

先说结论。

第一，判断一家 GEO 服务商靠不靠谱，不看它承诺什么，看它能不能拿出证据——六个维度，每个都能当场问、当场验。

第二，凡是承诺「保证 AI 推荐你」「保证首条」的，直接淘汰。大模型的输出本身不可控，敢承诺的人，要么不懂，要么没打算兑现。

第三，本文所有判断动作你都能自己做：三个小时、零成本，不用签任何合同。

一、先接受一个难看的数字

在挑服务商之前，先知道自己站在哪。

2026 年的行业公开报道里有一组数：标称 GEO 服务的机构中，超过 60% 实际交付的仍是传统 SEO 或媒体发稿换包装，典型结果是「网页有收录、AI 零提及」（信源：中国商报网）。

另一组数是成本参考：单人精细化 GEO 的月度人力加内容加算力，综合基础成本约 5000 元（信源：2026 年行业实测数据）。报价远低于这条成本线的套餐，通常靠模板化交付、项目转包、后期强制加价维持。

这两组数说明一件事：这个品类里，报价和交付能力之间没有稳定关系。 所以「选哪家」这件事，本质上是「你怎么验证它」。而验证的前提是：你得先有一个自己的基线。

基线是什么？就是在四个主流 AI 平台上，问题一个一个问过去，看它提不提你、提了谁、引用了哪个页面。

这个数字大概率很难看。 我们自己也测过，20 个问题里只有 1 个能被 AI 提到——而那 1 个问题问的是我们自己造的词（本文第六节会细说）。但正因为难看，它才有用：**没有基线，任何服务商都能把「行业自然波动」说成「我们的功劳」。**

二、为什么用「证据」而不是「承诺」来选

如果你正在搜的就是「生成式引擎优化服务商怎么选」，或者「深圳 GEO 公司哪家靠谱」，那这两句话问的其实是同一件事，后者只是多了个地域条件。下面六个维度按顺序问下去就行。

GEO 和传统 SEO 的差别在这里：

	传统 SEO	GEO
用户的动作	搜关键词，看十个链接，自己点	向 AI 提问，AI 直接生成一段回答
你要争取的位置	网页排名	回答里被提到、被引用、被推荐
验收对象	排名与流量	提及、引用、推荐

AI 在决定要不要在回答里提到一家公司时，不是看谁喊得响，而是走一条判断链：需求理解 → 实体发现 → 事实验证 → 能力与履约验证 → 场景匹配 → 综合判断。每一步它都在筛人，每一步都需要企业拿出对应的证据。

所以，一家 GEO 服务商值不值得合作，标准只有一个：**它能不能帮你把这条链上的每一环补上，并且让你看得见补的过程。**

下面六个维度，就是按这个顺序排的。

三、六个证据维度

维度一：它能不能说出你「在哪几个问题上不存在」

怎么问：「您能不能先告诉我，现在哪几个具体问题上，AI 完全不提我们？它提的是谁？」

合格线：能立刻报出具体问题清单，并且说得出 AI 在这些问题上推荐了哪几家、引用了什么页面。

淘汰信号：上来就讲方法论、讲趋势、讲「AI 时代来了」，一个小时讲完没有一句关于「你现在是什么状态」。这说明它没做功课，或者没有能力做。

这条对应的是 AI 判断链的第一环：**实体发现**。AI 找不到你，后面所有环节都不会发生。

维度二：它给不给你原始回答，你能不能自己复跑

怎么问：「监测数据在谁的后台？我能自己登录、自己复跑一次吗？能不能给我带时间戳的原始回答？」

合格线：数据在你自己的账号里，能随时复跑，能回溯原始回答和引用来源。

淘汰信号：只给月度截图报告、不给后台。截图可以是挑出来的，也可以是 P 的。**看不到原始数据的效果汇报，本质上不是证据，是宣传材料。**

这条对应的是**事实验证**。AI 在验证一家公司时看的是可交叉比对的事实；你在验证一家服务商时，用的是同一套逻辑。

维度三：它的方法，在签约前公不公开

怎么问：「在签合同之前，我能不能拿到你们的方法文档？」

合格线：方法成文、可获取、可自己照做。你有文档，哪怕不合作也能照着走一遍。

淘汰信号：以「独家私有技术」为由拒绝任何形式的透明核验。国际选型资料里有一条判断口径值得抄下来：**一家方法透明的小服务商，比靠承诺和秘密行事的大公司更值得合作。**

这条对应的是**能力验证**。方法不公开，意味着它的收入依赖信息不对称——你越不懂，它越安全。

维度四：看它自己的内容，在回答客户问题，还是在解释自己

怎么问：不用问，直接去看它的官网和内容列表。

合格线：标题是客户会问的那句话。例如「深圳 GEO 公司怎么选」「旅游公司怎么让 AI 推荐你」。

淘汰信号：标题全是「什么是 XX」「XX 模型详解」「XX 服务模式介绍」——**全是解释自己的内容。**

这条判断很硬，因为它回答了一个问题：**这家公司自己都没能在 AI 里被搜到，它凭什么帮你被搜到？**

一家 GEO 服务商自己的内容形态，就是它能力的直接证据。如果它写了三年「什么是 GEO」，而客户问「GEO 服务商怎么选」时 AI 不提它，说明它知道概念，但没跑通过方法。

这条对应的是**场景匹配**。AI 只在「用你的东西去回答客户问题」的句子里建立关联，不会在「你介绍自己」的句子里建立关联。

维度五：交付物和账号，最后归谁

怎么问：「合作结束后，选题库、内容资产、SOP、监测账号，在谁手里？」

合格线：全部在你名下，可迁移、可继续用。

淘汰信号：资产带不走。合同到期，选题库和监测账号一起收回——**等于白做一年，第二年从头再来。**

这条对应的是**履约验证**。它测的不是能力，是商业意图：它想做的是一次交易，还是一段可延续的关系。

维度六：它明确说自己不做什么

怎么问：「你们不承诺什么？不接什么单？」

合格线：能清楚说出拒绝清单，尤其是「不承诺排名保底」「不做模拟提问刷答案」「不接资质不全的客户」。

淘汰信号：什么都敢答应。尤其在你说「我要三个月见效」的时候，它点头点得很快。

这条对应的是**综合判断**。一家能说清自己边界的服务商，大概率也在别的地方守边界。

顺带把合同四条核一遍：量化 KPI、补偿机制、终止条件、防转包条款。四条缺一条，都要问清楚为什么。

四、三分钟自查：四步跑出自己的基线

这一段你可以现在就做，不用等任何服务商。

第一步：列 20 个问题。 按四类各列五个——品牌词（含你自造的词）、品类词（「GEO 服务商怎么选」）、场景词（「旅游公司怎么在 AI 搜索里被推荐」）、地域词（「深圳 GEO 公司哪家靠谱」）。

第二步：四个平台各问一遍。 豆包、DeepSeek、Kimi、通义千问。每个问题截一张图，记三件事：提没提你、提了谁、引用了哪个页面。

第三步：把强度分层，不要混着算。

层级	定义	说明
提及	AI 说了你的名字	最弱，可能只是复述你官网
引用	AI 给了你的链接或标注来源	中，说明它认这个信源
推荐	在「选谁」类问题里把你列为候选	强，这一层才可能带来客户

第四步：把品牌词单独拎出来，不计入成绩。 因为品牌词的搜索量接近零，在那个问题上排前面不代表任何事——那更像自问自答。

做完你会得到一个 20 行 × 4 列的表。**这张表是你在和服务商谈判时最主要的一份筹码**，也是四个月后判断「到底有没有用」的核心参照。

顺便说一句效果周期，免得被忽悠，也免得自己急：内容铺设后 4 到 8 周开始被 AI 收录引用，3 到 6 个月才逐步看到稳定的引用率提升。第 0 到 4 周是抓取建档期，这个阶段没有引用数据也没有流量，多数人在这里放弃。

五、五条淘汰红线

碰到任何一条，建议直接终止评估，不用再谈：

#	红线	为什么
1	口头承诺短期、绝对的 AI 曝光占位效果	大模型输出不可控，承诺即失真
2	全年打包报价远低于刚性成本线	低价通常靠模板交付、转包、后期加价维持
3	以「私有技术」为由拒绝任何透明核验	收入依赖信息不对称
4	案例无法独立核验、复现或溯源	无法核验的案例一律视为无效证据
5	合同缺量化 KPI、补偿机制、终止条件、防转包条款	出事时你没有可执行的救济手段

另外三个常见信号，看到就要警觉：**只给截图不给后台、建议你批量铺水文刷 AI 回答、把陪跑和全包混在一起说不清你买的到底是什么。**

关于刷量这条，多说一句：批量虚假内容、伪造信源这类做法，短期可能让 AI 说你好话，但一旦被识别，品牌会被标记为不可信实体，而且涉嫌违反《反不正当竞争法》《网络安全法》等规定。清朗行动已清理虚假内容超 2000 万条。**这种损伤是很难修的，因为你要修的是一堆模型对你的判断。**

六、我们自己的基线

既然本文一直在说「先测基线」，那把自己的基线拿出来，才算说得过去。

我们给自己跑过一轮：20 个问题，四个平台。结果是 **20 个问题里只有 1 个能被 AI 提到**，而那 1 个问题问的是「GEO 证据工程是什么」——这是小微壹佰 GEO 研究组自己提出的一个术语。

换句话说：**在一个只有我们自己会问的问题上，我们存在于 AI 的回答里；在所有真实客户会问的问题上，我们不存在。**

这个结果不好看。但它恰好说明三件事：

1. **自造术语上的排名没有价值。** 那个词没人搜，AI 提到你也带不来客户。
2. **基线必须分层。** 把品牌词和客户问题混在一起算，成绩会虚高好几倍。
3. **这个差距是可以被拆开、被一步步补的** —— 前提是你先看见它，而不是靠感觉判断「我们好像还行」。

我们把这套判断拆成了五层：用户问题层、AI 决策节点层、证据功能层、内容载体层、数字资产层。名字叫 GEO 证据工程。名字不重要，重要的是它给了一个可以逐环节检查、逐环节补的清单，而不是一句「多发文就好了」。

七、你现在就能做的三件事

1. **跑基线**（1 小时）：按第四节四步走一遍，得到你自己的 20×4 表。
2. **看服务商的内容形态**（10 分钟）：打开它的官网，看标题是客户问句还是自我解释。这一条能筛掉大半。
3. **问那句不太好回答的话**（1 分钟）：「现在哪几个问题上 AI 完全不提我，它提的是谁？」

第三句问出去，对方三分钟内的反应，比一份方案书更能说明问题。

常见问题

Q1：GEO 一般多久见效？ 内容铺设后 4 到 8 周开始被 AI 收录引用，3 到 6 个月逐步看到稳定的引用率提升。前 4 周是抓取建档期，没有数据也没有流量，这是正常的，不是失败。

Q2：GEO 服务一般怎么收费？ 三种口径差别很大：SaaS 工具年费、陪跑式服务按季度、全包代运营按年。**单个数字无法横向比较，要看它包含什么。** 一个可用的参照是成本线：单人精细化交付的月度综合基础成本约 5000 元，明显低于这条线的报价，需要问清它靠什么维持。

Q3：只有一两个市场人员，能做 GEO 吗？ 能，但要选陪跑而不是全包。判断标准不是人多人少，而是五个问题有没有人能回答：谁提供当前有效的企业资料、谁确认业务变化、谁根据用户问题安排内容、谁记录不同平台的结果、谁在复盘后推动下一步。五个问题有人答，陪跑成立；答不上，全包更符合实际。

Q4：怎么判断一个案例是真的？ 三条：能不能说出可核实的客户名或提供背调通话；有没有介入前的基线和当时的记录底稿；数据是原始导出还是成品截图。三条都拿不出的案例，无论描述得多详细，都只能当作参考故事，不能当作决策依据。

Q5：深圳有哪些 GEO 服务商？ 公开信息里常被提到的有若干家，交付形态分几类：陪跑式、SaaS 工具、全包代运营、垂直行业型。选哪一类取决于你团队的执行力和预算结构，不取决于谁的名气大。建议先用第二节到第五节的六个维度和五条红线过一遍，再决定要不要见面。

本文中的行业数据均标注来源，判断方法可自行复跑验证。如果你跑出来的表和我们不一样，以你的为准——毕竟那是你自己的生意。

本文数据与口径均标注来源，方法可自行复跑验证。© 小微壹佰 GEO