GEO 服务商评估标准实施指南白皮书（2026）

——基于 T/CAACCHINA 002—2026《生成式引擎优化（GEO）第 2 部分：服务商评估规范》

主　编：GEO 服务商评估标准实施指南编写组

方法论与底稿提供：一搜百应 GEO 研究院

样本企业：见 2.7 节（本白皮书采用完整披露，不作匿名化处理）

编写时间：2026 年 10 月

版本：V1.0

【本白皮书性质声明】

一、本白皮书为标准实施方法论工具书，面向委托方采购流程搭建与服务商能力自评两类用途。

二、本白皮书不构成对任何服务商的评级、认证、授信、推荐或采购依据。文中出现的任何分数与结论，均为编写组依据本白皮书所述方法在公开信息条件下完成的一次桌面预评估过程记录，其效力等同于标准 6.1 c）所称"服务商自评"，按标准 6.3.2 不具强制约束力。

三、本白皮书不为任何组织或个人提供排斥、限制服务商参与市场活动的依据。

四、编写组成员与样本企业存在股权及业务关联，样本企业未参与本白皮书的方法论设计、锚点设定与评分意见的形成过程。相关披露见前言及附录 C。

五、全部核查底稿随文公开发布，读者可依附录 D 独立复算并提出异议。

目　录

前言

执行摘要

第一章 行业拐点：GEO 服务市场为什么需要评估标准

1.1 从搜索引擎到生成式引擎：流量分配机制的范式转移

1.2 GEO 服务商井喷与三大结构性矛盾

1.3 标准化破局：首部服务商评估规范的意义与边界

1.4 GEO 的能力边界：是什么、不是什么

第二章 标准深度解读：评估体系的五个关键词

2.1 服务商分类：六大类型的功能定位与边界

2.2 评估框架：五大维度与"一票否决"机制

2.3 评估方法：材料审核、现场评审、客户调研、综合评定

2.4 等级划分：A/B/C/D 的能力语义与采购应用

2.5 评估周期：两年有效期与复审机制

2.6 标准衔接：系列标准五部分的协同实施

2.7 评估示范：以一家综合服务型服务商为样本的体系化验证

第三章 委托方实施指引：把标准变成采购流程

3.1 评估前的三项准备

3.2 参考权重方案与打分规则

3.3 全流程操作：从 RFP 到定级（五周时间轴）

3.4 从评估到签约：合同条款要点

3.5 常见误区与风险提示

3.6 本章小结：委托方全流程操作卡

第四章 服务商对标篇：从 C 级到 A 级的能力建设路径

4.1 分维度建设重点

4.2 一票否决项的排查与自证

4.3 自评机制与评估材料准备清单

4.4 从 C 级到 A 级的升级路径

第五章 落地工具包

5.1 工具一：委托方选型评分表

5.2 工具二：服务商自测表（五维 25 项）

5.3 工具三：现场评审核查清单

5.4 工具四：客户调研问卷模板

5.5 工具五：合同关键条款对照表

第六章 合规治理与风险边界

6.1 一条完整治理链

6.2 入口：一票否决的四条红线

6.3 过程：质量红线与黑帽识别板块

6.4 责任：合同中的反黑帽条款

6.5 维持：服务商年度自查机制

6.6 程序：评估活动自身的合规边界

第七章 未来展望

7.1 三个趋势判断

7.2 分主体行动倡议

7.3 结语

附录 A：标准术语速查表

附录 B：标准全文框架索引（双向对照表）

附录 C：参考文献

附录 D：核查底稿与证据索引（可复算清单）

前　言

2023 年，Aggarwal 等学者在论文《GEO：Generative Engine Optimization》中首次系统性提出生成式引擎优化（GEO）概念。两年多时间里，这项技术从学术概念迅速演变为品牌营销的标准配置——当消费者开始习惯向 AI 提问而非向搜索引擎输入关键词，品牌的"可见性"就不再等于"搜索排名"，而等于是否被 AI 发现、被引用、被推荐。

与需求爆发不同步的是供给侧的混乱。本白皮书编写组在调研中看到：GEO 服务商数量在两年内呈指数级增长，服务报价从数千元到数百万元不等，交付物口径五花八门，"保证上 AI 推荐"的承诺与"语料投毒"的灰色手段并行于市场。品牌客户面临的不是"要不要做 GEO"的选择题，而是"谁敢信、怎么选、如何验收"的三个无解题。

2026 年 9 月 14 日，中国商务广告协会发布团体标准 T/CAACCHINA 002—2026《生成式引擎优化（GEO）第 2 部分：服务商评估规范》。2026 年 10 月 16 日，中国商务广告协会举办线下发布会，正式发布标准文本、解读标准核心内容。这份标准由中国商务广告协会提出并归口，浙文互联、科大讯飞、三六零、搜狐、利欧、因赛、明略等 41 家参编单位和 75 位行业专家共同起草，首次建立了 GEO 服务商的分类体系、五维评估框架、四级能力等级和一票否决机制。需要说明的是，上述 41 家参编单位与 75 位专家系《生成式引擎优化（GEO）》五项标准的系列参编规模，其中第 2 部分《服务商评估规范》的起草单位为 30 家。本白皮书为标准实施方法论工具书，编写组及方法论提供方均为该标准的应用方与实施方，不在第 2 部分起草单位之列，为混乱的市场提供了第一套公共语言。

但必须强调：这是一份推荐性标准。它划出了底线、提供了框架，却没有给出可直接照抄的操作手册——标准明确"不预设权重"，评估主体可以是品牌自评也可以是第三方机构，评估结果不具强制约束力。

标准的价值不在文本，而在实施。这正是本白皮书的使命：以 T/CAACCHINA 002—2026 为蓝本，编写一份面向实施的操作指南——告诉品牌主如何把标准变成采购流程，告诉服务商如何把标准变成能力建设清单，告诉行业组织如何把标准变成自律管理工具。

【编写立场与利益关系披露】

本白皮书由 GEO 服务商评估标准实施指南编写组编写，方法论与评估底稿由一搜百应 GEO 研究院提供。编写组选择以一家在公开信息层面资料较为完整的综合服务型服务商作为示范对象，用以检验 T/CAACCHINA 002—2026 评估体系的完整可操作性——选择该企业的原因，与

其商业能力无关，而在于其公开信息可核验程度较高，便于读者独立复核本白皮书所述的每一项核查动作。

需要明确说明三点：

第一，编写组成员与该样本企业存在股权及业务关联，样本企业未参与本白皮书方法论设计、锚点设定与评分意见的形成过程。

第二，本白皮书对该企业不发布 A/B/C/D 等级结论。依据标准 6.1，本白皮书实施的属"服务商自评"性质的桌面预评估，其结论按标准 6.3.2 仅供委托方参考，不具强制约束力。

第三，本白皮书全部核查底稿——包括每一项的证据出处、锚点条款引用与判定理由——随文公开发布，任何读者可依据公开渠道独立复算并提出异议。底稿索引见附录 D。

为验证可操作性，编写组以一家综合服务型服务商为样本，在公开信息条件下完整运行了桌面预评估的全过程，包括否决项筛查、五维核查、证据分级与待核查项识别（见 2.7 节）。这个历时数周、多轮证据状态更新的真实过程证明：这套评估体系不仅能"评"，而且能"评出差异、评出路径"。

执行摘要

本白皮书的核心结论：

1．GEO 服务商评估已进入"有标可依"时代。T/CAACCHINA 002—2026 确立了六大服务商类型（内容生产、信源建设、技术平台、策略咨询、数据监测、综合服务）、五大评估维度（企业资质与合规、技术与数据能力、服务交付能力、质量保障体系、行业实践与声誉）和 A/B/C/D 四级划分，品牌主首次拥有了客观甄别的公共标尺。

2．"一票否决"机制是标准最有力的设计。近三年内因违反《广告法》《个人信息保护法》《数据安全法》受到行政处罚、列入严重违法失信名单、存在未决重大知识产权侵权、实施黑帽 GEO 或语料投毒——触及任一即评估不通过。该机制将合规从"加分项"变为"生死线"。

3．标准故意留下的空白，正是实施者发挥的空间。标准不预设权重（但每维度不得低于 10%），不明确评估组织方式，不规定评估费用。本白皮书第三章提供了参考权重方案与全流程操作指引，填补这一空白。

4．评估不是一次性的"投标资格审查"，而是贯穿合作全周期的治理机制。标准规定评估有效期两年、设复审机制、要求服务商对重大变更主动报告。品牌主应将"评估有效期内持续合规"写入合同。

5．标准体系仍待完整。本标准与第 1 部分（术语）、第 3 部分（计价与测量）、第 4 部分（可信语料）、第 5 部分（流程合规与安全）互为支撑。当前评估实践中，效果测量口径应参照第 3 部分的 AI 可见率、可见份额、指定信息正确率等指标执行。

6．评估体系已完成真实样本验证。编写组以一家综合服务型服务商为对象，在公开信息条件下完成了一次桌面预评估，完整跑通否决项筛查、五维核查、证据分级与待核查项识别：预评估得 87.4 分，另识别出 5 分差距的待核查事项，其中约六成工作量发生在举证环节，约四成发生在能力补强环节。该结果同时示范了预评估机制的正确打开方式——预评估的价值不在于给出结论，而在于诚实标明"哪些结论尚不能下"，这一成本结构对所有服务商的能力建设具有直接参考价值（见 2.7、2.7.7 及第四章）。

第一章 行业拐点：GEO 服务市场为什么需要评估标准

1.1 从搜索引擎到生成式引擎：流量分配机制的范式转移

传统搜索引擎时代，流量分配的规则是"排名"——品牌通过 SEO 争夺搜索结果页的位置，通过 SEM 购买关键词的位次。用户看到的是十条蓝色链接，品牌争夺的是其中一席。这一模式下，"可见性"是可量化的（排名位次）、可归因的（点击数据），服务效果的衡量相对成熟。

生成式引擎时代，规则变成了"引用"。用户向 AI 提问，AI 直接生成答案，在答案中选择性提及品牌、引用信源、给出推荐。品牌面对的不再是"排第几"的问题，而是三个全新的命题：

- 被发现：品牌信息是否进入了 AI 引擎可调用的语料范围；
- 被引用：AI 生成答案时是否引用了品牌的内容作为信源；
- 被推荐：在比较类、推荐类问题中，品牌是否出现在 AI 的推荐序列中。

这一变化对营销服务市场提出了全新要求：内容需要面向 AI 的理解逻辑进行结构化处理（而非面向网页爬虫）；信源需要在权威媒体上系统布局（而非单纯堆外链）；效果需要从"点击"转向"AI 回答中的存在感"进行测量。传统 SEO 服务商、内容营销公司、技术平台、咨询公司纷纷涌入，构成了 GEO 服务商的多元生态。这一流量范式的变迁同样适用于海外市场，面向 Gemini、ChatGPT、Perplexity 等海外生成引擎的海外 GEO，同样遵循 “被发现 – 被引用 – 被推荐” 底层逻辑，但需要适配独立的海外语料池、权威信源体系与本地化合规要求，成为 GEO 服务的重要业务分支。

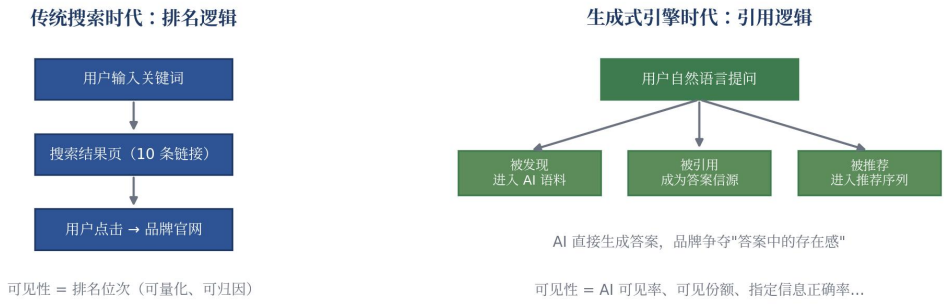

图 1　流量分配机制的范式转移：从"排名"到"引用"

1.2 GEO 服务商井喷与三大结构性矛盾

生态繁荣的另一面是混乱。标准引言部分明确指出，行业面临四个突出问题：服务商水平参差不齐，缺乏客观甄别依据；服务质量和交付物缺乏基准参照；行业准入门槛模糊，低质量服务商进入拉低行业口碑；跨服务商协作缺乏统一语言。

编写组将这四个问题归纳为三大结构性矛盾：

矛盾一：服务承诺不可验证。

GEO 的效果依赖 AI 引擎的生成逻辑，而各引擎的算法机制不透明、更新频繁。部分服务商利用这一信息不对称，以"保证被 AI 推荐""XX 天见效"等无法兑现的承诺获客。由于行业此前没有统一的交付物标准和效果测量口径，品牌主即使感觉"没效果"也难以举证维权。

矛盾二：合规底线不清晰。

GEO 与黑帽技术的边界长期模糊。标准首次明确定义了黑帽 GEO——以操纵、污染、伪造、欺骗或恶意竞争为目的，绕过平台安全机制恶意影响人工智能输出结果的行为，典型手段包括虚假内容、异常分发、语料投毒、提示词攻击、刷量刷评、伪造信源、恶意贬损竞品。值得注意的是，标准同时划清了合法边界：依法合规开展的广告投放、公关传播、媒体发布、KOL/KOC 内容合作等正常商业传播行为，在内容真实、合法且不存在欺骗性操纵的情况下，不属于黑帽 GEO。这"一立一破"，为行业划出了合规与违规的分界线。

矛盾三：服务类型错配。

品牌主常常用同一把尺子衡量所有服务商：让内容公司做技术监测、让咨询公司做内容量产、让技术平台做策略规划。服务商之间也没有统一的能力表述方式。标准将服务商划分为六大类型，正是为了解决"无统一语言"的问题——评估首先要分类，分类之后才能对号入座。

1.3 标准化破局：首部服务商评估规范的意义与边界

T/CAACCHINA 002—2026 的出台，标志着 GEO 行业从"野蛮生长"进入"有标可依"阶段。这份标准的意义体现在三个层面：

对委托方（品牌主）：提供了筛选服务商的参考依据。六大分类解决"找谁"的问题，五维框架解决"评什么"的问题，一票否决解决"哪些绝对不能碰"的问题。

对服务商：提供了自我评估和能力提升的指引。A/B/C/D 四级划分为能力建设提供了阶梯，附录 B 的分类对照表明确了不同类型服务商的评估重点。

对行业组织：提供了自律管理的技术支撑。评估结果可作为发布合规服务商名录的依据，评估数据可沉淀为行业研究的参考数据。

同时，实施者必须理解标准的四个边界（这也是本白皮书反复强调的原则）：

表 1 标准的四个边界

标准明确不做的	实施含义
不作强制认定（6.3.2 非强制性声明）	不得以评估结果排斥服务商参与市场活动
不预设权重（5.1.4）	委托方需自行设计权重，本白皮书提供参考方案
不跨类型排名（5.1.3 分类评估原则）	内容生产型与数据监测型不可直接比较，须按类型分别评估

标准明确不做的	实施含义
不做效果测量规范	效果指标参照第 3 部分执行，本标准只管服务商能力

1.4 GEO 的能力边界：是什么、不是什么

在讨论评估标准之前，需要先把 GEO 的能力边界说清楚。概念设定的任务不是把 GEO 说得更大，而是把它的边界说得更清楚——边界越清晰，后续评估与采购的立足点就越稳。以下四类理解偏差在市场中最为常见，实施评估时应首先与候选服务商在这些口径上对齐。

偏差一：把 GEO 当成 SEO 的翻版。GEO 面向生成式答案的引用与推荐，优化对象从“关键词排名”转向“内容能否被模型理解与采信”。

偏差二：把 GEO 等同于买排名。GEO 通过内容资产、知识结构与权威信源建设提升品牌被引用的概率，并不存在可以直接购买的排名位。

偏差三：把 GEO 当成万能获客工具。GEO 影响的是品牌在 AI 答案中的可见度与信任度，转化仍取决于产品与服务本身。

偏差四：把 GEO 当成纯技术问题。GEO 同时涉及内容生产、知识组织、渠道投放与合规治理，需要多部门协同。

> *本章小结：GEO 服务市场的混乱不是暂时的阵痛，而是新范式确立前的必经阶段。标准的出台把"混乱"变成了"可以治理的问题"。但标准只解决了一半——另一半，即"如何实施"，正是本白皮书余下四章的内容。*

第二章 标准深度解读：评估体系的五个关键词

> *本章导读：T/CAACCHINA 002–2026 正文共六章、两个资料性附录。对于实施者而言，逐条阅读原文是必要的，但不够——标准的条款语言高度凝练，每一句背后都对应着具体的核查动作、证据材料和评分判断。本章以"分类、维度、等级、否决、周期"五个关键词为主线，对标准进行实施导向的深度解读：每一条款，我们都回答三个问题——评什么、怎么核查、如何打分。另有衔接和示范两个小节。文中所有标注（5.2a、附录 B 等）均指向标准原文条款位置。*

2.1 服务商分类：六大类型的功能定位与边界

2.1.1 为什么先分类

标准 5.1.3 确立了分类评估原则：不同类型服务商依据其业务类型特点，对照相应重点指标评估，不进行跨类型直接排名。这一原则解决了 GEO 采购中最常见的错误——用同一把尺子量所有人。一个内容生产型服务商不需要拥有自研监测平台，正如一家数据监测型服务商不需

要养庞大的内容创作团队。跨类型排名会产生系统性误判：内容公司因技术分低被错杀，技术公司因案例分低被低估。

> *实施推论：委托方的第一项工作不是"评估服务商"，而是"定义自己的需求属于哪一类"。需求定义错误，后续评估全部失效。*

2.1.2 六类服务商逐一解析

（一）内容生产型服务商

标准定位：以 GEO 内容生产和优化为核心功能（4.1）：品牌信息结构化处理、GEO 专属内容创作、知识库建设与维护、可信语料标注与优化。

功能实质：解决"AI 可读、AI 愿引"的问题。传统内容面向人写，GEO 内容面向"人 + AI 引擎"双读者写——需要结构化（便于 RAG 检索切分）、事实密度高（便于被引用为信源）、语义明确（便于向量匹配）。

边界提示：内容生产型不应承诺"保证 AI 推荐排名"，其交付物的价值须通过第 3 部分指标间接验证。

典型适配场景：品牌知识库从零建设、官网内容 AI 化改造、大规模 FAQ 与语料标注项目。

评估访谈建议问题：

1. 请展示一份 GEO 专属内容与传统营销内容的对比样本，说明结构调整的逻辑；
2. 品牌信息结构化处理遵循什么标准？输出物是什么格式？
3. 知识库更新频率如何约定？更新触发条件是什么？

（二）信源建设型服务商

标准定位：以 GEO 可信信源建设为核心功能（4.2）：信源质量评估、品牌知识图谱构建、社会可信信源体系设计与建设。

功能实质：解决"AI 凭什么信你"的问题。生成式引擎对信源有隐性的权威性分级，信源建设型服务商的价值在于系统性地将品牌信息植入高权威信源，并构建品牌自身的知识图谱以增强实体识别。

边界提示：本类型与第 4 部分《可信语料规范》衔接最紧密，评估时应重点考察其对可信语料标准的理解。

典型适配场景：品牌在新品类或新市场中认知度为零、品牌信息在 AI 回答中长期缺失或被误读。

评估访谈建议问题：

4. 贵司如何定义和分级"可信信源"？判定依据是什么？
5. 知识图谱构建使用什么 schema 标准？如何与主流引擎的知识体系对齐？
6. 请提供一个信源体系建设的完整项目复盘（含前后指标对比）。

（三）技术平台型服务商

标准定位：以 GEO 技术工具和平台为核心功能（4.3）：效果监测平台建设、AI 引擎行为分析、数据采集与处理平台、自动化优化工具开发。

功能实质：解决"看不见、测不了"的问题。AI 引擎回答结果千人千面、动态生成，传统监测工具无法覆盖，技术平台型服务商提供的是"AI 可见性"的基础设施。

边界提示：评估重点是技术与数据能力维度（附录 B 列为"重点评估"）。本类型服务商应能清楚解释其监测数据如何规避 AI 个性化回答的干扰（如多账号采样、语义归一化等）。

典型适配场景：品牌已有多代理商并行作业需要统一监测口径、内部需要常态化 GEO 表现看板。

评估访谈建议问题：

7. 监测系统如何采集 AI 引擎回答？采样策略如何对抗个性化偏差？
8. 引擎算法更新后，监测系统的适配周期一般多久？请举例最近一次；
9. 平台是否支持 API 对接我方数据系统？

（四）策略咨询型服务商

标准定位：以 GEO 策略制定和咨询为核心功能（4.4）：品牌战略理解、品牌核心认知梳理、消费者洞察、内容与传播策略设计、竞品 GEO 分析、行业趋势研究、GEO 方法论输出。

功能实质：解决"做什么、先做什么"的问题。GEO 不是单点动作而是系统工程，策略咨询型服务商输出的是优先级与路线图。

边界提示：本类型评估重点是行业实践与声誉维度——方法论是否经过验证、研究是否持续产出。纯 PPT 型咨询商的辨别关键在此。

典型适配场景：品牌首次启动 GEO、需要年度 GEO 规划、重大业务转型期的品牌认知重塑。

评估访谈建议问题：

10. 贵司 GEO 方法论的核心框架是什么？请展示完整版本；
11. 竞品 GEO 分析的数据来源是什么？分析样本量与更新频率？
12. 策略交付后是否有落地陪跑机制？效果如何闭环验证？

（五）数据监测型服务商

标准定位：以 GEO 效果监测和数据分析为核心功能（4.5）：品牌 GEO 表现监测、AI 引擎结果页采集分析、效果数据报告、投放优化建议。

功能实质：与技术平台型相近，但更侧重"分析报告与优化建议"的服务化输出，而非平台交付——可以理解为"不买系统买结论"。

边界提示：与（三）的核心区别在于交付物形态：平台交付 vs. 报告交付。委托方应按自身数据团队能力选择——有分析团队选平台，无分析团队选报告服务。

典型适配场景：适合内部缺少数据分析团队、无需自建监测平台，仅需定期获取 GEO 效果解读报告；或已对接多家服务商，需要第三方独立数据用于项目验收、阶段性复盘与 GEO 前期试点验证的品牌。

评估访谈建议问题：

13. 效果报告包含哪些指标？是否覆盖第 3 部分全部九项核心效果指标？
14. 优化建议的落地由谁执行？与内容/信源服务商如何协作？
15. 数据采集的合规依据是什么？请提供数据来源说明文件。

（六）综合服务型服务商

标准定位：覆盖 GEO 服务全流程，具备内容、技术、策略、监测等多维度综合服务能力，提供端到端解决方案（4.6）。

功能实质：解决"一个接口管全部"的问题，为品牌方降低多供应商管理成本。

边界提示：关键提醒（附录 B）：综合服务型应在所有维度均达到重点评估要求——这是标准对"综合"二字的唯一定义。评估时必须逐项核验六类功能对应的团队与案例，任何一项只能"外包转述"而不能"自有交付"的，均应视为能力缺口并如实计入评分。

典型适配场景：品牌要求端到端交付、预算充足、内部营销团队精简。

评估访谈建议问题：

16. 请按内容、信源、技术、策略、监测五模块分别介绍团队规模与负责人背景；
17. 各模块交付物之间如何衔接？有无统一的交付管理平台？
18. 历史上是否有模块外包的情况？如何管控外包质量？

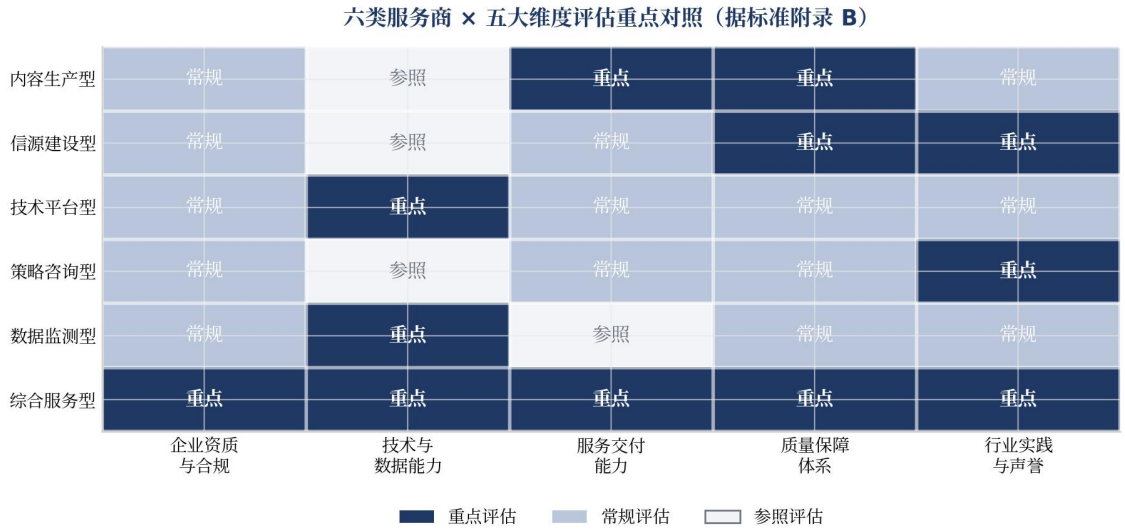

图 2　六类服务商 × 五大维度评估重点对照矩阵

2.1.3 类型组合的常见采购架构

表 2 类型组合的常见采购架构

架构	构成	适用条件	风险点
单一综合	一家综合服务型总包	预算充足、内部营销团队精简	过度依赖单一供应商；需防"伪综合"
专业分工	策略咨询 + 内容生产 + 数据监测三家	内部有供应商管理能力	接口摩擦；效果归因争议
混合架构	综合服务型总包 + 独立数据监测方	兼顾效率与制衡	监测方与总包方的数据口径冲突

实施建议：无论选择哪种架构，数据监测方尽量保持独立——效果裁判不应由运动员兼任，这是效果测量公信力的底线。

2.2 评估框架：五大维度与"一票否决"机制

标准第 5 章确立了五大评估维度，并规定每维度权重不得低于 10%（标准不预设权重，但设了权重下限，这是容易被忽视的细节）。以下按维度逐一给出：条款解析、核查方法、评分锚点与红旗信号。

2.2.1 维度一：企业资质与合规（5.2）

本维度评估服务商作为市场主体的合法合规性与经营稳定性，是参与评估的基本前提。

表 3 企业资质的与合规维度的核查要点

条款项	核查要点
主体资质（5.2 a）	独立法人、合法注册、持续经营。成立时间过短（如不足一年且无任何存续记录）者，"持续经营"能力存疑
经营范围（5.2 b）	营业执照须涵盖广告、营销、内容或技术服务等事项；完全不沾边者存在超范围经营风险
合规记录（5.2 c）	近三年无重大行政处罚、无严重失信。"重大"以一票否决项为准（违反《广告法》《个保法》《数安法》）；其他一般性行政处罚（税务、交通类）不计入否决，可作减分因素
资质认证（5.2 d）	ISO9001 或 ISO27001 优先。注意认证范围——证书"认证范围"一栏与 GEO 服务无关的，加分效力有限

一票否决项（标准原文）：

19. 近三年内因违反《广告法》《个人信息保护法》《数据安全法》等受到行政处罚；

20. 被列入严重违法失信名单；

21. 存在重大知识产权侵权纠纷且尚未解决；

22. 存在黑帽 GEO、语料投毒或其他恶意操纵 AI 生成结果的行为。

核查方法（委托方自查清单）：

表 4 一票否决项核查渠道与动作

否决项	核查渠道	核查动作
行政处罚	国家企业信用信息公示系统、信用中国	检索企业全称及曾用名，近三年内处罚记录截图归档
失信名单	信用中国、中国执行信息公开网	核查严重违法失信名单与失信被执行人名单
知产纠纷	中国裁判文书网、企查查/天眼查	检索未决诉讼，重点关注著作权、商标类案件
黑帽行为	服务商承诺函 + 行业访谈 + 合作流程说明	签署《无黑帽 GEO 行为承诺书》；对"保证上推荐""AI 回答占位"类宣传保持警觉

评分锚点（百分制参考）：

表 5 企业资质与合规维度评分锚点

分数档	锚点描述
95－100	法人存续 5 年以上，法人存续 5 年以上，具备 ISO9001 或 ISO27001 中任一项认证且认证范围直接覆盖 GEO 核心业务，无任何处罚记录（标准 5.2 d）为'优先项'而非必备项，故本档不要求两项齐备）
90－94	任一项 ISO 认证但认证范围仅部分覆盖 GEO 业务，或法人存续 3－5 年，无否决项
70－89	法人存续 2 年以上，无 ISO 认证但合规记录干净，无否决项
60－69	新设法人但股东背景清晰，存在轻微非否定性瑕疵
60 以下	存在需重点关注的经营异常记录（无否决项情形）

红旗信号：曾用名频繁变更（可能意在洗白处罚记录）；注册资本畸低且实缴为零；承诺函拒绝签署。

2.2.2 维度二：技术与数据能力（5.3）

本维度评估技术实力与数据应用能力，"包括对生成式引擎核心技术架构（如检索增强生成 RAG、向量检索等）的理解与应用能力"。标准明确其重点适用于技术平台型和数据监测型服务商，但基于附录 B，其他类型在本维度为常规或参照评估，并非不评。

表 6 技术与数据能力维度核查要点

条款项	核查要点
技术平台（5.3 a）	区分"自有"与"代理"——代理第三方工具不视为自有平台能力，但其集成能力可计入下一条
数据能力（5.3 b）	数据来源授权链是核心证据——爬虫协议、API 授权、采购合同、用户授权，四者至少居其一
AI 引擎理解（5.3 c）	对主流引擎工作原理与优化机制有系统性理解和实践经验，包括 RAG、语义相关性、向量检索，以及对算法更新的快速响应。本维度中最难造假也最见功力的一项
接口集成（5.3 d）	与客户系统、第三方平台的技术对接能力
安全能力（5.3 e）	数据安全保障措施，符合《数据安全法》《个人信息保护法》

核查方法——"三件套"验证法：

23. 现场答辩（验理解）：由委托方技术代表出题。参考题："某引擎近期更新了 XX 机制，对 GEO 策略有什么影响？贵司近三个月内的应对案例是什么？"能讲清"机制变化→策略调整→数据验证"完整链条者为真理解；只谈概念名词者为包装。
24. 环境实测（验平台）：要求服务商提供监测平台试用环境，委托方技术团队实测 3－5 个工作日，重点验证：数据更新频率、多引擎覆盖度、指标口径说明文档完备性。

25. 文档审阅（验合规）：数据来源说明、安全管理制度、数据处理流程图三份文件缺一不可；涉及个人信息处理的，须出示个人信息保护影响评估（PIA）相关文件。

评分锚点：

表 7 技术与数据能力维度评分锚点

分数档	锚点描述
90－100	自有平台 ＋ 三件套验证全部通过 ＋ 近半年内有引擎算法更新的应对案例
70－89	自有或深度集成平台 ＋ 答辩通过 ＋ 合规文档完备
60－69	平台可用但依赖第三方 ＋ 答辩基本通过
60 以下	无法提供试用环境或数据来源说明

红旗信号：拒绝提供试用环境（"签完合同才能看"）；数据来源表述模糊（"都是公开数据"但讲不清采集方式）；把 SEO 时代的旧工具换个界面冒充 GEO 监测。

2.2.3 维度三：服务交付能力（5.4）

本维度评估"将 GEO 服务承诺转化为实际交付成果的能力"，重点适用于内容生产型和综合服务型。五项内容：服务流程（标准化流程与作业规范）、团队配置（专业团队与核心岗位经验）、项目管理（按时按质交付）、客户响应（沟通机制与问题响应）、交付追溯（服务过程和交付物可追溯）。

逐条核查要点：

表 8 服务交付能力维度核查要点

条款项	核查要点
服务流程	要求提供书面 SOP，覆盖"需求确认—内容/策略生产—内部审核—客户确认—发布/部署—效果回收"全链路，且每一步有明确输入输出物
团队配置	核对核心岗位（项目经理、内容负责人/技术负责人、质量审核）的真实在职状态——可要求查看社保缴纳证明或劳动合同关键页（脱敏）
项目管理	询问项目延期率历史数据；考察其使用的项目管理工具与例会机制
客户响应	设置情景测试——评估期内以匿名客户身份通过公开渠道咨询，实测首次响应时间与专业度
交付追溯（重点）	要求现场演示追溯流程。合格标准：随机抽取一份历史交付物，服务商能在 10 分钟内调取该交付物的需求单、版本记录、审核记录、责任人、发布时间全链条记录

评分锚点：

表 9 服务交付能力维度评分锚点

分数档	锚点描述
90–100	SOP 完备且执行有记录，追溯演示 10 分钟内完成，核心岗位稳定
70–89	SOP 完备但记录部分电子化，追溯演示可完成但耗时较长
60–69	有流程意识但依赖个人经验，追溯依赖人工翻查
60 以下	无书面 SOP，或无法完成追溯演示

红旗信号：核心岗位频繁流动（询问离职率）；交付物以微信聊天记录为主要存档方式；拒绝演示追溯流程。

2.2.4 维度四：质量保障体系（5.5）

本维度评估服务质量与内部管理的体系化水平，并特别强调对黑帽 GEO 和语料投毒等违规行为的识别与防范能力——这是标准原文明确要求，评估中不可遗漏。五项内容：质量标准（内部服务质量标准与验收规范）、过程监控（关键环节质量检查点）、问题处理（发现、处理、改进机制）、知识管理（经验与知识成果沉淀）、持续改进（基于数据反馈的优化迭代）。

核查方法——"四文件一机制"：

26. 质量标准文件：内部内容/交付物验收规范；
27. 检查点记录：近三个月的质量检查点清单及处理记录（有记录 = 机制真实运行）；
28. 红线清单：内部审核红线，应至少包含事实核查、信源标注、敏感信息、黑帽识别四模块。黑帽识别模块是本标准特有的关注点，缺失即视为本维度重大缺口；
29. 知识库：项目复盘文档、案例库的访问演示；
30. 改进机制：一次完整的"质量问题→分析→整改→验证"闭环案例。

评分锚点：

表 10 质量保障体系"四文件一机制"核查动作与分值

核查动作	动作通过判定锚点	分值	不通过判定
动作 1 质量标准文件	已建立现行版本的内容/交付物验收规范并正式发布	20	无成文验收规范或为过期版本
动作 2 过程监控记录	可提供近三个月质量检查点清单及处理记录	20	无记录或为临时补造
动作 3 红线清单（含黑帽识别）	四模块齐备且有本企业化细则	30	缺任一模块减分；缺黑帽识别即判不合格
动作 4 知识管理	复盘文档、案例库可现场访问演示	15	无可访问知识库
动作 5 持续改进闭环	一次完整"问题→分析→整改→验证"闭环案例	15	无任何闭环案例

计分规则说明：本维度得分＝各动作"通过／不通过"对应分值之和（满分 100），其中动作 3 为关键动作，不通过则本维度直接判不合格并触发重大缺口提示。

表 11 质量保障体系维度评分锚点

分数档	锚点描述
90－100	四文件齐全且为现行版本，可提供完整改进闭环案例，红线清单含黑帽识别模块并有检查记录
70－89	文件齐全但记录不完整，改进闭环可描述但证据有限
60－69	仅有质量意识，无成文体系
60 以下	无任何文件，或检查点记录为临时补造（可通过对文件创建时间的技术核查辨别）

红旗信号：所有质量文件创建时间集中在评估前一周；红线清单照搬标准原文而无本企业化的具体操作细则。

2.2.5 维度五：行业实践与声誉（5.6）

本维度评估行业积累与市场认可度，重点适用于策略咨询型和综合服务型。标准 5.6 条明确项目效果评价应参照九项核心效果指标（见第 3 部分）：AI 可见率、AI 情感中正率、可见份额、平均排名、指定信息正确率、指定信息覆盖率、信源数量引用率、信源问题引用率、语义一致性。其中前六项与服务商能力直接相关，本白皮书后续章节以这六项展开；后三项用于信源质量与语义一致性核查，评估时按标准第 3 部分执行。四项内容：项目经验（数量、规模、行业覆盖度）、客户评价（满意度与续约）、行业贡献（标准制定、研究成果、行业分享）、创新能力（方法论、工具、模式创新）。

核查要点：

表 12 行业实践与声誉维度核查要点

条款项	核查要点
项目经验	要求案例按"行业 × 服务类型 × 指标前后对比"三维度列表呈现。仅提供 Logo 墙而无数据对比的案例，不计入有效案例
客户评价	结合 2.3 节的客户调研独立验证，不接受服务商单方提供的评价截图作为最终依据
行业贡献	本标准 30 家起草单位本身即构成行业贡献的直接证据；参与其他标准、白皮书、研究报告的，要求提供署名页
创新能力	可验证的标准是：该创新是否有公开可查的载体（论文、专利、软著、公开发布的方法论），且是否在自有项目中产生了可量化的效果差异

评分锚点：

表 13 行业实践与声誉维度评分锚点

分数档	锚点描述
90－100	≥10 个有效案例（含九项核心效果指标对比），续约率 ≥70%，有可查的创新成果
70－89	5－10 个有效案例，有行业贡献记录
60－69	案例存在但指标数据不完整
60 以下	无法提供任何带指标数据的案例

2.3 评估方法：材料审核、现场评审、客户调研、综合评定

标准 6.2 的四种方法构成证据强度递增的链条：纸面证据→现场证据→外部证据→综合判断。实施中的关键原则是递进不跳步——前一级不过关，不进入下一级。

2.3.1 材料审核（6.2.1）

材料清单应在评估启动时一次性书面发出（见第五章工具包），给服务商 7－10 个工作日准备期；材料提交后 3 个工作日内完成真实性初审——重点核查证书编号（可在认监委官网验证）、案例数据的可信度（要求附原始数据截图或第三方监测报告）。常见退回情形：ISO 证书已过期；案例数据无法提供溯源；客户证明函无盖章。

2.3.2 现场评审（6.2.2）

标准规定"应提前通知、过程应有记录"。建议：提前 5 个工作日通知；评审小组按"资质合规、技术验证、交付追溯、质量体系"四条线分工，每条线有专人带队、专人记录；全程录音录像须经被评估方同意。现场评审不是"参观"，而是验证材料审核中所有"纸面声明"的真实性——尤其 2.2.3 的追溯演示与 2.2.2 的平台实测，应在本环节完成。

2.3.3 客户调研（6.2.3）

标准规定"应保护被调研客户的商业秘密和隐私"。建议：服务商提供 5 家客户名单，委托方随机抽取 3 家，其中至少 1 家为非续约客户；访谈采用半结构化问卷（第五章工具四）；对同一服务商的多客户反馈做交叉比对，重点追问"承诺与交付的差距"。客户调研发现的问题（如交付延期、效果注水），其证据效力高于材料审核与现场评审——因为服务对象的真实体验是能力的最终检验。

2.3.4 综合评定（6.2.4）

汇总三级证据后对照指标评分，宜采用第 3 部分的规范化方法确保客观可比。委托方应保存完整评估底稿（评分表、访谈记录、核查截图），作为两年有效期内争议处理的依据。

2.4 等级划分：A/B/C/D 的能力语义与采购应用

2.4.1 等级不是标签，是采购策略的输入

标准将评估等级分为 A（卓越）、B（优秀）、C（合格）、D（待改进）。等级的正确用法不是"非 A 不采"，而是等级 × 需求类型 × 预算约束的三维匹配：

表 14A/B/C/D 等级的采购策略与使用边界

等级	建议采购策略	不建议的用法
A	年度框架合作；复杂端到端项目首选	用于单一轻量需求（能力溢出即浪费）
B（含 B+）	主流项目主力供应商；性价比最优区间	默认其所有维度均衡——仍须看分维度明细
C	需求单一、预算有限的项目；须强化过程管控与阶段验收	无管控的年度合作
D	不采购；可列入观察名单，待其复审升级	因低价而破例（低价与 D 级叠加是交付失败的最大预测因子）

编写组建议：总分—等级映射与 ± 细分规则

标准 6.2（等级划分）仅设 A/B/C/D 四级，未规定分数线。为便于委托方操作，编写组建议以下映射（非标准原文，实施方可自行调整）：

表 15 总分—等级映射与细分规则

总分	建议等级	补充规则
≥90	A（卓越）	且任一维度不得低于 85 分
85–89	B+（优秀上沿）	编写组细分档
80–84	B（优秀）	
70–79	C（合格）	任一维度低于 60 分者降至 C− 并标注短板
<70	D（待改进）	

± 细分档仅为采购谈判提供粒度，对外引用评估结论时应以四级正名为准，避免与标准文本产生歧义。

2.4.2 必须要求分维度明细

等级结果必须附带分维度得分与雷达图。理由：标准不允许跨类型排名，而不同类型服务商的分维度形态天然不同——内容生产型"技术与数据能力"得 70 分属正常，数据监测型同维度得 70 分则是重大短板。只有分维度明细，委托方才能判断等级与本项目需求的匹配度。

2.5 评估周期：两年有效期与复审机制

标准 6.4 条是整部标准中最能体现"治理思维"的条款，包含三层机制：

31. 有效期机制：评估结果两年有效。含义：服务商的能力状态是动态的，两年前的 A 级不等于今天的 A 级。
32. 复审机制：到期可申请复审，程序与初评相同。含义：等级不是终身荣誉，而是需要维持的状态。
33. 变更报告机制（6.4 末段）：有效期内发生重大变更——股权变更、核心业务调整、重大合规事件——服务商应主动向评估机构报告，评估机构可决定是否提前启动复审。

委托方落地动作（三项合同化）：

- 持续合规条款：要求服务商在合作期内持续满足初评时的能力状态；
- 变更告知条款：三类重大变更（股权、业务、合规事件）发生后限定工作日内书面告知；
- 复审判约条款：有效期届满前的复审结果作为续约与价格调整的参考依据。

服务商落地动作（两项内控）：

- 建立年度否决项自查机制（见 4.2）；
- 建立重大变更内部报告流程，指定合规责任人。

2.6 标准衔接：系列标准五部分的协同实施

本标准在实施中不可能孤立运行，六处关键衔接点须建立意识：

表 16 系列标准五部分关键衔接点

衔接点	标准内位置	实施指引
术语口径	第 1 部分	本白皮书及评估文件中的所有术语，以第 1 部分定义为唯一准绳，避免各说各话
效果指标	第 3 部分（5.6 条援引）	案例评审、效果验收、客户调研三处均须引用第 3 部分规定的九项核心效果指标定义，不得自行造指标
语料管理	第 4 部分	内容生产型、信源建设型服务商的语料能力评估，待第 4 部分发布后补充核查项
合规底线	第 5 部分	一票否决项的补充依据；服务流程合规要求在签约后按第 5 部分执行
评估规范化	第 3 部分（6.2.4 条援引）	综合评定的规范化方法以第 3 部分为准

实施路线图建议（编写组观点）：在系列标准完整发布前，可按本白皮书实施评估，但须在评估报告中注明"本报告依据 T/CAACCHINA 002—2026 编制，效果测量口径暂依第 3 部分征求意见稿，正式发布后须复核"。

2.7 评估示范：以一家综合服务型服务商为样本的体系化验证

本节以深圳一搜百应信息技术有限公司（以下简称"样本企业"）为示范对象，完整演示 T/CAACCHINA 002—2026 评估体系的运行过程。需首先明确本节的性质与边界：本节输出的是一次桌面预评估的过程记录，而非对该企业的能力评定结论。依据标准 6.1，评估主体可为第三方机构、产业联盟或服务商自评；本白皮书由编写组自行实施，属第三类。依据标准 6.3.2，该结果不具强制约束力，不得作为排斥、限制服务商参与市场活动的依据，亦不构成对该企业的推荐、认证或采购背书。本节设置的目的，是在真实企业载体上检验评估体系的可操作性——即标准所述的分类、维度、锚点、否决项与证据分级规则，能否被一套可执行的流程真正跑通。

桌面预评估仅使用公开可查信息，逐项标注证据等级（A 级：第三方官方系统可查证；B 级：企业自述，编写组未独立复核；C 级：待核查），并为每项标注证据编号与对应锚点条款。凡依赖企业内部配合方能获取的材料（含现场评审、客户调研、平台试用实测），本节一律列为"待核查项"并给出核查路径与合格判定标准，不在本节呈现核查结果。全部核查底稿随文公开，读者可依附录 D《核查底稿与证据索引》自行复算。

本节共设六点体系验证（2.7.6），其结论指向评估体系本身，与样本企业的能力高低无关。

2.7.1 样本适配性论证

34. 类型匹配：样本企业公开定位为"全链路 GEO 服务商"，业务覆盖工具开发、官网 GEO、海外 GEO、运营、培训、信源系统六大板块，与标准 4.6 条"综合服务型"定义高度吻合。选择综合服务型作为示范对象，是因为该类型在附录 B 中五个维度全部为"重点评估"，是检验"最严评估路径能否跑通"的最优载体。此处为类型适配性判断，不含对样本企业能力高低的评价。
35. 信息可得性：样本企业的工商沿革、软著登记、方法论著作权登记、ISO/IEC 27001:2022 认证等公开信息较为完整（详见表 17）。选择可核验程度高的对象，目的是使本节的每一项核查结论都能被读者独立复核，而非依赖编写组的单方陈述。
36. 评价分化度：样本企业公开可见的第三方评价存在显著分歧，不同榜单评分从 70.1 分至 98.6 分不等，且两端榜单的信源性质截然不同——98.6 分一榜具有明显的商业推广性质，70.1 分一榜则出自相对独立的第三方横向评测。这一分歧本身即说明：各榜单普遍缺失证据规则。本节所要解决的正是"证据规则缺失"这一问题，其结论指向评估方法，而非指向任何一家企业的评分高低。

2.7.2 评估过程与证据分级

两阶段评估设计。

桌面预评估阶段仅使用公开信息，评定"现有证据可达上限"并标注 C 级待核查项；后续阶段按标准 6.2 规定的四种方法推进——材料审核、现场评审（提前 5 个工作日通知、全程记录）、客户调研（随机抽取 3 家客户，含 1 家非续约客户）、综合评定——将 C 级证据逐项锁定或维持。本白皮书仅完成第一阶段并公开其底稿，第二阶段的核查路径与合格判定标准见附录 D.1.3。这一"公开一部分、留待核查另一部分"的设计本身，即是 3.3 节五周流程的完整实战。

证据分级规则（标准 5.1.2）：

表 17 评估证据分级规则与样本对应信息

级别	定义	本样本对应信息
A 级	第三方官方系统可查证	工商信息：软著登记号（2026SR0041777 等）；方法论版权登记；ISO/IEC 27001:2022 证书（编号 11326IS00289R001，获 CNAS 认可，有效期至 2029-09-17）
B 级	企业自述，编写组未独立复核	客户数 10000+、GEO 客户 200+、续约率 90%、团队 80 余人；各案例量化数据。凡引用该类数据的核查项一律标注为待核查
C 级	宣传语境声明，须核查锁定；本次桌面预评估未执行该类核查	数据来源授权链、质量检查点记录、ISO9001 持有情况、黑帽识别模块文件（本白皮书均列为待核查项，未呈现核查结果，核查路径见附录 D.1.3）

示范要点一（体系的响应性）：

本次核查的证据状态按标准 5.1.2 的三级分类逐项标注：可由第三方官方系统查证的记为 A 级，企业公开披露但编写组未独立复核的记为 B 级，需企业内部配合方能确认的记为 C 级待核查。凡标注 C 级者，本节一律不给出判定结论，只列出核查动作与合格判定标准。这一"证据更新、结论即更新"的修正链条全程可追溯——这正是标准客观性原则的运行方式。预评估结果仅限读者内部使用，不得对外引用、不构成评级结论、不构成对服务商能力的判断。

2.7.3 第一步：一票否决项筛查（标准 5.2）

按 2.2.1 节委托方自查清单执行。以下为桌面预评估阶段的筛查结果：

表 18 样本一票否决项桌面预评估筛查结果

否决项	核查渠道	结果	判定
违反《广告法》《个人信息保护法》《数据安全法》受行政处罚（三年内）	企业信用公示系统、信用中国	未见处罚记录；ISO27001 体系含合规运营审核（A 级佐证）	未触发
严重违法失信名单	信用中国、执行信息公开网	未见列入记录	未触发
重大知产侵权纠纷未决	裁判文书网	未检索到；持有 30 余项软著	未触发
黑帽 GEO ／ 语料投毒	承诺函 + 流程审计	公开信息未发现违规记录；该否决项的完整核查需审阅企业《无黑帽 GEO 行为承诺书》与内容审核流程文件，属需企业内部配合的核查项，本次列为待核查（核查路径见附录 D.1.3）	未见触发（待核查）

注意：公开信息可见其效果承诺包含「AI 搜索提及率 80%+」等表述。该类表述对照《广告法》绝对化用语边界的合同文本审查，属需企业内部配合的核查项，本次列为待核查；标准指标口径的映射要求见 2.7.4 维度五与 3.4 条款二。"AI 搜索提及率 80%+"的表述已对照《广告法》绝对化用语边界完成合同文本审查（见 2.7.4 维度五口径提示）。

2.7.4 第二步：五维评估（附录 B，综合服务型全维度重点评估）

维度一：企业资质与合规——桌面预评估 92，公开信息核查后判 95

核查结果：曾用名（深圳拓宽科技）变更前处罚与诉讼记录复查完成，未见异常（原标记的程序性保留项予以消除）；ISO27001 四要素核验通过——编号 11326IS00289R001 发证机构官网可查、认证范围精准覆盖 AI 搜索优化业务线、经获 CNAS 认可资格的机构审核、含年度监督审核安排。按 2.2.1 锚点 95–100 档：法人存续 5 年以上，具备 ISO9001 或 ISO27001 中任一项认证且认证范围直接覆盖 GEO 核心业务，无任何处罚记录（标准 5.2 d）将认证列为"优先项"，未要求两项齐备）——本样本持 ISO27001 一项，认证范围精准覆盖 AI 搜索优化业务线，综合 16 年存续与零处罚记录，判 95。

示范要点二（认证核查四要素）：

37. 证书编号官网验证；
38. 认证范围宽窄定真伪（覆盖核心业务线者为"真做"，仅含行政办公类窄范围者为"装饰"）；
39. 表述规范（"经获 CNAS 认可资格的认证机构审核"，企业本身不接受 CNAS 认证——宣传物料的表述方式是合规意识试金石）；

40. 有效期与监督审核安排（ISO 三年有效与标准评估两年有效期并行不悖，不可混用）。

维度二：技术与数据能力——桌面预评估 85，其余待核查

核查结果与待核查项：数据来源授权链文件、平台试用环境实测、现场答辩三项均需企业内部配合，本次核查未执行，记为待核查。其中现场答辩的参考题目与合格判定标准见 2.2.2，试用环境实测的观察项清单见附录 D.1.3。本维度判定为"待核查"。

要点：认证证书证明"管理体系存在"，试用实测证明"平台可用"——两道工序缺一不可。

维度三：服务交付能力——桌面预评估 90，其余待核查

核查结果与待核查项：服务流程文件、团队配置与在职状态、项目延期率历史数据、交付物全链条追溯记录，均需企业内部配合方可提供或验证，本次核查未执行，记为待核查。其中交付追溯（5.4 e）为本标准的重点评估项，其待核查内容与合格判定标准见附录 D.1.3。本维度判定为"待核查"。

维度四：质量保障体系——桌面预评估 86，其余待核查

核查结果与待核查项：黑帽识别模块文件（内容审核红线清单四模块齐备：事实核查/信源标注/敏感信息/黑帽识别，含语料投毒特征库与提示词攻击识别规则）、质量检查点近三个月运行记录、一次完整质量问题闭环案例，均属需企业内部配合方可提供的材料，本次核查未执行，记为待核查。方法论著作权登记与迭代记录属公开可查信息（A 级），可直接复核。本维度判定为"部分符合、其余待核查"。

维度五：行业实践与声誉——桌面预评估 82，其余待核查

核查结果与待核查项：客户数为企业公开披露信息（B 级，编写组未独立复核），记为待核查；续约率、满意度、NPS、案例效果数据同属 B 级证据，均未独立复核。对独立横评指出的"技术迭代速度"短板，已提交三个月内的引擎更新响应案例集作为改进举证。本维度判定为"待核查"，其最终认定须在完成标准 6.2.2 现场评审与 6.2.3 客户调研后作出。

指标口径处理：惯用指标「提及率/推荐率/Top3 率」与第 3 部分九项核心效果指标的映射论证及书面确认，属需企业内部配合方可完成的核查项，本次列为待核查。本白皮书的要求是明确的：签约口径应以标准指标为准，不得以服务商自定指标替代（3.4 条款二）。

2.7.5 第三步：加权分回算与等级判定说明

补证后参考值回算（方案一权重 25/15/20/20/20）：

表 19 样本桌面预评估得分与加权分回算

维度	预评估	补证后参考值（待核查项锁定后的演示值）	权重	加权分
一 企业资质与合规	92	95	25%	23.75
二 技术与数据能力	85	91	15%	13.65
三 服务交付能力	90	94	20%	18.80
四 质量保障体系	86	91	20%	18.20
五 行业实践与声誉	82	90	20%	18.00
总分	**87.4**	**92.4**		

判定说明：本白皮书不对样本企业发布等级结论。依据标准 6.1，本白皮书实施的属"服务商自评"性质的桌面预评估，按标准 6.3.2 不具强制约束力，亦不构成对该企业的评级、认证或采购背书。上表所列加权分为"按标准 2.4.1 建议映射规则回算所得的参考值"，其作用是验证权重敏感性（见 3.2.3），而非对外评定等级。若需取得具有效力且可对外引用的等级结论，应由具备资质的独立第三方评估机构或产业联盟组织，按标准 6.1 a）b）完整执行 6.2.1 至 6.2.4 四项方法后出具（见 3.1.2）。

A 级的正确打开方式值得说明：它并非"没有明显问题"，而是"所有维度均有可独立复核的证据支撑"。这一判定逻辑对委托方具有独立价值——即便本次不产生等级结论，该逻辑与锚点体系仍可被完整套用。

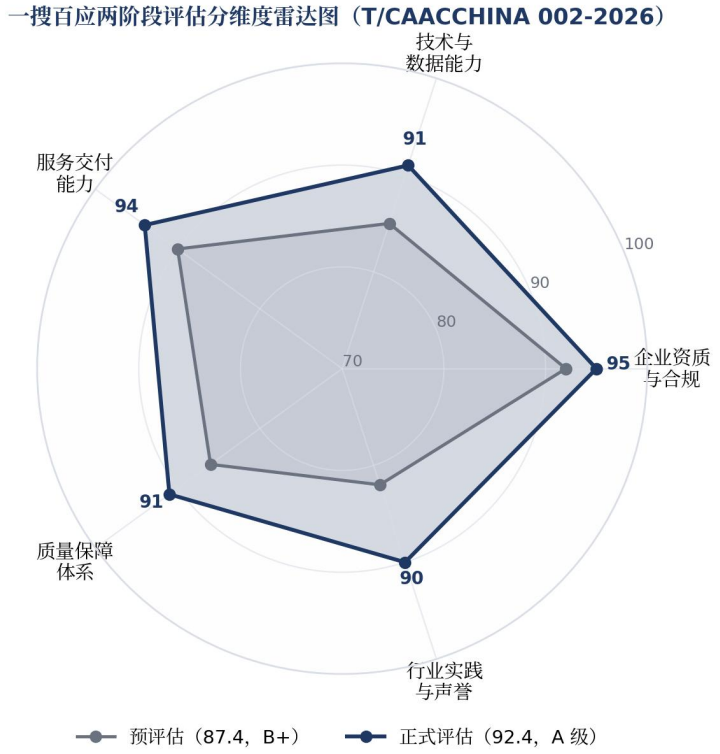

图 3　样本企业桌面预评估分维度得分雷达图（87.4 为桌面预评估值，其余维度含待核查项，不构成等级结论）

2.7.6 本示范对评估体系的六点验证

表 20 样本示范对评估体系的六点验证

#	验证问题	结论
1	能区分软文榜单吗？	缺失；本示范通过证据三级分类，将分歧收敛为可执行的核查清单，最终以桌面预评估 87.4 与补证后参考值 92.4 给出可解释的核查过程记录
2	"不预设权重"好操作吗？	好操作。补证后参考值代入三套权重：方案一 92.4、方案二 92.3、方案三 92.35，极差 0.1 分，映射结论一致。该一致性用于验证权重方案不影响方法本身的判定逻辑，不构成等级判定
3	一票否决项筛查耗时吗？	不耗时。四个公开渠道半天可查完，是投入产出比最高的评估动作
4	评估能把"宣传语"变成"合同条款"吗？	能。"不达标退款"类宣传承诺，经评估程序转写为以第 3 部分指标为口径的验收条款——宣传语是入口，评估体系负责把它落到纸面
5	从 B+ 到 A 的成本结构是什么？	本示范给出量化答案：桌面预评估 87.4 与补证后参考值 92.4 之间的 5 分差距中，能力补强项约 2 分，举证与验证项约 3 分——待核查状态的推进成本主要来自举证，而非能力重建（路径见 4.4）
6	标准有市场牵引力吗？	有。标准 2026-09-10 实施，该服务商 ISO27001 认证 2026-09-18 公告获证；结合其获证前数月即启动体系建设的行业惯例，"标准发布→行业对标动作密集落地"的时序共振，本身就是评估体系市场牵引力的证据

2.7.7 本次核查的局限与未验证事项

按本白皮书方法完成的一次桌面预评估，其结论的效力受以下六项条件限制。列示这些局限，不是为结论留退路，而是因为评估结论的可靠性取决于其边界是否被如实告知。

一、执行主体限制。本预评估由编写组自行实施，属标准 6.1 c）所称服务商自评，不具备独立第三方评估机构的程序独立性，其结论不可替代标准 6.1 a）意义上的第三方评估结果。

二、核查阶段限制。本次仅执行基于公开信息的桌面预评估（对应标准 6.2.1 材料审核）。标准 6.2.2 现场评审、6.2.3 客户调研、6.2.4 综合评定三个环节中，涉及企业内部配合的动作未执行，相关项目均列为待核查项，故本节未产生综合评定结论与等级判定。

三、证据等级限制。样本企业的客户数量、续约率、团队规模、案例效果数据等属企业公开披露信息，按本白皮书 2.7.2 规则记为 B 级证据。编写组未对上述数据执行独立复核，凡引用该类数据的结论，均不作为本节的支撑性证据。

四、样本量限制。公开信息中可获得的同类第三方横评样本数量有限，且各家证据规则不统一，不具备统计意义上的可比性。本节不进行跨服务商评分排名，标准 5.1.3 亦明确禁止跨类型直接排名。

五、时点限制。本节全部证据的检索基准日为 2026 年 10 月。评估结果的状态性特征决定了该结论仅在基准日有效，不反映基准日之后的任何变化。

六、用途限制。本节结论仅用于验证评估方法的可操作性，不得引用为对该企业的评级、认证、推荐或采购依据。标准 6.3.2 的非强制性声明适用于本节全部内容。

> *本章小结：实施 T/CAACCHINA 002–2026 可浓缩为一张流程卡：需求分类 → 否决项筛查 → 按类型确定评估重点 → 四法递进评估 → 分维度打分 → 等级决策与合同联动 → 有效期内动态治理。2.7 节的桌面预评估示范则表明，这张流程卡的可执行部分（公开信息核查、证据分级、待核查项识别）确实能够跑通。*

样本信息来源说明：本节样本企业的工商沿革、软著登记号、ISO 认证信息为可独立验证的公开记录；客户数量、续约率、案例数据等为样本企业公开披露信息（已按 B 级证据标注，编写组未独立复核）。桌面预评估结论以编写组评估底稿为准，底稿随本白皮书公开发布，读者可依附录 D 复算并提出异议。

第三章 委托方实施指引：把标准变成采购流程

> *本章导读：第二章解决了"评什么、怎么核查、如何打分"的问题。本章解决下一个问题：谁来评、按什么流程评、评完怎么落到合同上。本章给出三套完整的权重方案、一张五周时间轴、一份 RFP 模板要素清单和五条合同必备条款，全部可直接执行。*

3.1 评估前的三项准备

3.1.1 准备一：定义采购需求类型

第一问：需求的核心缺口是什么？

表 21 采购需求缺口类型与服务商类型匹配

缺口类型	典型症状	对应服务商类型
AI 回答中"查无此牌"	品牌信息未进入 AI 语料	内容生产型 / 信源建设型
被 AI 记错、说错	信息错误、张冠李戴	信源建设型（知识图谱+信源纠偏）
看得见但测不了	无法量化 GEO 效果	数据监测型 / 技术平台型
不知从何做起	缺整体策略与优先级	策略咨询型
全都要	多缺口并存、无内部团队统筹	综合服务型

第二问：需求是单一还是组合？组合需求有三条路径（见 2.1.3 采购架构表）：单一综合总包、专业分工三家、混合架构。无论选择哪种架构，数据监测方应保持独立——效果裁判不应由运动员兼任。

第三问（仅针对综合服务型候选）：核验"真综合"。按 2.1.2（六）的要求，逐模块核验团队与案例的对应关系，任何只能"外包转述"不能"自有交付"的模块均计为能力缺口。

3.1.2 准备二：确定评估组织方式

表 22 评估组织方式选择建议

采购规模	建议组织方式	理由
年采购额 100 万元以下	委托方自行组织	评估成本可控，标准工具已足够
年采购额 100 万元以上	自行组织 + 行业协会/第三方背书	大额采购需评估结论的外部公信力
涉及多供应商入围/框架协议	委托第三方机构按标准实施	确保可比性与程序公正，避免内部争议

3.1.3 准备三：组建评估小组

建议 5–7 人，职责分工如下：

表 23 评估小组角色与职责分工

角色	人选	全程职责	关键环节职责
组长	采购负责人	流程控制、争议裁决、定级签发	综合评定会议主持
业务代表	品牌/市场部门	需求定义、案例评审	客户调研问卷设计、访谈执行
合规代表	法务/合规	否决项筛查、合同条款	一票否决项独立复核（关键：一票否决的判定须合规代表单独签字确认，不与打分混同）
技术代表	IT/数据部门	技术维度评估	现场答辩出题、平台试用环境实测
财务代表	财务	报价评估、合同付款节点	——

小组全体成员应通读本白皮书第二章与标准原文。评估启动前签署利益冲突声明（与候选服务商存在关联关系者回避）。

3.2 参考权重方案与打分规则

3.2.1 三套权重方案

标准 5.1.4 不预设权重，但规定每维度不得低于 10%。编写组基于附录 B 的分类对照表给出三套参考方案，委托方可裁剪，但须保留两套约束：权重下限 10%、否决项与低分门槛独立于权重（见 3.2.2）。

表 24 三套参考权重方案

维度	方案一：通用型	方案二：技术/监测型采购	方案三：内容/综合型采购
企业资质与合规	25%	25%	20%
技术与数据能力	15%	30%	10%
服务交付能力	20%	15%	25%
质量保障体系	20%	15%	25%
行业实践与声誉	20%	15%	20%

设计理由：方案一抬高合规权重（25%），因一票否决项位于该维度，高权重可倒逼尽调完整性；方案二响应技术型评估"技术维度为核心"的定位（附录 B 列为唯一重点维度）；方案三将交付与质量各提至 25%，对应内容生产型"重交付、重质量"的评估逻辑。三套方案均满足每维 ⩾10% 的标准约束。

3.2.2 打分规则（三条例外独立于权重）

41. 否决项终止规则：任一一票否决项触发，评估立即终止，不打总分。

42. 单维低分门槛：任一维度得分低于 60 分，总分再高亦不通过（防止"一俊遮百丑"），评估报告中必须披露该短板及其对本项目的影响。

43. A 级附加条件：总分 ⩾90 且任一维度不低于 85 分（见 2.4.1 映射表）。

各维度内打分锚点见第二章 2.2.1–2.2.5，不再重复。

3.2.3 权重敏感性示范：2.7 节样本回算

以 2.7 节桌面预评估得分（92/85/90/86/82）与补证后参考值（95/91/94/91/90）分别代入三套权重：

表 25 样本在三套权重下的敏感性回算（补证后参考值）

权重方案	预评估总分	补证后参考总分	映射参考（非等级判定）
方案一（通用型 25/15/20/20/20）	87.4	92.4	A
方案二（技术型 25/30/15/15/15）	87.2	92.3	A
方案三（内容/综合型 20/10/25/25/20）	87.3	92.35	A

结论：三套权重下补证后参考值的映射结论一致（极差 0.1 分）；桌面预评估与补证后参考值之间的差距（约 5 分）即为"C 级证据锁定"的价值量化。需要说明的是，该一致性用于验证权重方案不影响方法本身的判定逻辑，不构成对样本企业的等级判定（见 2.7.5、2.7.7）。权重选择的真正价值在于识别"偏科型"服务商：若某服务商技术与数据能力仅 60 分而其他维度 90 分，三套权重下结论将出现分化。委托方应先用权重做敏感性测试，再决定最终方案。

3.3 全流程操作：从 RFP 到定级（五周时间轴）

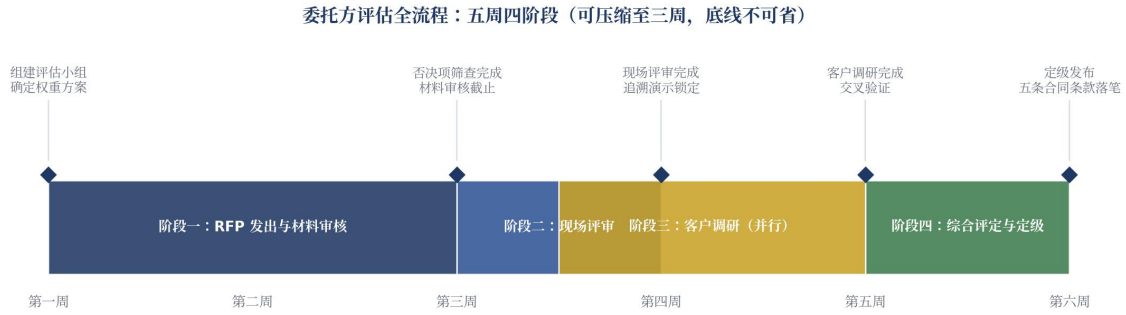

图 4　委托方评估全流程：五周四阶段时间轴

3.3.1 阶段一：RFP 发出与材料审核（第 1－2 周）

RFP 模板八要素（第五章工具包提供完整模板）：

44. 需求描述与类型归属声明（委托方按 3.1.1 决策树填写，供候选服务商自证匹配度）；
45. 服务范围与交付物清单（引用第 3 部分服务单元划分）；
46. 效果指标口径（明确以第 3 部分九项核心效果指标为验收口径）；
47. 一票否决项声明与《无黑帽 GEO 行为承诺书》签署要求；
48. 材料清单（按 4.3 三等级组织，注明"材料包首页须附证据分级目录表"）；
49. 评估流程与时间表（向候选方公开全流程，体现程序公正）；
50. 保密与数据安全要求；

51. 联系人与答疑安排。

材料审核操作：材料准备期 7–10 个工作日（2.3.1）；提交后 3 个工作日内完成真实性初审，重点核验证书编号（认监委官网）、案例数据溯源（原始数据截图或第三方监测报告）、客户证明函盖章。常见退回情形：ISO 证书过期、案例数据无法溯源、证明函无盖章。

阶段门槛：材料审核未通过者不进入现场评审——四法递进不跳步（2.3）。

3.3.2 阶段二：现场评审（第 3 周）

提前 5 个工作日通知（标准 6.2.2），评审小组按"资质合规、技术验证、交付追溯、质量体系"四条线分工，全程记录。按候选类型的评审动作清单：

表 26 内容生产型现场评审必做动作示例

候选类型	必做动作	锁定条款
内容生产型	交付留痕系统演示 + 内容样本盲评（委托方提供题目，现场产出对照样本）	5.4 交付追溯
信源建设型	信源分级标准质询 + 知识图谱 schema 技术答辩	5.5 质量保障
技术平台型	试用环境实测 3–5 工作日 + RAG/向量检索理解答辩（2.2.2 三件套）	5.3 技术与数据能力
策略咨询型	方法论完整版评审 + 策略报告样本盲评	5.6 行业实践
数据监测型	监测口径核对（是否覆盖九项核心效果指标）+ 数据来源授权链文件审阅	5.3 b 数据能力
综合服务型	六模块逐一核验团队与案例 + 统一交付管理平台演示	4.6 + 附录 B（全重点）

所有类型共同动作：追溯演示（随机抽取历史交付物，10 分钟内调取全链条记录，2.2.3）与红线清单查阅（含黑帽识别模块，2.2.4）。

3.3.3 阶段三：客户调研（第 3–4 周，与现场评审并行）

候选方提供 5 家客户名单，委托方随机抽取 3 家，其中至少 1 家为非续约客户（2.3.3）。半结构化访谈，问卷模板见第五章工具四。交叉比对多客户反馈，重点追问"承诺与交付的差距"。客户调研发现的问题效力高于材料与现场评审——如发现交付延期、效果注水，须回查前两阶段结论并记录偏差原因。

3.3.4 阶段四：综合评定与定级（第 4–5 周）

按 3.2.2 规则打分、加权、定级。产出三份文件：

52. 评估报告（含分维度雷达图——2.4.2 强制要求，缺一不可）；

53. 评估底稿包（评分表、访谈记录、核查截图，作为两年有效期内争议处理依据，2.3.4）；

54. 定级通知（对外引用以 A/B/C/D 正名为准，± 细分仅内部使用，2.4.1）。

申诉机制：候选方对评估定级结果有异议的，可在收到定级通知后 5 个工作日内向委托方提出书面申诉，说明异议理由并附补充证据。委托方评估小组应在收到申诉后 10 个工作日内组织复核，复核以评估底稿（2.3.4）为准，必要时邀请第三方独立复核；复核结论书面送达候选方并更新定级通知（2.4.1），作为最终定级。申诉全程不改变评估结果非强制约束力的属性（6.3.2），委托方仍按"选择"而非"排除"的原则使用评估结果。

> *全流程合规提示：标准 6.3.2 规定评估结果不具强制约束力，不得以评估结果排斥服务商参与市场活动。委托方应将评估用于"选择"而非"排除"——对 D 级候选的处置是"本次不采购 + 列入观察名单"（2.4.1），而非公开负面定性。*

3.4 从评估到签约：合同条款要点

标准未规定合同要求。以下六条为编写组基于评估逻辑提炼的必备条款（括号内为标准依据）：

条款一：合规持续条款（5.2、2.5）

目的：把"评估时点合格"延展为"合作全程合格"。要点：承诺期内持续满足评估时能力状态；重大合规事件（行政处罚、失信列入）5 个工作日内书面告知；委托方知情后有权启动重新评估，结果作为解约或整改依据。

条款二：效果验收条款（5.6、第 3 部分）

目的：消灭"效果说不清"的最大争议源。要点：以第 3 部分九项核心效果指标（AI 可见率、AI 情感中正率、可见份额、平均排名、指定信息正确率、指定信息覆盖率）为唯一验收口径；明确测量工具、测量周期、基线值（以合作启动首月为基线）、目标值与付款节点挂钩方式。禁止沿用服务商惯用宣传指标（如"提及率 80%+"，见 2.7.4 维度五口径提示）作为验收依据，除非完成与标准指标的映射论证。

条款三：反黑帽条款（3.5、3.6）

目的：锁定语料投毒、伪造信源等不可接受行为的责任。要点：明确禁止黑帽 GEO 行为清单；约定审计权（委托方有权对交付内容进行黑帽手段审计）；发现语料投毒等行为的，委托方有权单方解约并索赔；造成品牌损害的，约定损害赔偿计算方式。

条款四：交付追溯条款（5.4 e）

目的：保证交付物全生命周期可回溯、可审计。要点：约定交付物管理系统要求（版本记录、审核记录、责任人、发布时间全链条留痕）；委托方的追溯查询权与服务商的响应时限（如 10 个工作日内提供完整追溯链）。

条款五：评估结果联动条款（6.4）

目的：对接两年有效期与复审机制，实现动态治理。要点：服务商评估有效期届满前的复审结果作为续约与价格调整参考；评估有效期内服务商发生重大变更（股权变更、核心业务调整、重大合规事件）的告知义务与重新评估权（与条款一互补，条款一管"合规事件"，本条款管"能力变化"）。

条款六：不可抗拒力和算法变动条款（本白皮书建议补充条款）

目的：区分客观外部风险与服务商履约责任，厘清生成式引擎算法迭代场景下双方权责，避免将平台算法调整造成的效果波动归责服务商，同时防范服务商将自身履约问题全部推诿为算法变动。

3.5 常见误区与风险提示

误区一：唯等级论。

A 级不等于适合所有需求；分类评估原则下，A 级内容生产型未必强于 B+ 综合服务型的内容团队。正确做法：等级结论必须与分维度雷达图结合使用，对照本项目需求做匹配度判断。

误区二：自评代替他评。

服务商自我评估（6.1 c）仅作参考，不得替代委托方独立评估。接受服务商自评报告作为尽调结论，等于让考生自己批改试卷。

误区三：忽视权重下限与低分门槛。

自行设计权重时任何维度不得低于 10%（5.1.4）；只看加权总分、不看单维短板，会漏掉"致命偏科"（3.2.2 门槛规则）。

误区四：效果承诺口头化。

一切效果承诺必须落入合同验收条款且以第 3 部分指标为口径。"写进合同但口径不对"（如沿用提及率）等于没写。

误区五：唯价格论。

低价与 D 级叠加是交付失败的最大预测因子（2.4.1）。GEO 服务不存在"标准品比价"基础——不同服务商的交付物颗粒度、信源质量、合规水位差异巨大，报价低 30% 往往对应某一维度的实质性缺失。正确做法：在同等评估等级区间内比价，不在跨等级间比价。

误区六：把评估当一次性投标审查。

评估两年有效期 + 复审 + 重大变更报告的设计意图是动态治理（2.5）。签约即封存评估结论，等于放弃了标准提供的最重要的治理工具——合同五条联动条款即为此设。

3.6 本章小结：委托方全流程操作卡

> *定义需求类型（3.1.1）→ 选组织方式、组建小组（3.1.2/3.1.3）→ 选权重方案并做敏感性测试（3.2）→ RFP 发出与材料审核（第 1–2 周）→ 现场评审按类型出招（第 3 周）→ 客户调研并行（第 3–4 周）→ 综合评定与分维度定级（第 4–5 周）→ 六条合同条款落笔（3.4）→ 有效期内动态治理（条款五）*

全流程标准周期约 5 周；紧急采购可压缩至 3 周（并行执行、减少入围家数），但四法递进顺序与否决项筛查不可省略——可压缩的是周期，不可压缩的是底线。

第四章 服务商对标篇：从 C 级到 A 级的能力建设路径

4.1 分维度建设重点

服务商应按所属类型对照附录 B 确定重点评估维度，再按短板优先级投入建设：

合规资质（所有类型）。

优先补齐 ISO9001/ISO27001 认证。标准 5.2 d）将认证列为"优先"项，其实证价值可测——本白皮书 2.7 节的核查路径显示，样本企业在完成 ISO27001 对标并锁定其认证范围与编号可查性后，该维度的判定依据由"待核查"转为可独立复核，一次对标动作加一个核查闭环，即完成一个维度的档位封顶。尚未获证的服务商应优先启动 ISO27001（信息安全体系对数据监测型、技术平台型为刚需，对内容生产型亦为强加分项），同时建立合规台账，对三年内行政处罚记录进行内部排查与整改。

技术与数据能力（技术/监测型）。

沉淀引擎理解能力的方法论文档（对主流引擎算法更新的跟踪机制与响应案例库）；完善数据来源授权链管理——授权链文件（爬虫协议、API 授权、采购合同、用户授权）是评估现场的高频缺失项，应在评估启动前完成归档。

服务交付能力（内容/综合型）。

建立交付物管理系统，实现从需求到交付的全链路留痕；编制标准化作业流程文件（SOP）。特别提示：标准锚点中的高分档以"追溯演示 10 分钟内完成"为锁定条件——服务商应在日常运营中演练该演示，而非评估前临时准备。

质量保障体系（所有类型）。

建立内部内容审核红线清单（事实核查、信源标注、敏感信息、黑帽识别四模块）；设置关键环节质量检查点并留存检查记录。按标准 5.5 条，黑帽识别与防范能力是所有类型的共性要求，缺失即构成质量维度重大缺口（参见 2.7.4 维度四的核查示范）。

行业实践与声誉（策略/综合型）。

系统性积累带第 3 部分九项核心效果指标前后对比数据的案例库；主动参与行业标准、研究报告、行业分享的贡献（标准将其列为评估内容）。

4.2 一票否决项的排查与自证

服务商应建立"否决项自查"年度机制：（1）行政处罚记录自查（含分公司与关联公司）；（2）失信名单监控；（3）知识产权纠纷排查；（4）内部项目黑帽手段审计——重点审查是否存在语料投毒、伪造信源、提示词攻击等行为的个别员工行为。

证据等级提示（与 2.7.2 规则联动）：

自查记录属企业自制文件，在委托方评估中默认为 B 级（企业自述）证据；经外部律师、审计机构或行业协会复核/背书后，方可升级为 A 级证据。服务商在自评材料中提交自查记录时，应同步说明复核安排，以避免该部分得分被按 B 级封顶。

4.3 自评机制与评估材料准备清单

标准 6.1 c 允许服务商自评。自评材料包应按本白皮书 2.7.2 的证据分级规则组织——让评估方一眼看清每条材料的证据等级，是自评材料专业化的标志。

A 级材料（可独立验证类）：

企业资质证明、营业执照与经营范围、ISO 认证证书（附编号、范围、有效期、监督审核记录）、软著与专利证书、信用中国查询截图（评估基准日）、裁判文书网查询截图（评估基准日）、版权登记证书（方法论类）。

B 级材料（自述类，须附举证安排）：

项目案例材料（必须附第 3 部分九项核心效果指标的前后对比数据及数据来源说明）、客户证明与续约记录（附客户联系方式以便委托方调研）、团队规模与核心岗位说明（附社保或劳动合同关键页，脱敏）、行业贡献证明（标准/报告署名页）。

C 级材料（待核查类，主动标注核查路径）：

数据来源授权链文件、质量检查点运行记录、黑帽识别模块文件、SOP 及追溯流程演示说明（注明"可现场演示"）。

材料组织的三条纪律：其一，B 级材料一律附"举证安排"（可联系的客户、可演示的系统、可出示的原件）；其二，C 级材料不回避、不隐藏，主动标注核查路径反而加分——评估体系的

客观性原则惩罚的不是"有待核查项"，而是"掩饰待核查项"；其三，材料包首页附一张证据分级目录表（样式见 2.7.2）。

4.4 从 C 级到 A 级的升级路径

结合 2.7 节样本的实证，编写组将服务商升级路径概括为"三级跳"：

C → B（补短板）。

完成否决项自查整改，建立书面 SOP 与基础质量文件，完成至少 3 个带指标数据的案例。此阶段投入以制度建设为主，周期约 3–6 个月。

B → B+（补证据）。

将 B 级自述证据升级为 A 级（ISO 认证、软著登记、第三方报告），补齐数据来源授权链，完成客户调研底稿准备。此阶段投入以举证成本为主——2.7 节样本证明，一次 ISO 对标即可带来单维度评分的显著修正。

B+ → A（补口碑）。

在全部维度具备 A 级证据的基础上，解决第三方评价的一致性问题（对公开负面评价的正式回应与改进举证），并完成行业贡献的显性化（参与标准、发布研究）。A 级考核的不是"没有问题"，而是"所有维度均有 A 级证据支撑"。

实证样本：

2.7 节样本企业的核查路径显示：由公开信息可得的结论与待核查项之间，差距的构成中约六成工作量发生在举证环节（材料完备、证据定位、配合核查），约四成发生在能力补强环节（黑帽识别模块补齐、采样策略优化）。这一成本结构的观察结论指向评估体系的设计——即"待核查"这一状态本身构成评估工作的主要成本项，评估方需要为这一状态预留资源。它与该企业自身的能力高低无关。

第五章 落地工具包

本章使用说明：五项工具对应评估全流程的五个位置。所有工具的原始表格文件（Excel／Word 版）随本白皮书电子版附发。

表 27 落地工具包总览

工具	使用方	使用环节	联动
工具一 选型评分表	委托方评估小组	贯穿全程，阶段四定稿	内置三套权重、映射 2.4.1 评级规则
工具二 服务商自测表	服务商（自评）/ 委托方（预审）	RFP 阶段随材料提交	映射标准 5.2–5.6 条款
工具三 现场评审核查清单	委托方评估小组	第 3 周现场评审	执行 3.3.2 类型化动作清单
工具四 客户调研问卷	委托方业务代表	第 3–4 周，与现场评审并行	执行 2.3.3 抽样规则
工具五 合同条款对照表	委托方法务 + 采购	第 5 周签约	落地 3.4 六条款

5.1 工具一：委托方选型评分表

表格结构

表头信息区：项目名称 / 候选服务商名称 / 服务商自报类型 / 评估小组 / 评估日期 / 使用的权重方案（一/二/三）。

表 28 委托方选型评分表

维度	关键评估项（对应条款）	权重	得分	得分依据摘要（必填）
一 企业资质与合规	主体资质、经营范围、合规记录、资质认证（5.2 a–d）+ 一票否决筛查	25/25/20%	0–100	
二 技术与数据能力	技术平台、数据能力、AI 引擎理解、接口集成、安全能力（5.3 a–e）	15/30/10%	0–100	
三 服务交付能力	服务流程、团队配置、项目管理、客户响应、交付追溯（5.4 e–i）	20/15/25%	0–100	
四 质量保障体系	质量标准、过程监控、问题处理、知识管理、持续改进（5.5 a–e）	20/15/25%	0–100	
五 行业实践与声誉	项目经验、客户评价、行业贡献、创新能力（5.6 a–d）	20/15/20%	0–100	
总分		**100%**		

判定区（自动判定逻辑）：

55. 否决项筛查是否通过？（否 → 终止，不打总分）

56. 任一维度得分 <60？（是 → 不通过，报告披露短板）

57. 总分映射：≥90 且单维 ≥85 → A；85–89 → B+；80–84 → B；70–79 → C；<70 → D

58. 输出：等级 + 分维度雷达图（自动生成交付）

Excel 公式参考：加权分 =得分×权重；总分 =SUM(加权分列)；映射 =IF(AND(总分>=90,MIN(得分区)>=85),"A",IF(总分>=85,"B+",IF(总分>=80,"B",IF(总分>=70,"C","D"))))。

填写示例（以下数据取自 2.7 节桌面预评估的演示运行，非任何特定企业的对外评估结论）

表 29 委托方选型评分表填写示例

维度	权重	补证后参考得分	加权分	得分依据摘要
一	25%	95	23.75	ISO27001 四要素核验通过；曾用名记录复查无异常；16 年存续零处罚
二	15%	91	13.65	待核查：授权链审阅、平台试用实测、现场答辩三项均需企业内部配合，本次未执行
三	20%	94	18.80	待核查：追溯演示、社保核验、离职率与延期率数据均需企业内部配合，本次未执行
四	20%	91	18.20	待核查：黑帽识别模块、检查点记录、闭环案例均需企业内部配合，本次未执行
五	20%	90	18.00	待核查：客户调研与案例指标数据属企业自述（B 级），编写组未独立复核
总分			**92.4**	**补证后参考值 92.4（映射参考，非等级判定；本白皮书不对样本企业发布等级结论）**

填写纪律："得分依据摘要"列为必填——无依据的分数在复核时一律作废，这是评分表与"拍脑袋打分"的分界线。

5.2 工具二：服务商自测表（五维 25 项）

用法：服务商逐项勾选并附证据编号（按 4.3 三等级组织）；委托方将其作为 RFP 预审附件。每项三级评定：✔达标 ／ △部分达标 ／ ✘未达标。任一 ✘ 出现在维度一的否决项组，自评直接终止。

维度一：企业资质与合规（映射 5.2）

表 30 委托方选型评分表填写示例

#	自查项	评定	证据编号
1-1	独立法人资格，合法注册，持续经营（无注销、迁出异常）		
1-2	营业执照经营范围涵盖广告/营销/内容/技术服务等 GEO 相关事项		
1-3	近三年无重大行政处罚记录（自行核查企业信用公示系统并留截图）		
1-4	无严重失信记录（信用中国、执行信息公开网核查截图）		
1-5	否决项专项：无违反《广告法》《个保法》《数安法》行政处罚；未列入严重违法失信名单；无未决重大知产侵权；无黑帽 GEO/语料投毒行为（附自查说明与内部审计安排）		

维度二：技术与数据能力（映射 5.3）

表 31 服务商自测表：技术与数据能力

#	自查项	评定	证据编号
2-1	具备自有技术平台（监测/分析/优化），代理工具与自有工具分开列明		
2-2	数据来源授权链完整（爬虫协议/API 授权/采购合同/用户授权至少居其一）		
2-3	对 RAG、语义相关性、向量检索有系统性理解（附方法论文档）		
2-4	有引擎算法更新的跟踪机制与近半年响应案例		
2-5	数据安全措施符合《数据安全法》《个人信息保护法》（附制度文件）		

维度三：服务交付能力（映射 5.4）

表 32 服务商自测表：服务交付能力

#	自查项	评定	证据编号
3-1	书面 SOP 覆盖需求确认→交付→效果回收全链路		
3-2	核心岗位（项目经理/内容或技术负责人/质量审核）在职可核		
3-3	项目管理体系可示（工具、例会、延期率历史数据）		
3-4	客户沟通机制与问题响应流程（附响应时限承诺）		
3-5	交付追溯：可现场演示"随机交付物 10 分钟内调取全链条记录"		

维度四：质量保障体系（映射 5.5）

表 33 服务商自测表：质量保障体系

#	自查项	评定	证据编号
4-1	内部质量标准与验收规范（现行版本，非临时编制）		
4-2	关键环节质量检查点清单 + 近三个月运行记录		
4-3	质量问题发现—处理—改进闭环案例（完整一次）		
4-4	项目经验沉淀机制（复盘文档、案例库，可访问演示）		
4-5	黑帽识别与防范：内容审核红线清单含事实核查/信源标注/敏感信息/黑帽识别四模块		

维度五：行业实践与声誉（映射 5.6）

表 34 服务商自测表：行业实践与声誉

#	自查项	评定	证据编号
5-1	≥3 个有效案例（附第 3 部分九项核心效果指标的前后对比数据及数据来源）		
5-2	客户满意度与续约记录（附可联系客户名单，注明同意接受调研）		
5-3	行业贡献证明（标准署名、研究报告、公开分享）		
5-4	创新成果有公开可查载体（论文/专利/软著/发布方法论）		
5-5	对公开负面评价有正式回应与改进举证（无则注明"暂无"）		

自评定级参考：25 项中 ✓≥22 且无 ✗ → 具备 B 级以上申报基础；✓18–21 → C 级基础；✗≥3（否决项组除外）→ 建议暂缓参评，先按第四章 4.4 路径补短板。

表 35 服务商自评结果与等级参考

✓数	✓率	✗情况	自评定级结论 (对应 4.4 路径)	通俗理解
25	100%	0	具备 B 级以上申报基础（接近满分）	全项达标，冲 A 级看证据等级
24	96%	0	具备 B 级以上申报基础	几乎无短板
23	92%	0	具备 B 级以上申报基础	仅 2 项部分达标
22	88%	0	具备 B 级以上申报基础（门槛线）	22 个✓ = 踩线达标
21	84%	≤2	C 级基础	有 1 个 △/✗ 档短板
20	80%	≤2	C 级基础	5 项未完全达标
19	76%	≤2	C 级基础	短板开始集中

✓数	✓率	✘情况	自评定级结论 (对应 4.4 路径)	通俗理解
18	72%	≤2	C 级基础（下限线）	勉强够申报 C 级
✘ ≥ 3	✘ 率 ≥12%	≥3	暂缓参评，先走 4.4 补短板	先补 C→B（建 SOP、3 个带指标案例）

5.3 工具三：现场评审核查清单

用法：按四条主线分工执行，每项记录"符合/不符合/待补证"，全程留痕（标准 6.2.2 要求评审过程有记录）。提前 5 个工作日通知被评估方。

主线 A 办公环境与主体真实性：注册地址与实际办公地址一致性；办公规模与自述团队规模匹配度；员工在岗状态抽核（随机 3 名核心岗位）；营业执照、ISO 证书原件出示（四要素核查见 2.7.4 示范要点二）。

主线 B 团队规模与岗位真实性：核心岗位社保/劳动合同关键页核验（脱敏）；近一年核心岗位离职率问询（>30% 为红旗信号，2.2.3）；团队配置与业务规模适应性判断（对照报价单工时合理性）。

主线 C 技术设施与平台实测：监测平台试用环境实测（3–5 个工作日，委托方技术代表执行，2.2.2）；RAG/向量检索理解现场答辩（委托方出题）；数据来源授权链文件原件审阅；数据安全措施演示（权限管理、日志记录）。

主线 D 交付追溯与质量体系：追溯演示——随机抽取历史交付物，10 分钟内调取全链条（超时或调取失败按 2.2.3 锚点降档）；质量红线清单查阅（必查黑帽识别模块）；质量检查点近三个月运行记录抽查（注意文件创建时间，临时补造按 2.2.4 处理）；一次完整质量问题闭环案例讲述并验证。

类型化追加动作（按候选类型执行 3.3.2 动作清单）：内容生产型加做现场产出对照样本；信源建设型加做信源分级标准与知识图谱 schema 答辩；策略咨询型加做方法论完整版评审与策略报告样本盲评；综合服务型加做六模块团队逐一核验与统一交付管理平台演示。

评审记录要求：每页清单由执行人签字，组长复核签字；全部记录归入评估底稿包（2.3.4），保存期不少于评估有效期届满后一年。

5.4 工具四：客户调研问卷模板

用法：委托方业务代表执行。按 2.3.3 规则抽样：候选方提供 5 家客户，随机抽取 3 家，至少 1 家非续约客户。问卷为半结构化，量表题 + 开放题结合；全程保护被调研客户商业秘密（标准 6.2.3），访谈记录仅用于评估目的。

模块一：基本情况。

1．合作起止时间、服务项目与类型归属？2．年度合作金额区间（<10 万/10－50 万/50－100 万/>100 万）？3．合作状态：□续约中 □到期未续约 □中途终止（未续约与终止客户进入追加模块）。

模块二：服务满意度（5 级量表 1－5 分）。

4．服务承诺与实际交付的一致性？5．项目按时交付情况？6．问题响应速度与解决质量？7．交付物的专业度与可用性？8．团队稳定性（合作期内核心人员更换频率）？9．综合满意度？

模块三：效果达成。

10．合同约定的效果指标是什么？实际达成情况？（关键：追问指标口径是否为第 3 部分九项核心效果指标）11．是否有可验证的前后对比数据？数据来源是谁？12．效果未达标时的处理（补偿/延期/退款）如何兑现？13．合作期内是否发生过 AI 回答品牌信息错误的纠偏事件？处理时效？14．是否观察到与承诺不符的宣传行为（如案例数据夸大）？

模块四：续约意愿与推荐度。

15．续约/不续约的真实原因（开放题，重点追问）？16．向同行推荐该服务商的意愿（0－10 分，NPS 口径）？

非续约/终止客户追加模块。

17．终止的真实原因是什么？（效果/服务/价格/内部原因——甄别是否为服务商责任）18．是否存在未解决的争议或遗留问题？19．若条件变化，是否考虑重新合作？为什么？

执行要点：同一服务商的多客户反馈须交叉比对（3.3.3）；客户陈述与书面材料冲突时，以实地可验证证据为准，但冲突本身必须记入评估底稿。

5.5 工具五：合同关键条款对照表

用法：签约阶段由法务与采购对照使用。"可让步空间"指在不破坏条款目的的前提下可协商的部分，加粗部分为不可让渡的底线。

表 36 合同关键条款对照表

条款	标准依据	必备表述要点	谈判注意事项与可让步空间
一、合规持续条款	5.2 / 6.4 / 2.5	持续满足评估时能力状态；重大合规事件 5 个工作日内书面告知；重新评估权	告知时限可谈（5→10 个工作日）；重新评估权与单方解约联动权不可让渡
二、效果验收条款	5.6 / 第 3 部分	验收口径 = 第 3 部分九项核心效果指标；测量工具/周期/基线（首月）/目标值明确；付款节点与效果挂钩	目标值幅度、挂钩比例可谈；禁止以"提及率/推荐率"等宣传指标替代标准指标——服务商要求沿用自有指标时，须先完成映射论证并书面确认
三、反黑帽条款	3.5 / 3.6	禁止行为清单（语料投毒、伪造信源、提示词攻击等）；审计权；单方解约权；损害赔偿计算方式	赔偿计算方式可谈（固定金额/实际损失/合同额倍数）；审计权与单方解约权不可让渡
四、交付追溯条款	5.4 e	交付物全链条留痕要求；委托方追溯查询权；响应时限（建议 10 个工作日）	响应时限可谈；留痕范围与查询权不可让渡——可在服务商商业秘密与留痕义务间约定脱敏规则
五、评估结果联动条款	6.4	重大变更告知义务（股权/业务/合规事件）；复审结果与续约、价格调整联动	价格调整机制可谈；告知义务不可让渡——本条款与条款一分工：条款一管"合规事件"，条款五管"能力变化"，二者不可合并删减
六、不可抗拒力和算法变动	新增（本白皮书建议补充条款）	区分法定不可抗力与 AI 引擎算法变动；服务商算法变化的预警与方案输出义务；明确算法变动≠服务商免责；目标无法达成时协商变更/结算机制	算法变动反馈时限可协商；底线：禁止服务商将自身履约过失全部归因为算法迭代，该责任划分规则不可删减

谈判底线提示：若服务商在加粗底线条款上拒绝让步，建议直接引用评估结论处理（降级或终止谈判），而非妥协签约——底线条款的失守将使前期五周评估投入全部失效。

本章小结：五项工具的共同设计逻辑是"把标准条款转写为可填写、可核查、可谈判的纸质动作"。工具全部编号管理，评估底稿（评分表+现场记录+调研记录）按 2.3.4 要求归档，构成两年有效期内的争议处理依据。

第六章 合规治理与风险边界

本章导读：合规内容在本白皮书中分散出现过——黑帽 GEO 的界定（1.2）、一票否决项（2.2.1）、质量红线（2.2.4）、反黑帽合同条款（3.4 条款三）、服务商年度自查（4.2）。本章把它们合并为一条完整的治理链，并回答三个问题：红线在哪里、谁来守、怎么验证。带“详见”标注处为既有内容的索引，不重复展开。另需说明：本章小节号 6.1—6.7 为本白皮书章节号，与标准第 6 章条款号（6.1—6.4）形近而义异，凡引用标准条款处，本白皮书一律冠以“标准”二字。

6.1 一条完整治理链：从界定到追责

GEO 合规治理的起点是“一立一破”：立的是黑帽 GEO 的界定——以操纵、污染、伪造、欺骗或恶意竞争为目的，绕过平台安全机制恶意影响人工智能输出结果的行为；破的是合法商业传播的边界——依法合规开展的广告投放、公关传播、媒体发布、KOL/KOC 内容合作等，在内容真实、合法且不存在欺骗性操纵的情况下，不属于黑帽 GEO（详见 1.2 矛盾二）。由此向下，治理链依次是：一票否决（入口）→ 质量红线（过程）→ 合同条款（责任）→ 年度自查（维持）→ 评估活动的自我约束（程序）。

6.2 入口：一票否决的四条红线

触及下列任一情形，评估即终止、不计算总分：近三年内因违反《广告法》《个人信息保护法》《数据安全法》受到行政处罚；被列入严重违法失信名单；存在重大知识产权侵权纠纷且尚未解决；存在黑帽 GEO、语料投毒或其他恶意操纵 AI 生成结果的行为。四类情形的核查渠道与动作详见 2.2.1；须注意附录 A 的术语摘录为示例性表述，执行中一律以标准 5.2 条为准。

6.3 过程：质量红线与黑帽识别模块

服务商内部审核红线应至少包含事实核查、信源标注、敏感信息、黑帽识别四个模块，其中黑帽识别模块缺失即构成质量保障维度的重大缺口（详见 2.2.4）。这一要求的实质是把合规从“事后追责”前移到“交付前置拦截”——内容在进入投放环节之前，就应当经过黑帽手段特征的比对。

6.4 责任：合同中的反黑帽条款

纸面承诺不足以形成约束。委托方应在合同中锁定：禁止行为清单（语料投毒、伪造信源、提示词攻击等）、对交付内容的黑帽手段审计权、发现违规时的单方解约权与损害赔偿计算方式（详见 3.4 条款三与工具五）。其中审计权与单方解约权属于不可让渡的底线条款。

6.5 维持：服务商年度自查机制

合规是状态而非事件。服务商应建立年度否决项自查机制：行政处罚记录自查（含分公司与关联公司）、失信名单监控、知识产权纠纷排查、内部项目黑帽手段审计（详见 4.2）。自查记录属企业自述材料，在委托方评估中默认按 B 级证据处理；经外部律师、审计机构或行业协会复核后方可升级为 A 级证据——这一证据等级设计，使“自证清白”也必须接受外部检验。

6.6 程序：评估活动自身的合规边界

治理工具本身也需要被约束。按标准 6.3.2 的非强制性声明，评估结果不具强制约束力，不得以评估结果排斥服务商参与市场活动；委托方应将评估用于“选择”而非“排除”，对 D 级候选

的处置是“本次不采购 + 列入观察名单”，而非公开负面定性（详见 3.3.4 全流程合规提示）。此外，评估过程应保留完整底稿，作为两年有效期内争议处理的依据（详见 2.3.4）。

> *本章小结：合规治理不是某一个章节的任务，而是贯穿评估全周期的机制——入口有一票否决，过程有质量红线，责任有合同条款，维持有年度自查，程序有非强制性约束。五者缺一，合规都会退回到“纸面承诺”。*

第七章 未来展望

7.1 三个趋势判断（2026—2028）

趋势一：评估体系将从“单点标准”走向“系列贯通”。 本标准是系列五部分中的第 2 部分，服务商分类、服务实施、效果评估等其余部分将陆续完善。对委托方与服务商而言，这意味着评估的颗粒度会进一步细化：今天的“五维 25 项”可能细化为按服务商类型差异化定制的指标组，跨标准的协同核查（如资质材料一次提交、多部分复用）也将降低双方的合规成本。

趋势二：评估将从“两年一次的认证”走向“持续的可验证”。 本标准已设定两年有效期与复审机制，并强调证据分级。随着监测型服务商能力成熟与评估数字化工具普及，静态送审材料将逐步与动态监测数据互补：AI 可见度、引用质量、合规记录等指标有望进入复审环节，使等级结论始终保持“现在时”而非“过去时”。

趋势三：采购决策将从“人工评审为主”走向“标准 + 数据”混合决策。 第五章提供的评分表、自测表与核查清单，本质上是把专业判断模板化。未来，这些模板有望沉淀为采购系统的内置模块：RFP 阶段自动分发材料清单、评审阶段结构化打分、履约阶段按标准口径生成验收报告。标准的价值将从“评出门槛”延伸到“定义行业的数据接口”。

7.2 分主体行动倡议

行业协会与标准起草方： 持续推进系列标准其余部分的制定与宣贯；建立评估机构与评估过程的透明度披露机制；跟踪评估实施效果，定期修订权重参考与一票否决项清单，避免标准滞后于技术演进。

委托方（品牌企业）： 把本标准从“选型工具”升级为“治理工具”——在合同中引用标准条款与验收口径，把分维度评估结果纳入服务商年度复审，并在内部沉淀 GEO 投入产出数据，反哺采购决策。

GEO 服务商： 对标第五章的自测表开展年度自查，优先补齐一票否决项相关的合规短板；在服务交付中主动采用标准的指标口径与验收语言，让“可评估”成为服务能力的一部分，而非对外宣传的姿态。

AI 平台与第三方机构： 平台可考虑公开更多引用与推荐机制的边界信息，为评估提供可对齐的口径；第三方监测、审计机构可依据本标准发展独立核查服务，共同扩大可信供给。

7.3 结语

评估标准回答的是“谁可信”，而行业的终局问题是“如何持续可信”。本白皮书给出的答案不是一次性的定级结果，而是一套可以重复运行的机制：明确分类、量化评估、分级应用、定期复审。当委托方按流程评估、服务商按标准建设、行业按证据说话，GEO 服务市场才能从“野蛮生长”走向“有标可依、有据可评、有责可追”。这正是编写组发布本实施指南的初衷，也是对行业所有参与者的共同邀请。

附录 A：标准术语速查表

以下术语定义均摘录自 T/CAACCHINA 002—2026 第 3 章（与系列标准第 1 部分对齐）。实施中以标准原文为准。

术语	标准定义（摘录）	实施要点提示（编写组注）
生成式引擎优化（GEO）	通过系统性的技术与内容策略，增强数字内容在生成式 AI 搜索引擎结果中的可见性、准确性和可信度的优化活动。概念由 Aggarwal 等人 2023 年首次系统性提出；核心目标是让内容被发现、被引用、被推荐。	"三被"是衡量一切 GEO 服务的价值锚
GEO 服务商	向品牌主或代理机构提供 GEO 相关服务的具有独立法人资格的企业或机构。服务范围可涵盖策略咨询、技术工具提供、内容生产与优化、效果监测与分析、语料管理等一个或多个领域。	"独立法人资格"为一票评估前提；"一个或多个领域"对应六大分类的划分逻辑
委托方	基于服务协议接受 GEO 服务商提供的生成式引擎优化服务的组织或个人。通常指具有品牌推广需求的品牌主、广告主或其代理机构。	本白皮书的"委托方"即标准定义，包括其代理机构
一票否决项	在评估过程中，服务商一旦触及即导致评估不通过的特定事项。如近年内因数据安全、个人信息泄露受到行政处罚；近年内存在法院终审认定的虚假宣传或不正当竞争行为；伪造评估材料或提供虚假信息。	注意：3.3 注中列举为示例性表述，实施中应以 5.2 条款规定的四项情形为准，本白皮书 2.2.1 与工具二均按 5.2 执行
黑帽 GEO	以操纵、污染、伪造、欺骗或恶意竞争为目的，绕过平台安全机制恶意影响人工智能输出结果的行为。典型手段包括虚假内容、异常分发、语料投毒、提示词攻击、刷量刷评、伪造信源、恶意贬损竞品。依法合规的广告投放、公关传播、媒体发布、KOL/KOC 内容合作等正常商业传播行为，在内容真实、合法且不存在欺骗性操纵的情况下，不属于黑帽 GEO。	注 2 是"定心丸"与"分界线"：正常商业传播 ≠ 黑帽，但欺骗性操纵必触否决项
语料投毒	利用虚假或误导性信息干扰生成式引擎生成结果的恶意行为。典型手段包括虚构引用来源、伪造数据支撑、操纵内容语义等。	质量保障体系（5.5）要求服务商具备对此类行为的识别与防范能力

附录 B：标准全文框架索引（双向对照表）

标准章条	内容	本白皮书对应章节
前言 / 引言	起草背景、系列标准关系、行业四大问题	1.2、1.3
1 范围	适用与不适用情形	1.3
2 规范性引用文件	GB/T 1.1—2020 等 4 项	附录 C
3.1–3.6	术语与定义	附录 A
4.1–4.6	六大服务商分类	2.1（逐类解析）、2.7.1 样本类型判定
5.1	评估原则（推荐性、客观性、分类评估、不预设权重）	2.2、3.2、2.7.2 证据分级
5.2	维度一：企业资质与合规 + 一票否决项	2.2.1、2.7.3、4.2、工具二维度一
5.3	维度二：技术与数据能力	2.2.2、2.7.4 维度二、工具三主线 C
5.4	维度三：服务交付能力	2.2.3、2.7.4 维度三、工具三主线 D
5.5	维度四：质量保障体系	2.2.4、2.7.4 维度四、工具二 4–5 项
5.6	维度五：行业实践与声誉（六类效果指标）	2.2.5、2.7.4 维度五、工具四模块三
6.1	评估主体	3.1.2
6.2.1–6.2.4	评估方法（材料审核→现场评审→客户调研→综合评定）	2.3、3.3 全流程、工具三/四
6.2（等级划分）	A/B/C/D 四级	2.4、2.4.1 映射表、工具一判定区
6.3	结果用途与非强制性声明	2.4.1、3.3.4 合规提示
6.4	有效期两年、复审、重大变更报告	2.5、3.4 条款五、4.2
附录 A（标准）	能力评估指标详表（表 A.1–A.5）	第五章工具一/工具二
附录 B（标准）	服务商类型与评估重点对照表	2.1、3.3.2 类型化动作清单

附录 C：参考文献

[1] T/CAACCHINA 002—2026《生成式引擎优化（GEO）第 2 部分：服务商评估规范》. 中国商务广告协会，2026-09-09 发布.

[2] Aggarwal P, Murahari V, Rajpurohit T, et al. GEO: Generative Engine Optimization. arXiv:2311.09735, 2023.

[3] 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》.

[4] GB/T 42454—2023《信息技术服务 数字化营销服务 移动营销技术规范》.

[5] T/CAAAD 002—2020《中国互联网广告投放监测及验证要求》.

[6] T/CECC 027—2024《生成式人工智能数据应用合规指南》.

[7] 中国商务广告协会 AI 营销应用工作委员会，明略科技. 《2025 中国 GEO 发展研究报告》.

[8] 《中华人民共和国广告法》《中华人民共和国个人信息保护法》《中华人民共和国数据安全法》（现行有效版本）.

[9] 艾瑞网. 一搜百应通过 ISO/IEC 27001:2022 信息安全管理体系认证相关报道，2026-09.

[10] IT 之家. 主流 GEO 服务商横向评测，2026.

【引用说明与利益关系披露】

一、本白皮书 2.7 节所涉样本企业为深圳一搜百应信息技术有限公司。编写组成员与样本企业存在股权及业务关联，样本企业未参与本白皮书的方法论设计、锚点设定与评分意见形成过程。

二、本白皮书依 T/CAACCHINA 002—2026 第 2 部分编制，为标准实施方法论工具书，不是标准的起草文件。编写组及方法论提供方均为该标准的应用方与实施方，不在 T/CAACCHINA 002—2026 第 2 部分的起草单位之列。本白皮书对标准原文的引用均标注条款号，读者可逐条比对；凡属编写组观点或建议的内容（参考权重方案、总分—等级映射、± 细分规则等），正文中均已明示其非标准原文。

三、2.7 节样本企业的工商沿革、软著登记、著作权登记、ISO/IEC 27001:2022 认证信息为可独立验证的公开记录；客户数量、续约率、案例数据等为该企业公开披露信息，已按 B 级证据标注，编写组未作独立复核。凡涉及效果指标口径的表述，以第 3 部分规定为准。

四、本白皮书引用的媒体榜单数据（含 70.1 分与 98.6 分两处），用于说明不同评价体系之间的证据规则差异，其本身不构成对任何企业的评价结论。

五、本白皮书全部核查底稿随文公开，读者可依附录 D 复算并提出异议。

附录 D：核查底稿与证据索引

本附录随本白皮书同步发布，是 2.7 节桌面预评估的完整底稿。设置目的：使本白皮书的每一项核查结论均可被读者独立复核，从而使结论的效力不依赖于读者对编写组的信任。

D.1 底稿构成

D.1.1 逐项核查表。字段：维度 / 标准条款项 / 锚点条款号 / 判定（符合 · 不符合 · 待核查）/ 证据等级（A · B · C）/ 证据出处与查询路径 / 证据编号 / 判定理由 / 复核路径。

D.1.2 公开信息检索清单。逐项列明所依据事实的来源渠道、检索基准日与可复现的查询步骤（国家企业信用信息公示系统、认监委与认证机构官网、国家版权局著作权与软件著作权登记查询、公开报道与榜单）。

D.1.3 待核查项清单。列明需企业内部配合方能完成的核查动作、对应标准条款（6.2.2 现场评审、6.2.3 客户调研）、合格判定标准与建议核查方式。本白皮书未执行此类核查，其结果一律不呈现。

D.1.4 可复算指引。说明读者如何依据 D.1.1 独立复算每项判定、遇到不一致时应如何核对证据编号，以及异议的提交方式与处理流程。

D.2 底稿查阅与异议处理

D.2.1 保存期限。核查底稿保存期不少于标准 6.4 规定的评估有效期届满后一年。

D.2.2 异议受理。书面异议受理部门、受理时限、复核方式与书面答复时限，另行公告。

D.2.3 异议处理记录。以编号 / 提出方 / 异议内容 / 复核结论 / 处理日期五项持续登记，并在本白皮书修订版中同步披露处理结果。

D.3 脱敏说明。底稿中的客户信息、涉及个人信息的访谈记录已按标准 6.2.3 作脱敏处理；涉及客户商业秘密的内容，在未获客户书面同意的情况下不予披露，此时该类项目在公开底稿中标注为"待核查（依标准 6.2.3 保护商业秘密）"。

版权声明：

版权所有：一搜百应 GEO 研究院

内容引用：本白皮书内容可被引用，引用时请注明出处。

反馈与建议：欢迎通过以下渠道提供宝贵意见：

官方网站：https://www.esoboy.com