程、产品知识、数据分析和法务支持；系统化运营再增加国家、行业、商品数据或 Agent 接口负责人。具体人数取决于产品线、语言、监管和内部成熟度，不能套用固定编制。

26.4 总拥有成本

【作者框架】TCO 至少包括：

内部人力 + 研究与内容 + 技术与数据 + 监测工具 + 第三方证据 + 合规审核 + 供应商 + 纠错/删除 + 机会成本

【管理建议】预算不按统一固定比例强制切分。企业应分别核算事实和证据、技术数据、实验监测、第三方工作与合规；若大部分预算流向批量发布而事实、技术和测量投入不足，项目结构通常失衡。

26.5 ROI 与决策规则

【作者框架】GEO 价值由增长、效率和风险三部分构成：

$$\text{净价值} = \text{受影响毛利} + \text{获客/研究效率收益} + \text{避免的错误与合规损失} - \text{TCO}$$

该公式是管理框架，不是可直接套用的财务标准。受影响毛利必须明确归因方法；风险避免值应用情景范围，不应凭空制造精确数字。

【管理建议】扩大条件包括：准确性改善；非品牌候选进入稳定改善；推荐语义符合真实适用场景；品牌搜索、销售或行动出现可解释信号；错误和行动失败率下降。停止条件包括：仅品牌词变化；无法重复；引用增加但错误加重；边际成本上升而业务无信号；供应商不可审计；风险大于收益。

27. 指标字典与实验设计

27.1 指标必须有明确分母

指标	建议定义	必须记录的分母/边界
抓取成功率	成功获取正文的目标页面数 / 被测试页面数	crawler、时间、状态码、正文阈值
候选进入率	品牌进入合适候选的问题数 / 适用问题数	排除不适用问题，记录平台与重复次数
可见引用率	展示企业来源的回答数 / 有回答的目标问题数	引用可见不等于被采用
主张准确率	正确且边界完整的关键主张 / 被检查主张	主张级，不以整篇回答笼统计分
推荐语义合格率	推荐理由、适用/不适用和风险均正确的问题数 / 推荐问题数	由双人或盲审编码
来源独立率	独立证据簇 / 引用来源簇	按所有权、原始稿和商业关系去重
AI 引荐转化	AI 引荐完成目标事件会话 / 可识别 AI 引荐会话	只覆盖可点击流量，不能代表零点击影响
受影响机会金额	销售明确记录 AI 参与研究的机会金额	自报偏差、时间窗口和多触点归因
Agent 行动成功率	正确完成授权行动 / 合法目标行动	价格、库存、权限、取消和异常分别记录
纠错时长	从确认错误到各渠道修正的时间	发现、确认、发布和传播更新时间分开

27.2 最小实验协议

【管理建议】一次实验至少记录：假设；目标平台和模型；问题集与价值权重；账号、地区、设备、语言、是否联网；基线时间；改变的唯一主要资产；等待窗口；重复次数；原始回答与引用；盲审规则；商业观察窗口；失败和停止条件。

【管理建议】测试避免五类错误：

- 同时发布大量内容、做公关、投广告和改网站，导致无法归因；
- 只挑成功 Prompt，删除失败或负面答案；
- 把一次运行当固定排名；
- 改写问题后仍声称与基线同条件；
- 用平台可见度变化直接宣称收入因果。

27.3 样本与统计边界

【管理建议】40—80 个 Prompt 适合管理试点和发现问题，不自动满足统计推断。需要估计小幅效果、比较多个平台或向董事会承诺量化提升时，应由数据人员根据基线比例、最小可检测效应、重复相关性和多重比较计算样本量。(CLM-035)

高价值问题可以按业务价值加权，但加权规则必须在测试前固定；人工编码应提供定义、双人复核和分歧处理；平台版本更新、地区变化和突发新闻应记为外部事件。

27.4 反向验证

【管理建议】不仅问“怎样提高出现”，还应主动测试：不适用人群是否被正确排除；产品限制是否保留；旧版本是否仍被引用；竞品和负面来源是否有合理证据；删除/纠正后错误是否消失；不同平台是否给出冲突结论。没有反向验证的推荐率优化容易把风险当成果。

28. 服务商 RFP、POC 与合同验收

28.1 RFP 必答问题

【管理建议】要求服务商以原始材料回答：

- 如何定义 GEO，哪些环节可控、不可控；
- 覆盖哪些平台、地区、语言、账号和设备；
- 如何构建 Prompt 集，是否保留失败样本；
- 如何区分被检索、采用、引用、推荐和转化；
- 如何追溯来源、同源转载、付费关系和内容所有权；
- 如何证明内容、数据、案例和评论真实；
- 使用哪些账号、站点、工具、crawler 和发布网络；
- 数据和账号归谁，合同终止后如何完整导出与删除；
- 如何处理平台更新、错误答案、投毒、投诉和监管事件；
- 哪些结果绝不承诺，哪些假设会导致项目停止。

28.2 建议评分

【管理建议】可按以下建议权重设计 POC 评分，企业应自行校准：证据与合规 25%；测量和可复现 25%；技术与数据能力 20%；内容与行业能力 15%；团队与治理 10%；价格 5%。价格不宜成为第一权重，因为低成本批量发布可能显著增加纠错和合规成本。（CLM-036）

28.3 POC 必须交付

【管理建议】POC 至少交付：冻结的问题集；基线与完整原始回答；测试条件；技术审计；主张/证据样例；发布资产和变更日志；前后对照；失败样本；来源血缘；商业和风险指标；导出文件；扩大和停止建议。不得只交付引用截图、品牌出现次数、无法复跑的仪表盘或缺少分母的“提升率”。

28.4 合同条款

【管理建议】SOW 明确范围、平台、语言、问题集、交付格式、责任和变更；数据条款明确账号、Prompt、回答、日志、素材和模型输出归属；真实性条款禁止虚构和未披露关系；审批条款规定高风险内容的人类批准；纠错条款规定发现、确认、修正、升级和删除；退出条款要求账号移交、数据导出、内容清单和删除证明。

【管理建议】任何 24 小时、72 小时或 30 天 SLA 仅是企业协商建议，不是法定期限。合同不得要求服务商保证平台排名、推荐率或不受平台更新影响。（CLM-037）

29. CEO 与董事会的 90 天议程

【管理建议】 董事会不应讨论“要不要追热点”，而应批准六个具体问题：

- 业务问题： 优先解决事实风险、候选缺席、比较失真还是交易断点；
- 试点边界： 一个品类、市场、平台组合和冻结问题集；
- 事实责任： 谁批准公司和产品事实，谁批准高风险主张；
- 红线： 禁止虚假证据、站群共识、隐藏关系和删除风险信息；
- 指标： 准确、候选、商业、风险和成本共同衡量；
- 退出： 何时停止扩张，如何撤回和纠错。

第 30 天审阅事实与风险基线；第 60 天审阅证据、技术和内容交付；第 90 天审阅重复测试、成本、商业信号和风险。若证据不足，董事会应批准“继续小规模验证”或“停止”，而不是用主观热度替代结果。（CLM-035）

第七部分 最终判断

30. 六项研究判断

判断一：GEO 是真实命题，但广义定义属于作者框架

【已验证事实】生成式回答、助手检索、平台引用和 Agent 购物已经存在，企业信息会在新入口中被发现、重述和使用。

【作者框架】把这些工作统称为“可信信息基础设施”是本报告的经营框架，不代表行业已经形成统一定义。企业不应围绕名词争论，而应明确要解决的事实、候选、推荐、行动和风险问题。

判断二：SEO 是地基，但业务目标已经超出点击

【已验证事实】搜索索引、技术可访问、质量系统和优质内容仍是重要底座；平台没有公开特殊“GEO 排名公式”。

【管理建议】继续做好经典 SEO，同时增加答案准确、候选进入、推荐语义、结构化数据、Agent 行动与纠错治理。

判断三：原创证据和事实一致性比内容数量更可持续

【观察性关联】成熟品牌监测提示第三方提及、视频存在和品牌共识与 AI 可见度相关，但不是因果权重。

【管理建议】把投资重点放在真实产品、原始研究、可验证案例、独立佐证、更新机制和机器可读数据。页面数量和批量分发只在真实问题与证据支持下有价值。

判断四：3 · 15 证明风险面，不证明正规 GEO 的 ROI

【已验证事实】虚构内容可以污染特定检索增强回答；监管和安全研究也确认检索、知识库和内容供应链存在攻击面。

【管理建议】可信 GEO 必须把来源血缘、反证、独立性、现实验证、纠错和责任审计写进流程与合同。

判断五：商业价值存在，但公开数据不能代替企业实验

【观察性关联】零点击压力、AI 引荐增长和部分高意图转化在不同研究中出现；这些样本、分母和时期不一致。

【管理建议】企业用本地 Prompt、销售、CRM、商品和风险数据决定投资，不使用不同厂商数据拼接统一 ROI。

判断六：Agentic Commerce 方向可信，时间和格局不确定

【已验证事实】ACP、AP2 等协议与商品发现产品已公开。

【趋势预测】机器将更多参与比较、授权、支付和履约，但规模化采用可能慢于市场宣传，且协议会并存。

【管理建议】优先建设可跨渠道复用的商品/服务事实、实时数据、权限、日志和履约能力，再按真实流量与转化接入入口。

31. 最终建议

【管理建议】企业现在应做的不是全面押注 GEO，而是完成五个最小动作：

建立公司和产品唯一事实源；

用高价值 Prompt 建立跨平台基线；

给关键主张绑定原始证据、责任和有效期；

修复错误、过期、不可访问和不可行动的信息；

做一个有对照、有失败样本、有停止条件的 90 天试点。

【作者框架】GEO 最终不会只是“营销部如何写文章”，而是企业能否把真实能力表达成机器可发现、可验证、可比较、可授权和可纠错的信息系统。能被 AI 看见不是终点；值得被推荐、适合被行动、出现错误时能够负责，才是长期竞争力。

附录

附录 A：证据记录最低字段

每条关键主张进入结构化证据台账：

字段组	必填内容
身份	主张 ID、主张文本、主张类型、负责人、状态
来源	原始链接、发布机构、标题、发布日期、核验日期、存档
口径	指标、单位、分母、样本、地区、时间范围、产品版本
方法	数据来源、采样、实验/观察方法、比较组和计算方式
边界	因果属性、外推限制、已知偏差、冲突证据
评级	来源权威性、方法质量、外部适用性、时效性
生命周期	生效日、到期复核日、替代/撤回记录
使用	对应正文、图表、简报和公开页面

【管理建议】证据评级不得只看来源名气。平台官方最适合证明产品如何自述工作，不天然证明商业效果；厂商大样本可发现方向，不自动具有因果和普适性；学术实验若环境受控，也不能直接外推开放互联网。

附录 B：Prompt 基线样例

以下仅为结构示例，企业应替换为真实客户语言：

品类认知：“什么是 X？它解决哪些问题，哪些场景不适合？”

场景诊断：“一家 500 人、跨三地的企业出现 Y，可能有哪些原因？”

方案比较：“方案 A、B、C 在成本、安全、实施周期和退出方面如何取舍？”

品牌发现：“在中国/某地区有哪些可提供 X 的企业？列出筛选依据和来源。”

竞品比较：“品牌 A 与 B 在指定版本、地区和预算下有哪些差异？不要推测未知信息。”

证据验证：“某厂商声称提升 40%，原始样本、方法、适用条件和独立验证是什么？”

风险：“购买 X 最常见的失败、合规和锁定风险是什么？哪些企业不适合？”

采购：“为 X 制定 RFP、POC 和验收清单，区分必须项和建议项。”

行动：“在预算、地区、库存和交期约束下列出可购买候选；信息过期时明确提示。”

纠错：“关于品牌 A 的价格、资质和产品能力有哪些来源冲突？以最新原始来源为准。”

附录 C：主张真相卡模板

项目	填写要求
标准主张	一句可承担责任的话
主张类型	客观、检测、市场统计、因果、体验或预测
适用条件	产品、版本、地区、人群、时间、配置
原始证据	不用二次转载替代
方法/样本	数据怎么来、多少样本、是否有对照
可用表述	对外允许的准确语言
禁用表述	绝对化、扩大化或缺证据语言
反例/限制	不适用、失败和风险
负责人	业务、产品、法务和更新责任
有效期	到期复核或触发条件

附录 D：核心一手与研究来源

平台官方

- Google：AI 搜索内容优化与技术要求
- Google：Search Console 生成式 AI 表现报告
- OpenAI：ChatGPT Search
- OpenAI：ACP 与商品发现
- Anthropic：Research
- Perplexity：Crawlers
- Google Cloud：AP2

学术与独立研究

- GEO: Generative Engine Optimization
- Source Coverage and Citation Bias in LLM Search
- PoisonedRAG
- Attribution Crisis in LLM Search
- Pew：AI Overview 与点击行为

市场、用户与商业观察

- CNNIC：第 57 次《中国互联网络发展状况统计报告》
- QuestMobile：AI 应用行业报告
- SparkToro：2026 零点击观察
- Ahrefs：品牌可见度相关性
- Ahrefs：LLM 引用与新鲜度
- Shopify：AI Search Insights
- Adobe Analytics：AI 引荐流量

监管与案例

- 中央网信办：2026 年 AI 治理专项行动
- 中央网信办：专项行动第一阶段数据
- 央视 3·15 晚会案例视频

央视财经报道文字记录

人民网：公安部门启动调查

FTC：消费者评论与证明规则问答

EU AI Act：欧盟正式文本

附录 E：署名、AI 使用、利益冲突与纠错

本报告由吴兆华以个人研究身份发布，不代表 e签宝或任何任职、合作机构的正式立场。作者可能基于职业经验关注企业服务、电子合同、B2B 销售和数字化场景；案例用于解释方法，不构成对相关公司、产品或服务的正式评价。

Codex 用于资料整理、交叉核验、结构化写作、图表和视觉辅助。所有关键事实应以配套证据台账和原始来源为准。AI 生成概念图均需标注“AI 生成概念图，不作为事实证据”；数据图表不得由生成式图像承担数值表达。

报告证据截止 2026 年 7 月 12 日。AI 搜索、平台规则和监管变化较快；动态主张应设置到期复核日。发现错误后执行：登记问题；冻结相关主张和图表；回到原始来源复核；评估受影响页面；同步更新网页与两份 PDF；记录旧值、新值、原因、日期和责任人。重大错误应在公开页面展示更正说明，不以静默覆盖替代修订记录。

研究与责任声明

报告属性

吴兆华个人署名研究，不代表 e签宝正式立场，不构成法律、投资或平台排名承诺。

AI 使用

Codex 参与资料整理、结构校验、写作协作、制图和技术构建；关键事实由原始来源和证据台账约束。

概念图片

AI 生成图片仅用于解释机制与场景，不作为事实证据；数据图表由结构化数据确定性生成。

更新原则

平台、协议和监管口径快速变化；动态主张设置复核日期，读者应以最新官方来源为准。