

METHODOLOGY WHITEPAPER

闭环型 GEO：Agent 自进化、内容资产与留资结算方法论

智多GEO引擎 方法论白皮书 · v2.0

编制与发布：JingYu Media · 智多研究院

方法论的实现系统：智多GEO引擎（主用简称 智多引擎；平台名 智多智能GEO）

版本：v2.0（方法论母版） 发布日期：2026-09-21

文档定位：跨行业方法论母版，行业专版在母版基础上实例化

本 PDF 由官网白皮书页导出，正文与线上版本一致。

目录

- 0. 版权声明、数据来源与免责说明
- 1. 摘要：本文的问题、主张与结构
- 2. 企业已经在 AI 回答里了，钱却没算出来
- 3. 主张：把 GEO 做成一台会自己变强的机器
- 4. Agent 自进化：一次失败怎么变成一条能力
- 5. 内容资产：一次投入，反复使用
- 6. 留资测量：从被提到，到算出谁来了
- 7. 引擎跳转归因：豆包、DeepSeek 各自带来什么
- 8. 一张看板上要有的九个数字
- 9. 成熟度五级：你现在在哪一级
- 10. 实施节奏：先做什么，什么时候看到什么
- 11. 适用边界、合规底线与数据纪律
- 12. 智多研究院的主张与作业红线
- 13. 结语 附录 A-H：术语表 / 可复现清单 / 合规与红线速查 / 数据来源分类 / 常见问题 / 关于智多研究院 / 常见质疑与回应 / 反馈方式

版权声明、数据来源与免责说明

0.1 编制与发布

项目	内容
文档名称	闭环型 GEO：Agent 自进化、内容资产与留资结算方法论
版本	v2.0
发布主体	JingYu Media · 智多研究院
编制单位	智多研究院
官网	ZhiduoGEO.com
发布日期	2026-09-21
适用对象	房地产开发企业、区域品牌方、代运营与咨询机构

本文可以当方法说明读，也可以当验收清单用。第 9 章给出五级成熟度判据，第 10 章给出实施节奏，附录 B 给出可复现清单，附录 E 回答了被问得最多的十二个问题。

0.2 品牌与署名规范

对外出现三个品牌名，分工固定，不能互换。这是本院对外的统一规范，避免同一份材料里出现两种叫法。

品牌名	使用场景
智多GEO引擎	正式全名。首次出现、署名、合同与出版物封面
智多引擎	主用简称。正文叙述与日常沟通
智多智能GEO	平台名，简称“智能GEO”。指产品系统本身

本文正文统一使用“智多引擎”，指代产品系统时使用“智能GEO”。出版主体在封面、页脚与附录 F 三处出现，写法一致。

0.3 本文数字的三类来源

读者有权知道每个数字是怎么来的。本文出现的数字全部归入下面三类之一，混列即为我们的失误。

类别	含义	举例
实跑值	系统里跑出来，可复算	2,420 条问法、660 覆盖网格
估算值	有明确假设的推算	到访比例区间的推演
公开值	第三方公开来源，标出处	公开统计与法律法规原文

推算值一律标注，不与实跑值并列。正文里凡是带“大概”“约”字样的数字，都会写明它的假设是什么。这条纪律在本文内部同样生效。

0.4 术语与主张归属

“闭环型 GEO”的提法，以及三个部件（感知、资产、进化）的命名与定义，由智多研究院提出并在此固定。其他机构的近名提法与本研究院无关，也不存在沿用关系。

核心术语的释义集中在附录 A。首次出现的术语在正文里用「」标出，方便读者对照查阅。

0.5 本文的写作约定

为了让读者读得省力，本院在本文里固定了几个写法。

- 不写公式，不用技术常量。第 4 章讲机制时用的是日常语言，例如把统计上的下置信界说成“一次成功不算数”。
- 不写与结论无关的行业铺垫。需要引用外部资料时，只在用到的那一节标注。
- 术语第一次出现时用「」标出，释义统一收在附录 A，正文内不重复解释。
- 每一章结尾给出下一步可读的章节，方便跳读。

正文里的“本院”指智多研究院。凡涉及我们没有做到的事，文中直接写明“在建”或“不适用”，不做含糊处理。

0.6 免责与引用

本文不构成投资建议，也不构成任何购房决策建议。文中涉及的行业判断基于公开信息与本院实践，可能随政策与市场变化而失效，不构成对未来结果的担保。

引用本文的观点或数据时，请注明“JingYu Media · 智多研究院《闭环型 GEO》v2.0”。需要转载授权、数据来源说明或提交勘误，可通过官网 ZhiduoGEO.com 联系我们。转述数字之前，建议先核对附录 D 的来源说明，避免把估算值当成实跑值引用。

摘要：本文的问题、主张与结构

1.1 一句话主张

企业做 GEO，真正要买的不是一批内容，而是一台会自己变强的机器。

这台机器有三个部件，加一个仪表，分工如下：**看得见**（知道 AI 怎么提我们）、**产得出**（把每次产出留下来，变成下次能直接用的东西）、**改得动**（系统自己发现哪里不行，并且改掉）。仪表是**结算**，把 AI 回答带来的客户一路算到到访和成交。

1.2 四笔算不清的账

问题从一组很常见的困惑开始。企业已经在 AI 回答里被提到，但几乎没有人能把这件好事折算成钱。

- **提到了，人没来。** 品牌出现在 AI 回答里，却没有一个明确的入口把好奇的人接住。
- **人来了，不知道谁送的。** 客户进门，说不清是从豆包来的还是从 DeepSeek 来的，也就分不出该往哪个方向加投入。
- **内容做完就废。** 每一轮投放都从零开始，上一次踩过的坑、写好的句子，下次没人接得住。
- **越做越乱。** 判断好坏的标准每换一个人就换一套，两个月的数字没法摆在一起看。

这四笔账的共同原因只有一个：从“被提到”到“算得出来”之间，缺了可以测量的环节。

1.3 我们给出的三件事

第一件，把可见性做成可复算的数字。同一批问法、同一套判断标准，什么时候重跑都能得出同一个结论。第 4 章讲这套标准怎么被锁住：改一次，历史数据就要重算一次。

第二件，把内容做成资产。一次投入产出的东西，第二个月、第三个月还能直接拿来用，并且越用越准。第 5 章讲资产库由什么构成、怎么复用、什么时候折旧。

第三件，把效果算到人头上。第 6 章讲留资测量，第 7 章讲 AI 引擎跳转归因。留资测量今天已经可用，引擎跳转归因分批上线，本文会写明哪一部分已经能跑、哪一部分还在建。

1.4 本文结构

章	内容	读者能得到什么
2	问题	判断自己是否处在同一处境
3	主张	一台机器该有哪几个部件
4	Agent 自进化	系统自己变强的机制与边界
5	内容资产	资产怎么积累、怎么折旧
6	留资测量	效果怎么算到到访与成交
7	引擎跳转归因	豆包、DeepSeek 各自带来什么
8	看板	该盯哪九个数字
9	成熟度五级	自评当前处在哪一级

章	内容	读者能得到什么
10	实施节奏	先做什么、什么时候看到什么
11	边界与合规	什么情况下不适用
12	研究院主张	我们的作业红线与客户承诺
13	结语	收束与下一步

1.5 本文不写什么

一份方法说明的价值，有一部分来自它主动排除的内容。本文不含下面三类内容。

不含与同行的比较。 本院不以对比他人来立论。读者如果觉得某一节像是在影射谁，那是我方表述不周，欢迎指出。

不含效果承诺。 全文没有任何位置、排名或成交额的承诺。

不含未经标注的推算。 每一个数字都归入实跑值、估算值或公开值三类，见第 0.3 节。

上面三类是本文主动排除的。本文能给出的是判断与验收方式，行业走向不在其中，这是本文的局限。如果确认需要的是前者，建议从第 2 章读起，那一章从客户的真实处境讲起。

企业已经在 AI 回答里了，钱却没算出来

2.1 第一次被 AI 提到，是一件让人高兴的事

场景：一位区域开发商的市场负责人，在豆包里输入“某某板块改善型三房怎么选”。答案里出现了他们项目的名字，还带了一句卖点介绍。

团队把这条截图发进工作群，当天开了个短会。会上有人问了一个大家都答不上来的问题：这条提到，值多少钱？

没有人知道。这不是因为团队不重视，而是因为从“被提到”到“有人上门”之间，缺了可以测量的环节。于是这条提到只能停在截图阶段，变成一次情绪上的好消息，进不了预算表。

2.2 第一笔账：提到了，人没来

数据上看，这件事很常见。品牌出现在 AI 回答里，但回答里没有一条能落到企业自己手上的入口。

用户的动作是这样的：看到答案里有个楼盘名字，去搜索框里再搜一遍。搜索结果的第一个页面可能是竞品的官网、可能是中介平台、也可能是某个自媒体文章。等企业自己的人接到电话，客户早就不知道从哪个环节过来的。

问题不在于 AI 回答不够好，而在于企业没有为“看到答案之后的那一步”做准备。答案里没有可承接的入口，流量就自动流向了别人。这件事无法靠多发内容解决，只能靠把入口补齐。

2.3 第二笔账：人来了，不知道谁送的

假设入口补齐了，客户也留了联系方式。下一个问题紧跟着来：这一单算谁的功劳？

客户进门，说不清自己是从豆包来的还是从 DeepSeek 来的。两边都提到了品牌，两边都可能带来人。没有区分办法，就只能靠感觉分配预算。感觉最不可靠的地方在于：它会偏向最近一次看到的东西。

要区分，不能靠问客户。客户自己也记不住。可区分的前提是：企业在分发内容的时候，就已经给不同来源留了不同的记号。这件事必须在分发之前做完，事后补不上。

2.4 第三笔账：内容做完就废

每做一轮投放，团队都要重新开会、重新选题、重新写文案。

上一轮写好的句子、验证过好用的判断标准、踩过的坑，散落在各个人的电脑和聊天记录里。下一个人接手，等于从头摸索。这类浪费不会显示在任何报表上，但它是成本里最大的一块。

内容只有变成“下次还能直接用的东西”，才算资产。文章的草稿不是资产，因为下次用不上。能反复调用的判断标准、句子模板、禁用清单才是。

2.5 第四笔账：越做越乱

判断标准会漂。这一轮用“被提到就算好”，下一轮改成“要排进前三”，再下一轮又加了一条“必须带上楼盘名”。

标准一改，两个月的数据就不能摆在一起看了。更麻烦的是，团队会得出互相矛盾的结论：上个月说这个板块有效，这个月说无效，其实两次量的根本不是同一件事。

标准不是不能改，但改的动作要有记录，改之前的历史数据要标注清楚。否则系统会一边变强、一边失去可比性。

2.6 一次真实的测量事故

为什么本院对数字这么谨慎，原因来自一次自己查出来的错。

系统在统计引用条数时，出现过一次虚高。顺着记录往上查，问题出在一处兜底逻辑：上游返回的流式响应里，除了真正的结果帧，还有心跳帧和结束帧。兜底代码把任何不含结果字段的响应行，都展开成了一条“标题、链接、正文全为空”的记录。

后果是：每一条问法都会稳定地多出约 1 条空记录。这一批 264 条全部跑下来，引用计数被虚高了约 264 条。

问题没有出在对外发布上：这件事在发布前就被查出来了。它改变的是流程。本院此后要求，任何一个对外引用的数字，都要能回答“它是怎么算出来的”。算不出来的数字不出报告。附录 D 记录了本文全部数字的来源分类。

2.7 这四笔账的连带后果

四笔账单独看都不致命，合在一起会产生三个连带后果。这三个后果比账本身更难处理。

第一个后果是预算错配。分不出哪个渠道带来的客户多，预算就按“感觉哪个渠道要紧”来分。多数团队会把钱投在最近一次看到效果的地方，而不是效果最好的地方。

第二个后果是结论互相矛盾。判断标准不固定，两个季度得出的结论就可能互相打架。上季度说某板块有效，这季度说无效，翻回去看，两次量的不是同一件事。团队会在这种矛盾里失去对数据的信任。

第三个后果是人的流动带来重置。内容与判断标准存在个人手里，人一走，积累就归零。这件事在有经验的写手离职时最容易看出来：新来的人要花几个月重新摸一遍，摸出来的还是同一批结论。

三个后果里，最需要早处理的是第二个。标准不固定，后面所有动作都建立在不稳的基础上。

2.8 判断标准：四笔账能不能算清

本院用一个简单的动作判断企业是否处在上述处境：拿最近一个季度问自己四个问题。

1. 这段时间里，AI 回答提到过我们几次？答案能不能复算。
2. 这些提到里，有多少条带了能落到我们手上的入口。
3. 过来的客户里，有多少能说出是从哪个 AI 引擎看到的。
4. 这一季度的内容里，有多少是复用上一季度的成果。

四个问题里能清楚回答两个以上的企业不多。答不上来的那几个，就是本文后面几章要处理的对象。

本院不主张所有企业都要把四个问题答满。如果业务本身不依赖线上咨询，第二与第三个问题可以不追。但如果企业的主要客源来自线上，这四个问题就构成一套最低限度的自查表。

2.9 结论：问题不在曝光，在结算

四笔账加上一次事故，指向同一个判断：企业缺的不是曝光量，而是一套能把曝光折算成结果的机制。

- 缺入口，就把入口补上。
- 缺区分，就把记号做在分发之前。
- 缺复用，就把产出存进资产库。
- 缺可比性，就把标准锁住再开跑。
- 缺可信度，就让每个数字都能被复算。

下一步读什么，看你的角色：只想知道结论的，可以跳到第 9 章做一次自评，看看自己的团队现在处在哪一级。想先看整套机制怎么设计的，第 3 章给出完整主张。本院建议的动手顺序是先补入口，再补区分，最后才做资产库。

主张：把 GEO 做成一台会自己变强的机器

3.1 一次性交付与长期系统的差别

市面上多数 GEO 服务，本质是一次性交付：写一批内容、发出去、给一份报告，然后结束。

问题出在交付物上：这种模式能解决“被提到”，解决不了“被提到之后怎么办”。因为它交付的是成果，不是能力。成果会过期，能力会积累。两者的差别可以在三个月后看出来，一次性交付的团队要重新开局，长期系统的团队在已有基础上继续走。

本院的主张：把 GEO 做成后者，也就是一台会自己变强的机器。机器的判断标准有两条：同样的投入，第二个月的产出质量应当高于第一个月；同样的目标，第二个月需要的人工干预应当少于第一个月。做不到这两条，就还是一次性交付。

3.2 三个部件和一个仪表

一台能自己变强的机器，由三部分组成。少任何一部分，它都会退化成普通服务。

部件	通俗说法	承担什么
感知	看得见	知道 AI 在各家引擎里怎么提我们、提得准不准
资产	产得出	把每次产出留下来，变成下次能直接调用的东西
进化	改得动	自己发现哪里不行，生成改进候选并决定要不要用

再加一个仪表：**结算**。感知告诉你“被看到了”，资产告诉你“有本钱”，进化告诉你“在变强”，但只有结算能告诉你“值不值”。结算的对象是客户，也就是留资、到访与成交。

四者缺一会出现什么问题，本院在实践中都见过。只有感知没有资产，报告很漂亮但没有产出。只有资产没有进化，库越来越厚但越来越不准。只有进化没有结算，系统会朝着自己定义的目标跑偏。只有结算没有感知，就退回到靠运气投放。

3.3 为什么“会自己变强”是分水岭

一次性交付和自进化系统，在第一个月的表现几乎一样。差别从第一次失败开始出现。

一次性交付遇到失败的处理方式是：等人发现、等人复盘、等人写进下一版方案。这三步都要人来做，而且依赖发现者的水平。自进化系统的处理方式是：失败在发生时就被记录，系统按预定规则生成改进候选，过判定后才被采用。

本院把这个过程叫做「闭环」。它的关键不在于系统会不会改，而在于三件事：改动从哪里来（必须来自真实执行记录，不能来自推测）、谁能决定放行（必须有判定门槛，不能自己说好就好）、改完能不能退回（必须留痕，出问题能退）。

第 4 章把这三件事讲透。这三件事也是本院与“用大模型自动生成一批内容”这类做法的分界。

3.4 这台机器长什么样

把三个部件和一个仪表连起来看，这台机器的工作顺序是这样的。

顺序	部件	输入	输出
1	感知	一批固定问法	AI 在各家引擎里的回答与引用来源
2	资产	引用来源与缺口分析	选题方向、内容成品、可复用的句子与判断
3	进化	每一次执行的结果记录	改进候选，过门后写入资产库
4	结算	客户的留资与成交记录	哪条来源带来了人、值不值

四步之间是顺序关系，不是并列关系。上一代的输出是下一代的输入，中间断一步，后面的数字都会失真。本院在实施时按这个顺序推进，不跳步。

还有一点要提前说明。第一代的产出质量通常不高。系统还没有足够的使用记录，判定门偏保守，资产库也薄。本院建议客户把第一代当基线，而不是当成果。真正值得比较的是第二代与第三代之间的差距。

3.5 这台机器替不了你做的三件事

本院必须把边界说在前面。下面三件事，系统不做，也建议不要交给任何系统做。

第一件，替你定目标。 系统能算出哪个板块的问法更容易带来留资，但不要主攻改善型客群，是经营判断。目标定错了，系统只会更快地朝错误方向走。

第二件，替你承诺结果。 AI 引擎的答案由引擎自己生成，任何机构都无法保证排名，也无法保证提问一定被提到。声称能保证的，请要求对方出具可复算的过程记录。

第三件，替你承担合规责任。 内容合法、宣传表述合规、客户数据使用有授权，责任在发布主体。系统提供检查项与留痕，不替代法务与合规审查。

3.6 客户需要准备什么

先说明：这套机制需要客户投入的东西不多，但有几样不能省。缺了任何一样，节奏会停。

需要准备	具体是什么	缺了会怎样
固定的问法清单	客户业务里最常被问到的问题	测量无从开始
一个确认人	有权确认对外内容的人	放行关卡住
现有内容的清单	已经发过的内容与投放渠道	资产库没有起点
私域数据的授权	线索与成交数据的接入许可	结算层算不出

这四样里，关键的是第二样，原因值得说清：不少项目前期推进顺利，卡在第一次要对外发布内容的时候，因为内部没人敢拍板。本院建议在启动会上直接问一句：这一批内容发出去，谁签字。

第一样也常被低估，问题在于：客户往往觉得“客户会问什么我们都知道”，真正写下来才发现，团队内部对这件事的判断并不一致。把问法写成清单这个动作本身就有价值，它把分歧提前暴露在开始之前。

后两样可以边做边补。内容清单在第一阶段结束时补齐即可，数据授权可以留到第四阶段。这两样晚一点不影响前面的推进。

把这三件事说清楚，剩下的部分才好谈。第 4 章从一次真实的失败开始，讲这台机器怎么改自己。如果你更想先看结果怎么算，可以直接翻到第 6 章。

Agent 自进化：一次失败怎么变成一条能力

4.1 从一条失败记录说起

过程是这样的。系统在给某个板块生成内容时，连续三次调用失败，报错落在同一步：拿不到楼盘的可售状态。

按旧的做法，这件事要等人看报表时才会发现，然后写进下一次的优化清单。本院的做法是，失败发生的那一刻就被记成一条执行轨迹：走到哪一步、卡在哪、错误是什么。

接下来系统做三件事，顺序是固定的：先把这条轨迹转成一个改进候选，内容是“遇到可售状态拿不到时，改用上一次的有效快照，并在输出里标注数据时间”。再把这个候选放进判定队列，等它过门。过门之后，它才被写进正式资产库。

整个过程没有人参与，但每一步都有记录。这就是本院所说的「自进化」：不是模型自己变聪明了，而是系统把用过的经验收下来，变成下次能直接调用的能力。

4.2 三种变化：新长出、修旧、派生

系统的自我改动来自三个方向，各自的触发条件不同。

方向	触发条件	产出
新长出	出现从未见过的失败	一条防错规则
修旧	某条已有资产用起来效果差	该条资产的改良版
派生	某条资产在特定场景下效果突出	它的专用版本

新长出的用途是补洞。系统遇到没处理过的异常，就长出一条防错规则，避免同一个坑重复踩。

修旧的用途是止损。某条资产被调用多次，成功率偏低，系统会生成一个改良版。这里有一条纪律：**改良版不会直接覆盖原资产**。它作为一个新候选参与判定，通过了才替换。原资产在过渡期继续可用，避免一次失败就让某个环节停摆。

派生的用途是放大。某条资产在特定场景下表现突出，系统会把它派生出一个专用版本，继承它的来源记录，但收窄适用范围。这样做的结果是一个库越来越细，而不是越来越乱。

4.3 能不能放行：四道判断

候选生成之后不能直接用。它要过四道判断，任何一道不过就退回候选池。

- 第一道，看结构。** 候选写全了吗。一条能用的能力资产，至少要讲清楚：什么时候触发、按什么步骤做、失败了长什么样。缺任何一项，都不是可执行的东西，只是一句感想。
- 第二道，看语义。** 它说的是不是这件事。候选必须跟触发它的那次失败对得上。如果一条防错规则来自“拿不到可售状态”，却在讲“如何写标题”，那它就是把两件事混在了一起。
- 第三道，看重复。** 库里是不是已经有同一条。如果已有可用的同类资产，系统直接复用，不新建。这条判据的作用是压制库存膨胀：仓库变厚的速度一旦超过变准的速度，维护成本会吃掉全部收益。

第四道，看实战。 它被真正用过之后，有没有产出结果。前三道量的都是候选正文自己写得好不好，只有这一道回答一个更重要的问题：它在真实业务里到底有没有用。第 4.4 节专门讲这一道。

四道判断里，前三道是自动的，第四道要看使用记录。使用记录不够的时候，第四道不给分，候选就留在池子里等。

4.4 小样本不算数

这是本院在适应度判定上最坚持的一条：**一次成功不构成证据。**

道理很直白。一条资产被调用过一次、恰好成功了，说明不了它好用，只能说明这一次没出问题。如果系统据此给它高分，库里会迅速堆满“运气好”的东西。

把这条道理落成规则，本院的处理方式是对成功次数打折，规则如下：一次就成功、之后没再试，按 0.21 分算；试了三次全成功，按 0.44 分算；试到十次全成功，才算 0.72 分。同一个“全成功”的结论，因为试的次数不同，得分差了三倍多。

这条规则的副作用是系统在早期显得保守：新资产要反复使用才拿得到高分。本院接受这个副作用，因为代价是不准，收益是可解释。

还有一个配套设计：先验与实测的配比。系统对新候选有一个初始判断（来自它属于哪个来源、继承自哪条老资产），但实测记录越多，初始判断的权重就越低。用到足够多次之后，得分基本由实测决定。这样的话，任何一条资产的分数都能回答“它是靠历史还是在靠真实表现”。

4.5 放行线比留任线严

新候选要进库，和已入库资产要被清退，用的是两条线。

两条线的松紧不同：新候选要进正式资产库，实测分必须过一条更严的线；已经入库的资产要停用，线设得更宽。本院把这个不对称当作刻意设计：放进来的东西会立刻被用在真实业务上，代价高；踢出去只是回到候选池，代价低。

系统中还有一条硬地板。实测分低于地板的候选，哪怕前三道判断给了很高的评价，也不许放行。这条地板存在的意义是防止一件很具体的事：一条写得很漂亮、但从没被用出结果的能力，因为文字好看而混进生产环节。

这里有一个容易被忽略的地方：录取线严于留任线，会带来短暂的“资产荒”。库里的东西不够用的时候，团队会想放宽标准。本院建议不要放，宁可让某一步先由人工顶上，也不要让没验证过的能力进入生产。

4.6 机器做什么，人管什么

本院把决策点分成两类，界限写进产品，不靠约定俗成。

环节	谁做	原因
从执行记录生成候选	系统自动	量大、重复、规则清晰
候选过四道判断	系统自动	判据可写成规则
候选升格为可用资产	系统自动	过线即升，不需要二次确认
资产状态变更	人确认	停用会影响在跑的业务

环节	谁做	原因
对外内容的最终编辑	人确认	责任在发布主体
回退到上一代资产	人确认	回退是在承认决策失误

这张表可以看作本院对客户的承诺边界：系统不会在你不知道的情况下改动对外内容，也不会自行停掉正在产生效果的东西。反过来说，凡是这张表里标了“人确认”的环节，客户不能要求系统全自动，本院也不会答应。

4.7 标准先定死，再开跑

标准会漂，是自进化系统最容易出的问题。

这里的标准指判断好坏的那套规则：什么算被提到、什么算提到得准、什么算合格内容。一旦这套规则在跑的过程中被改动，之前所有按旧规则算出的分数就不能直接跟新分数摆在一起比较。

本院的做法是给标准加版本号。开跑前定好一版，跑的过程里不动。确实要改的时候，升一版，历史数据标注它属于哪一版。这样做的结果是，任何两个月的数字放一起时，第一个要回答的问题是：它们是同一版标准算出来的吗。

这套做法会带来一点麻烦：想快速试新想法的时候，改标准要走升版流程。本院建议不要绕过它。绕过的代价不是这一次的分数不准，而是此后所有历史数据的可比性被一次性破坏。

4.8 每一代都留痕，随时退得回去

系统里的每条资产都知道自己从哪来。

候选继承了它的来源：是从哪次失败长出来的，还是从哪条旧资产改出来的。这条来源记录一直保留，资产升级时向后传递。出了问题的排查路径因此是清楚的：沿着来源往回走，找到第一次出现的地方。

留痕还支撑一个动作：回退。客户可以退回到上一代资产，系统保留每一代的完整记录。本院要求这项能力常备可用，理由很实际：自动改动跑得越多，出错的可能就越多，没有回退手段的系统最终会被客户关掉。

还有一条边界与留痕配套：**机器只改机器产出**。系统自动生成和自动维护的资产，机器可以改。人工撰写的文案、人工确认过的表述，系统不会自动改写，只会在旁边给出修改建议。

4.9 一次完整的改动走一遍

前面讲的是规则。这一节把一次改动从头到尾走一遍，读者可以对照自己的工作流。

起点：一条失败记录。系统在给某个板块生成月度简报时，连续两次拿不到可售状态，两次都停在同一处。每次执行都会被记成一条轨迹，包含走到哪一步、耗时、结果、错误信息。

第一步，生成候选。系统读完这条轨迹，生成一条改进候选。候选的正文大意是：遇到可售状态获取失败时，改用上一次的有效快照，并在输出里注明数据时间；连续三次失败则中止并提示人工介入。

第二步，过结构判断。候选讲清了触发条件（拿不到可售状态）、执行步骤（改用快照并标时间）、失败特征（连续三次失败）。这一道过。

第三步，过语义判断。候选处理的问题与触发它的失败是同一件事，没有串到别的话题上。这一道过。

第四步，过重复判断。系统检索资产库，发现已有一条“数据获取失败时的降级规则”，但那条只覆盖单一数据源，不覆盖可售状态这一类。判定为不重复，允许新建。这一道过。

第五步，看实战记录。新候选没有使用记录，这一道不给分。它进入资产库之前，先以受限状态存在，只在明确的场景下被调用。

第六步，人来确认。候选涉及对外简报的数据表述，落在“资产状态变更”这一行，需要人确认。负责人确认后，候选转为可用。

第七步，留下来源。系统记下：这条能力由某年某月的一次失败长出来，关联到当时的执行轨迹编号。此后这条能力如果再被改良，来源记录继续向后延伸。

第八步，观察与回退。投入使用的头几轮里，系统继续记录它的使用结果。成功率低于留任线时，它退回候选池；期间如果发现对外表述出了问题，负责人可以在系统里退回上一代。

八步走完，一条失败变成了一条可用的能力，全程有记录。这套流程的价值不在快，在于每一次改动都能被解释、被追溯、被撤回。

4.10 机器不能碰的三样东西

到这一节，本院要把系统主动放弃的能力说清楚。下面三样东西，系统有能力做，但设计上不做。

数字。系统不能自己生成对外引用的数字。所有对外数字只能来自实跑值，或者来自标注了假设与来源的估算值。这条限制直接来自第 2.6 节记录的那次测量事故。

标准版本。系统可以对标准提出修改建议，但不能自己升版。升版需要人确认，因为升版意味着历史数据的可比性要重新处理。

对外发布。系统不能未经人确认把内容发到任何公开渠道。它可以完成撰写、检查、排期、准备发布动作，最后一步的确认由人做。

把这三样交出去，系统跑得会更快。本院选择不交，因为这三样恰好对应客户承担的三类风险：数据失真、结论失效、内容违规。

4.11 这套机制什么时候会失效

任何一种机制都有适用范围。本院把自进化机制可能失灵的情形列出来，供客户在评估时对照。

情形一，执行记录太稀。系统能自进化的前提是有足够多的执行过程被记录下来。如果一个功能一周只被调用几次，候选永远攒不到足够的实测记录，判定门第四道一直不给分，机制等于停在原地。这类低频环节，本院建议继续由人工维护。

情形二，失败被当成噪声清掉。有些团队会把报错当成需要屏蔽的东西，处理方式是加一层静默重试，让错误不再出现。这样做的好处是表面干净，代价是失败记录一起消失，系统失去了改进的原料。本院要求失败必须留痕，不允许静默跳过。

情形三，标准频繁变动。前面讲过标准要锁版本。如果标准一个月升一次，历史数据的可比性会反复被切断，系统的改进判断也会失去参照。标准升版要有理由，不能因为换了负责人就重来一遍。

情形四，人把机器的产出改回去。 自动改动生效后，如果每次都被人工改回原样，系统会不断生成同一条候选。出现这种情况，本院建议先停下自动改动，把分歧整理进标准，再重新开启。

四种情形里，第二种最隐蔽。它不会报错，也不会被指标发现，只会让系统越来越难变聪明。本院建议把它列为巡检项，定期检查失败记录的数量有没有异常偏低。

如果读者在自己的团队里发现了其中任何一种，下一步动作不是加功能，而是先把这个情形解决。机制的前提是原料干净，原料出了问题，改得再多也不会变准。

到这里，机器的“改得动”讲完了。第 5 章讲它的另一个部件：产出的东西怎么变成能反复用的资产。如果你想先看效果怎么结算，第 6 章从留资测量开始。

内容资产：一次投入，反复使用

5.1 资产的定义很窄

本院对内容资产的定义只有一条：下次还能直接用的东西，才叫资产。

按这条标准，一篇发出去的文章不算资产。它下次用不上。一段写好的句子、一条验证过好用的判断标准、一份“这些话不能说”的清单，才是资产。它们能反复调用，用一次多一次价值。

这个定义看起来很窄，但它决定了整件事的经济性。内容生意的成本结构里，最大的一块不是发布，是重新想。资产库要解决的就是“重新想”这件事。

5.2 五类库构成能力存量

本院的资产库由五类内容组成，合计 102 个能力组件。

类别	作用	通俗说法
角色	定义“谁来写”，不同内容交给不同的写作身份	谁来说
模板	定义“按什么骨架写”，保证结构不散	怎么说
技能	定义“遇到什么情况怎么做”，可执行的动作	怎么做
判定标准	定义“什么叫好”，用于打分与筛选	算不算好
禁令	定义“什么不能说”，含平台敏感表述与套话	不许说

五类库各管一段，互相之间不重叠。这个划分的意义在于定位问题：一篇内容不好，能问出是“角色选错了”还是“模板用错了”还是“判定标准太松”，而不是笼统地说“写得不行”。

需要说明的是，这 102 个组件不是一次做完的。它们由第 4 章的机制逐步长出：每一次失败、每一次效果偏差，都可能带来一条新组件或者一次改良。

5.3 问法、角度、网格三层

资产库里最容易被误解的，是三个互相嵌套的数字。本院按从外到内的顺序解释，读者可以拿自己的业务对照。

最外层是 2,420 条问法。 这些是真实用户会问出口的句子，例如“某某板块总价三百万能买什么”。它们不是关键词，是完整的问题。

用关键词和用问法，得到的结果差别很大。关键词是行业内部的说法，问法是客户的说话方式。拿关键词去问 AI，得到的答案里没有客户；拿问法去问，答案才接近客户看到的那个版本。这就是本院坚持建问法库、不建关键词库的原因。

中间层是 22 维。 上面那些问法不是散着的，每一条都能归到某个购房决策的判断角度上，例如总价预算、通勤时间、学区、户型面积、交付时间、开发商口碑。

这一层的作用是做体检。问法按角度归好之后，能一眼看出企业在哪些角度上有内容、在哪些角度上一片空白。多数企业的分布很集中：户型与地段写得多，通勤与交付写得少。空白的角度不是不重要，是没人写。

最内层是 660 个覆盖网格。 它的算法很直白：22 个判断角度 × 6 类购房人设 × 5 个决策阶段。6 类人设指首次置业、改善、投资等不同身份的提问者；5 个决策阶段指从“随便看看”到“准备签约”的过程。

这一层解决的是一个具体问题：同一个话题，不同身份、不同阶段的人问法不一样。“首套房总价三百万怎么选”和“置换改善三房怎么选”看起来相似，实质上要的答案不同。网格让这些差别被显式地列出来，而不是靠写手的直觉去补。

三层合起来回答一个问题：我们到底覆盖了多少种真实情况。数字大不代表好，因为每一格都要有对应内容才算真覆盖。本院建议用这三个数字做自评时，先看网格的空白比例，再看问法总数。

5.4 区域决策切片

合作伙伴常问的一个问题是：你们的楼盘实体库有多少条。

本院对外统一使用「区域决策切片」这个说法，按城市加板块的方式切片，不对外披露精确条数。原因有两条：一是这类数据更新频繁，公布的数字很快会过期；二是库的大小和效果没有直接关系，能被调用的切片才有意义。

放到使用上，切片的价值在于让内容落到具体板块。写“某市楼市口碑”是无效的，写“某市某板块改善型三房的选择逻辑”才有可能被 AI 引用。

5.5 八个环节与四个关口

资产从无到有经过八个环节。本院按顺序列出，并标出每个环节的产出。

顺序	环节	产出
1	探照	知道 AI 在各家引擎里怎么提我们
2	采集	拿到可用的引用与来源记录
3	评估	算出当前的可见性水平
4	选题	定下这一轮做哪些方向
5	生产	得到内容成品
6	质检	确认能对外发
7	分发	内容进入公开渠道
8	结算	算清这一轮有没有效果

八个环节里有四个关口，每一项都要通过才能往下走。这四个关口是本院在实践里固定下来的，也是验收时的检查点。

- **采集关**：拿到的数据能不能用。来源不可追溯的数据在这里被拦下。
- **选题关**：这个方向值不值得投。没有真实问法支撑的方向在这里被拦下。
- **放行关**：内容能不能对外发。表述不合规、与事实不符的在这里被拦下。
- **结算关**：这一轮有没有效果。算不出效果的，不能进入下一轮的放大清单。

四个关口都挡住过东西。一个从不拦东西的关口等于没有关口，本院把它当作流程失效的信号。

5.6 资产会折旧

资产不是放进库里就永久有效。

环境会变：平台的引用偏好会变，用户的问法会变，政策表述会变。一套三个月前很好用的句子，今天可能已经失效。

本院的处理方式是给资产标上时间。内容有新鲜度标记，超过一定时长的内容在调用时会被提示。判定标准与模板的更新周期更长，但同样有标记。

折旧机制带来的一个结果是：资产库需要一个维护动作。本院建议把维护作为固定安排，而不是等效果下滑了再回头清理。具体节奏见第 10 章。

5.7 批量生成为什么会失效

有一种做法听起来很划算：用大模型批量生成几千篇内容，一次铺满。

本院不建议这么做，原因不是成本，而是效果会自我抵消：同一个模型、同一批提示词产出的内容，在结构、句式、论证顺序上都很接近。AI 引擎在挑选引用来源时，会优先选择信息密度高、彼此不重复的来源。大批同质内容的结果是：它们互相稀释，谁也没被选中。

禁令库里的 11 条规则就是针对这件事设的。它限制的不只是违规表述，还包括套话与固定句式。一条内容如果需要靠套话撑起篇幅，说明它本来就没有信息量。

本院给出的替代做法有两条：一是把生产量降下来，把每个方向做深；二是把资产库用起来，让复用替代重写。这两条都比批量铺量慢，但产出会被引用。

5.8 资产库需要有人维护

资产库不是建完就放着。本院把维护动作定成三个，按固定节奏做。

第一个动作，看调用记录。 哪些资产被用得最多，哪些从来没有被调用过。长期闲置的资产占位置，也占检索时间。

第二个动作，看失效迹象。 同一类内容的引用比例在下降，通常说明这一类的表达方式已经跟不上平台偏好的变化。这一类的判断标准要重做。

第三个动作，合并重复。 判定门会拦掉大部分重复，但语义相近、措辞不同的条目仍会积累。合并的目的是让同一个判断只存在一条，避免调用时出现两个版本。

三个动作里，第一个最容易做，也最容易被忽略。本院建议在月度例会上固定花一点时间看调用记录，把从来不用资产清出去。库小一点、准一点，比大而全更有用。

5.9 一套内容从选题到进库走一遍

前面讲的是分层。这一节把一份内容从选题到进资产库的过程走一遍。

第一步，从缺口出发。 感知层算完，发现企业在“交付时间”这个角度上没有可被引用的内容，而这个角度下的问法数量不小。这个缺口成为本轮选题的第一优先。

第二步，查库。 系统先检索资产库，看有没有现成的模板、句子与判断标准可以复用。假设找到一个通用结构模板，以及三条关于交付说明的句子。复用从这里开始，不从空白页开始。

第三步，生产。 写的人在这个基础上补，把需要事实支撑的部分填实。系统提供方向与骨架，事实由客户提供。

第四步，过质检。 检查表述是否合规、数据是否有出处、有没有踩禁令。这一步不过就回去改，不改标准。

第五步，分发与记录。 内容对外时带上入口标记，同时把这一批的选题依据、用到的资产、预期结果记下来，作为后续对照的基线。

第六步，回库。 这一轮新写出来的好句子、验证有效的判断，经过判定进入资产库，成为下一轮的起点。

六步走完，一份内容既完成了当下的任务，也留下了下次能用的东西。本院衡量这一步是否做好的方式很直接：下一轮生产时，有多少是从库里取出来的。

到这里，机器“产得出”这一半讲完了。第 6 章开始讲仪表：效果怎么算到人头上。如果更关心 AI 引擎各自带来多少客户，可以直接翻到第 7 章。

留资测量：从被提到，到算出谁来了

6.1 唯一算得准的数字来自客户

前面几章讲的都是站在企业外面看：AI 有没有提我们、提得准不准。这些观察有用，但都有一个共同的弱点：它们是间接的。

真正算得准的数字只有一个来源：客户自己。他留了联系方式、他来了、他签了，这些动作发生在企业的系统里，不是推断出来的。

本院把这一层叫结算：它在整台机器的最后，也是唯一能回答“值不值”的地方。第 6 章讲留资怎么变成可算的数字，第 7 章讲怎么把客户跟 AI 引擎对应起来。

6.2 漏斗三段：线索、到访、成交

线索从进门到成交，经过三段。每一段的定义本院都写得具体，避免不同人算出不同的数。

阶段	判定条件	谁记录
线索	留下可联系的号码，且通过有效性检查	系统或销售
到访	客户实际到了现场，或完成了线上深聊	销售更新
成交	签署认购或合同，金额写入	销售更新

三段之间是单向流转，不能跳级。系统只认记录，不认口头判断。这一条看起来机械，但它解决了一个很常见的麻烦：月底对账时，销售说“这个客户基本定了”，报表里却还是“到访”。金额不上表，回本进度就算不出来。

6.3 没有对照组，只能叫相关

留资测量里最难的一步，是判断“这笔成交跟 GEO 有没有关系”。

本院的态度写得很直白：**在拿不出有效对照组的情况下，只能说相关，不能说因果。**

落到产品里，这条态度变成一个封顶动作：当客户来源信息缺失到一定程度时，系统算出的到访归因分最高只给到 70 分，并在报告里注明“相关性而非因果”。70 分不是一个随便定的数字，它的含义是：数据能说明这件事值得继续，不能说明这件事是它造成的。

本院宁可在报告里显示不那么好看的数字，也不愿把相关性包装成因果。原因很实际：一旦把相关当因果，下一轮的预算分配就会建立在错的前提上。

6.4 天然对照组

有没有不用做实验就能拿到的对照组？有，而且它就在数据里。

客户的来源可以分成两类：一类自述是从 AI 回答来的，一类是其他来源（搜索、社交、线下、说不清）。这两组天然构成对照。把两组的到访比例摆在一起，差值就是本院使用的一个指标。

这个做法的优点是诚实：它不是让系统去证明因果，而是让数据自己说话。缺点也清楚：客户自述不一定准，有人记错了，有人不愿意说。所以本院给它配了一个前提条件：自述来源的覆盖比例要过一条线，才计算这个差值。覆盖不够时不算，报告里显示“数据不足”。

6.5 留资质量分

线索多不等于效果好。一百条打不通的号码，比二十条能约到现场的线索更没有价值。

本院对留资质量打一个分，由四个部分组成。

组成部分	看什么	权重
号码有效性	能否接通、是否重复、是否同行	0.30
意向等级结构	意向 A 与 B 的占比	0.25
线索到访率	留了电话的人有多少真来了	0.25
到访成交率	来了的人有多少签了	0.20

四个部分的权重是固定的，不随客户调整。本院这样定的原因是：如果权重可以按客户的偏好改，这个分就不能跨客户比较，也就失去了作为参考的意义。

无效号码会被打上标记，标记内容包括号码无效、重复提交、同行探盘、代运营骚扰。打了标记的线索仍然保留在库里，但进入质量分时会先被剔除。保留是为了可核对，剔除是为了不失真。

6.6 把这笔账算出来

有了三段漏斗，算账就变成四步。

第一步，算出单个线索的成本。 用这一轮投入除以有效线索数。这一步最容易出错的地方是把无效号码也算进分母，本院在计算时会剔除。

第二步，算出成交金额。 把这一轮里所有成交线索的金额加总。

第三步，算回报。 用成交金额减去投入，再除以投入。结果是负数时，本院不会在报告里做任何修饰。

第四步，算回本进度。 用成交金额除以预设的投入目标。这个数字回答的问题是“投进去的钱回来了多少”。

本院建议客户在开始之前先写下一个投入目标，而不是事后补。事后补的目标通常会被现场表现影响，失去参照作用。

6.7 数据从三条通道进来

私域数据不会自己长到系统里。本院设计了三条约定的通道。

第一条是手工录入。 销售在系统里直接录。适合线索量不大的情况。

第二条是批量导入。 从客户现有的表格导入，分两步走：先预览，确认字段对应关系没问题，再正式提交。整批可以一次性回滚，避免导入错了要一条条删。

第三条是表单回传。 客户的表单系统在客户留资时把数据实时送回。这条通道带签名校验，防止伪造；同一批可以重复提交而不会产生重复数据。

三条通道之上有三条安全线，本院把它当作硬要求，不接受协商。

安全线	具体做法
授权	客户未开启数据授权时，不采集、不出分
最小必要	授权未覆盖手机号时，该字段在入口处直接丢弃
脱敏	号码进入系统后立即加密存储，明文不留在库里；姓名只保留姓氏
可核对	每条线索带统一去重标识，任意批次可整体回滚

脱敏这一条需要单独说明：号码加密用的是“客户编号加号码”一起算的，同样的号码在不同客户名下会得到不同结果，因此不能用系统去做跨客户比对。这是本院有意设置的限制。

6.8 漏损定位

算出一个总分之后，更有用的动作是找出问题在哪一段。

本院的做法是把三段之间的转换比例分别列出来，和上一周期对比，看是哪一段掉下来。

- 如果线索到访率掉，问题通常在触达内容与意向匹配上。
- 如果到访成交率掉，问题通常在案场，不在内容。
- 如果号码有效性掉，问题通常在选择投放位置上。

第 10 章会给出这套定位动作在实施节奏里的位置。它的价值在于把“效果不好”这种笼统判断，换成一句可以派活的话。

6.9 一个季度的账怎么算

规则讲完之后，用一个例子走一遍。下面的数字是示例，不是任何客户的真实数据。

假设某项目一个季度的投入是 30 万元。这一季度里，三条通道合计进来 210 条线索，剔除无效号码与重复提交 30 条，有效线索 180 条。

第一步，单个线索成本。 用 30 万除以 180，得到每条有效线索约 1,667 元。如果分母用了 210，这个数会变成 1,429 元，看起来更好，但它是错的。

第二步，漏斗转换。 180 条有效线索里，到访 36 条，成交 4 套。线索到访是 36 比 180，到访成交是 4 比 36。这两段要分开看，因为它们指向的问题不同。

第三步，成交金额。 4 套合计 1,200 万元。

第四步，回报与回本进度。 用 1,200 万减去 30 万，再除以 30 万。回本进度用 1,200 万除以预设的投入目标。

第五步，质量分。 号码有效性、意向 A 与 B 的占比、线索到访率、到访成交率四项按固定权重合成一个分。

第六步，来源拆分。 180 条线索里，能归到具体引擎的假设是 70 条，另外 110 条归入未归因。本院的做法是只对 70 条做引擎间比较，不用这 70 条的比例去推算全部 180 条的分布。

第六步最容易做错，问题在这里：把可归因部分的比例用到全部线索上，会让报表好看很多，但那就是编数据。本院建议客户在看报告时先问一句：这里的比例是只算可归因部分，还是推到了全部。

6.10 三条通道怎么选

第 6.7 节列了三条数据通道。选哪一条，取决于线索量与内部系统的条件。

情形	建议通道	原因
每月线索在几十条以内	手工录入	建通道的成本高于收益
已有台账或表格系统	批量导入	复用既有数据，导入可回滚
已有表单或客服系统	表单回传	留资即入库，时效最好

三条通道可以并用。本院见过的常见组合是：表单回传承接线上留资，批量导入承接历史与线下记录，手工录入补漏。三条并用时，去重靠统一标识完成，不需要人工逐个核对。

有一条纪律要说明：通道数量不改变质量分的算法。不管线索从哪条通道进来，进入质量分之前的清洗动作没有区别。这一条排除了“换个通道分数就变好”的可能。

如果客户的线索量在未来增长，本院建议先把通道从手工切到表单回传，再考虑批量导入。切换时不需要重做历史数据，统一标识会把老数据与新数据接起来。

6.11 什么时候不出分

本院设定了三种不出分的情形，全部显示“数据不足”而不是零分。

第一种，授权没开。第二种，有效线索数太少，样本不足以支撑结论。第三种，客户自述来源的覆盖比例过低，对照组不成立。

不出分看起来像系统能力不足，本院把它当作正常状态，理由很直接：少给一个分数，比给一个会误导决策的分数要好。如果客户的报告连续几个月显示“数据不足”，本院建议先解决数据通道的问题，再谈优化，下一步动作从第 6.7 节的三条通道开始排查。

引擎跳转归因：豆包、DeepSeek 各自带来什么

7.1 为什么看浏览器来源一定不准

一个直觉做法是：用户从 AI 回答点进官网时，浏览器会带上来源信息，读一下不就知道是哪个引擎了。

这件事在实践中不成立，原因有两条：AI 问答引擎给出的引用链接，多数不携带可识别的来源标识；有的还会经过中间跳转，把来源信息洗掉。指望浏览器来源来区分豆包和 DeepSeek，结果会是一大片“直接访问”。

本院把这条路排除掉了。不是因为实现难，是因为它建立在别人的行为上，对方改一次跳转规则，整套统计就失效。

7.2 可行做法：在分发时给入口贴标记

可靠的出路只有一条：把区分信息放在我们自己能控制的地方，也就是链接本身。

具体做法是，同一条内容在不同渠道分发时，使用带不同标记的入口链接。这个标记很短，不长到引起平台注意，但足以区分来源。用户从 AI 回答点进来的那一刻，标记随链接一起抵达，不需要任何引擎配合。

标记的设计要回答三个问题：这条内容是从哪个渠道发出去的、属于哪一批、对应哪个目标方向。三个信息都短的标记，才能在不改变链接外观的前提下工作。

本院必须说明这条路的边界：它能区分“内容从哪条渠道被引用”，不能区分“用户问了什么问题”。后者要靠第 6 章提到的客户自述问题字段来补。

7.3 落地页要承接什么

标记到了落地页，还需要有人接住它。这一步是目前工程上的主要工作量。

落地页要做的动作有三个。第一，打开时读取标记，识别它属于哪一批内容。第二，把标记暂存在用户本地，避免用户在站内翻几页之后丢失。第三，在客户提交表单时，把标记与客户信息一起送出。

本院对落地页还有一条排版上的要求：首屏必须能看清“这是什么、能解决什么、下一步做什么”。标记能算清来源，但前提是用户愿意留下来。一个加载缓慢、重点不清的落地页，标记再准也接不到人。

7.4 结算方式：算人，不算次数

标记带回来的数据，本院按“人”结算，不按“次数”结算。

指标	含义	用途
跳转量	带某个标记的入口被打开的次数	判断内容被引用的活跃度
到访量	该来源带来的有效到访人数	判断哪个引擎的用户更接近成交
留资量	该来源带来的有效线索数	分配下一轮投入的主要依据
留资质量	该来源线索的质量分	避免只看数量

本院不建议按“跳转量”分配预算。跳转量大而留资质量低，通常说明这个方向的用户离决策还很远。真正该加投入的是到访与留资双高的方向。

7.5 今天已经能做到什么

本院在这里把现状说清楚，不用模糊表述。

已经可用的部分：

- 渠道级归因。线索记录里带渠道标记，能区分是诊断环节、内容环节、发布环节还是手工录入带来的。
- 平台触达基线。系统按引擎统计内容的触达次数，能看出同一批内容在不同引擎之间的活跃差异。本文把它叫「触达基线」，因为它是触达量，不是客户量。
- 客户自述来源。线索记录里有一个来源字段，选项包括 AI 回答、搜索、社交、线下、说不清，由销售在与客户沟通时确认。
- 到访归因与质量分。第 6 章的对照组分析与留资质量分已经在跑。

分批上线的部分：

- 带标记的入口链接自动生成：随内容分发动作一起产出，不需要人工拼链接。
- 落地页的标记读取与暂存，以及表单回传时带上标记。
- 线索记录增加引擎字段，并把未归因部分单列。

本院把这三项列为在建，是因为它们还没有在全量客户上实际投用。已经能跑的部分按上一版可用；还在建的部分按分批上线推进，具体进度以交付时的版本为准。我们不把在建项写成已具备。

7.6 已知失真与四种限制

即使机制建成，本院也要把失灵的情形列清楚。客户在验收时应当把这些情形写进预期。

二次跳转丢标记。 用户从 AI 回答点到一篇第三方文章，再从文章点进官网。如果那篇文章里的链接没带标记，标记就在这里丢失。

跨设备断链。 用户在手机上看到 AI 回答，回头用电脑搜索再进站。两次动作之间没有可接续的标识。

多引擎同时提及。 同一次提问可能被多家引擎同时提到，用户究竟看了哪一家，从数据上分不出来。这类线索计入“不确定来源”。

平台清洗参数。 部分平台在展示外链时会剥掉参数。被剥掉的部分同样计入“不确定来源”。

本院对这些情形的处理方式是：不摊派。归不到具体引擎的线索，单独列一项，不按比例分给各家。摊派能让报表好看，但会让分配决策建立在编造的比例上。

7.7 把这三项做出来要动什么

客户在做技术评审时，通常会问“这件事要改多少东西”。这一节把改动面说清楚。

第一处，分发环节。 内容对外时，入口链接由系统自动生成，同一篇内容对应多个带不同标记的链接，标记与这一批内容、这一条渠道绑定。这里不能靠人工拼参数，否则一定会漏。

第二处，落地页。 落地页打开时读取标记，识别它属于哪一批；把标记暂存在用户本地，避免站内来回翻页后丢失；表单提交时随客户信息一起送出。三件事都在落地页内完成，不依赖第三方统计工具。

第三处，线索记录。 线索表增加一个来源字段，记录引擎标识与批次号。原有字段保留，新增字段为空时不影响老数据。

第四处，未归因单列。 报表增加一项“未归因”，把标记缺失的线索单独统计。这一项的存在比它的数值更要紧：它让可归因部分的占比随时可见。

四处改动都不涉及算法，工作量集中在两件事上：落地页的表单接续要接通，线索表要允许新字段。本院建议客户在评审时重点确认这两件事的负责方与时间，而不是笼统地问整体工期。

还要提醒一句：标记机制生效的前提是分发动作走系统。手工发出去的连接带不上标记，也就无法参与引擎归因。

7.8 标记怎么设计

标记看起来是个小事，设计不好会让整套归因失效。本院把三条要求写出来。

第一条，短。 标记挂在对外链接上，过长会引起平台注意，也可能被截断。本院的做法是只用少量字符承载三层信息：哪一批内容、哪条渠道、对应哪个方向。

第二条，不重复使用。 同一批内容的标记不复用给下一批。复用会让两批数据混在一起，事后无法拆开。本院要求标记带上批次号，批次号只增不退。

第三条，不带隐私。 标记里不放任何与个人有关的信息，也不放内部项目名称。它的作用只有区分来源，不需要更多信息。

标记的生成必须由系统完成，随内容分发动作一起产出。本院不建议让运营同事手工拼参数。人工拼参数一定会出现错拼、漏拼与重复使用三种问题，而且事后很难查出来。

如果客户已经有自己的统计工具，本院建议标记由服务方与客户方的工具各自独立记录，两边结果交叉核对。核对不一致时，先查标记生成环节，再查上报环节。

7.9 触达基线是先行指标

既然引擎级归因还没全量投用，现在用什么看趋势？用触达基线。

它的逻辑是：内容在一个引擎里被引用的活跃度先变，客户量随后才变。所以看趋势用触达，做结算用留资。两者的关系是“先行与滞后”，不是“替代”。

本院把这条写进方法论的原因是防止一种误用：拿触达量的增长去汇报效果。触达增长只能说明内容被引用了，不能说明有人来了。这两件事在报告里必须分开列。

到这里，结算这一层讲完了。第 8 章把前面所有指标收拢到一张看板上，并说明哪些数字不该出现在上面。

一张看板上要有的九个数字

8.1 数字分四层

看板上的数字如果平铺在一起，读的人会失去重点。本院按层次划分，每层回答一个问题。

层	回答的问题	指标数
可见性	我们在 AI 里被提到得怎么样	3
资产	我们的本钱在变厚还是变薄	2
留资	有没有人真的来了	3
结算	值不值	1

四层之间是因果关系，不是并列关系。可见性是入口，资产是产能，留资是结果，结算是结论。看板上从上往下读，就能从“被提到”一路看到“回本”。

8.2 九项指标

本院把九项写出来，并说明每一项该被怎么读。

可见性层

- 1. **提及率**：在一批固定问法里，有多少条回答提到了品牌。这是最基础的指标，也是最容易被误读的指标，因为只算提到、不算提得准。
- 2. **前三提及率**：品牌出现在回答前三个位置的比例。它比提及率更接近真实曝光。
- 3. **引用来源构成**：AI 引用的是哪些来源。这一项决定后面所有动作的方向，也是本院最看重的一项。

资产层

- 4. **资产复用率**：本轮产出的内容里，有多少比例来自资产库而非重新撰写。它衡量的是这台机器有没有在积累。
- 5. **自动改动通过率**：系统生成的改进候选里，有多少过了判定门。它有两个读法：过低说明门槛太严，过高说明门槛太松。

留资层

- 6. **有效线索量**：剔除无效号码与重复提交之后的线索数。
- 7. **线索到访率**：留了联系方式的人里，实际到访的比例。
- 8. **留资质量分**：第 6.5 节那个由四部分构成的分。

结算层

- 9. **回本进度**：成交金额与预设投入目标的比值。

8.3 领先与滞后

九项指标里，有的先动，有的后动。分清这一点，能避免把正常滞后当成失败。

引用来源构成属于领先指标。公开研究显示，主流 AI 引擎的引用来源构成本身波动很大，周与周之间的变化幅度可以达到三到四成，而品牌提及的变化幅度明显更小。这个事实的用法是：看来源构成的变化，可以比看提及率更早发现趋势转向。

成交属于滞后指标。从内容被引用，到客户留资，到访，再到签约，每一段都有时间消耗。用成交来判断这一周做得好不好，结论一定是错的。

本院的建议是分层看：来源构成看周，资产指标看月，留资与结算看季度。

8.4 波动纪律

单周不看趋势。这一条本院写进交付规范。

理由很直接：样本量小的指标，单周的波动大部分来自噪声。一周的数字掉下来就调整策略，等于在噪声上做决策。本院的处理方式是看四周移动，并且只在样本量过线时才出分。

还有一条纪律是标注来源。看板上每一项数字都要能回答“它是实测还是推算”。混合显示会让人误以为所有数字的分量一样。

8.5 不该出现在看板上的数字

有几类数字本院不建议放上去，即便它们看起来很能说明问题。

没有出处的行业百分比。 凡是引用“行业平均”的说法，必须能说出样本与出处，否则不放。

把触达写成留资。 触达是内容被引用的活跃度，留资是客户。两者混列会直接导致预算错配，第 7.7 节已经把这条单独讲过。

把推算值写成实测值。 这是最需要防的一条。本文第 0.3 节把数字分成三类，看板上同样适用。

总量型虚荣指标。 例如“累计生成内容多少万字”。这类数字只增不减，读不出好坏。

看板的价值不在于数字多，而在于每一个数字都能导向一个动作。本院建议在新增一项指标之前，先回答一个问题：如果它变差，我们要改什么。答不上来的，先不放。

8.6 例会怎么读这张看板

指标定好之后，还需要一个读的规则。本院建议按下面这个顺序过一遍，每项只回答一个问题。

第一步，看来源构成有没有变。 变了就往下追是哪个渠道进来的，没变就直接跳到第四步。这一项动得最快，最值得每周花时间。

第二步，看资产指标。 复用率与自动改动通过率。这两项回答“这台机器是在积累还是在原地打转”。连续几个周期复用率不涨，说明资产库没被用起来。

第三步，看留资层。 有效线索量、线索到访率、留资质量分。这一层看月度，不看周。

第四步，看结算层。 回本进度与未归因占比。这一层看季度。

第五步，只定一个动作。 例会结束时只定一件事：下一轮把资源往哪里挪。定三件事等于没定。

本院建议按“来源构成看周、资产指标看月、留资与结算看季度”的节奏安排会议。把四个层的时间尺度混在一起讨论，会得出互相打架的结论：一层在进步，一层在波动，会上就会吵起来。

如果某一层连续多个周期不动，本院建议先检查这一层的数据是不是断了，而不是先调策略。数据断了的情况下，策略调整拿不到任何反馈。

8.7 看板上的数字从哪里来

每一项指标都要能回答数据从哪来，否则看板会变成猜谜。本院把对应关系列出来。

指标	数据来源	采集方式
提及率	固定问法在各引擎的回答	系统定期提问并解析
前三提及率	同上	记录品牌出现的位置
引用来源构成	回答里的引用链接	按来源归集
资产复用率	内容生产记录	统计复用与新建的比例
自动改动通过率	执行轨迹与判定结果	按候选统计
有效线索量	线索记录	剔除无效与重复
线索到访率	线索记录的状态流转	按状态统计
留资质量分	线索记录四个部分	按固定权重合成
回本进度	成交记录与投入目标	按季度统计

这张表同时是一份验收清单。客户在验收时可以对每一项问一句：如果上面这个数据源断了，这个指标会显示什么。正确的回答是显示“数据不足”，不是显示零。

为什么这一点重要：零是一个看起来确定的数字，它会被人当成结论用。数据源断掉却显示零，会让团队去追一个不存在的问题。本院要求所有缺失都显示为不确定状态。

第 9 章把这套指标换算成五个成熟度等级，方便团队做自评。看过第 8.2 节的九项指标之后再读第 9 章，会更容易对上位置。

成熟度五级：你现在在哪一级

9.1 五级判据

本院用五个等级描述企业的 GEO 成熟度。等级不看投入多少，看有没有形成机制。

等级	名称	标志	常见缺口
L1	有内容	有人在持续产出内容并对外发布	不知道有没有被 AI 提到
L2	有测量	按固定问法定期测量，标准固定成版	产出用完就散，下次重来
L3	有资产	产出进入可复用的库，有复用比例	系统不会自己发现哪里不行
L4	自进化	系统自己生成改进候选并过门放行	效果算不到客户头上
L5	有结算	效果算到留资与成交，并能按引擎区分	未归因部分仍占一定比例

五级是递进的。跳过一级的做法在实践中都失败过：没有固定标准就建资产库，库里会堆满互相矛盾的内容；没有回退机制就上自进化，一次改错要人工全量回滚。

9.2 自评九问

本院准备了九个问题，按顺序回答，落在哪一个“是”上，就大致在哪一级。

1. 我们每个月有没有稳定对外发布内容。
2. 我们有没有一批固定的问法，用来定期检查 AI 的回答。
3. 这套问法的判断标准有没有固定的版本号，改动有记录。
4. 我们产出的句子、模板、判断标准，有没有存在一个可被调用的地方。
5. 第二个月的内容里，有没有一部分是复用上一轮的成果。
6. 系统会不会自己提出“这里可以改”的建议。
7. 系统提出的建议里，有没有一条真的被采用并投用。
8. 我们有没有退回过一次自动改动。
9. 我们的报告里，有没有出现留资、到访与成交的数字。

第 1 到 3 题全为“是”，说明处在上半段。第 4 到 5 题全为“是”，到 L3。第 6 到 8 题全为“是”，到 L4。第 9 题为“是”，到 L5。

本院建议先答第 8 题。如果系统已经在自动改动，却一次没有回退过，通常不是因为它没出错，而是因为回退路径没准备好。

9.3 三道坎

从自评结果看，团队卡住的位置很集中，一共三道坎。

第一道，从 L2 到 L3：愿不愿意把产出交回库里。 这道坎是人的问题。写手担心自己的东西被替代，所以宁愿每次重写。本院的看法是：交回库里的是句子与判断标准，不是人的位置。库越厚，写的人越省力。

第二道，从 L3 到 L4：敢不敢让系统改。 这道坎是风险承受力的问题。要过这道坎，前提是第 4.8 节的留痕与回退先做好。本院不建议在没有回退能力的情况下上自进化。

第三道，从 L4 到 L5：数据通道能不能接通。 这道坎是工程与合规的问题。客户的数据能不能接进来、授权怎么拿、号码怎么脱敏，都要事先谈清楚。很多客户卡在这里，不是技术做不到，是内部没人拍板。

三道坎里，第三道最容易被低估，问题出在部门协作上：本院见过团队把内容与测量都做得很整，唯独结算一直空着，原因是销售部门不愿意共享数据。这种情况下，前四级做得再好，也无法回答“值不值”。

9.4 每一级最常见的自欺

本院做过的自评里，团队给出的等级往往比实际高一级。原因集中在几种固定的说法上，列出来供读者自查。

L1 常见的自欺：把发布量当成绩。 一个月发了三十篇，就说自己有内容能力。发布量只说明动作在发生，不说明有人看。

L2 常见的自欺：把临时抽查当测量。 想起来查一次 AI 回答，看到品牌在，就说测量在跑。真正的测量要有固定问法、固定标准版本、固定周期。

L3 常见的自欺：把文件柜当资产库。 存了一堆文档，但生产环节从不调用，复用率是零。存起来不等于用得上。

L4 常见的自欺：把自动生成当自进化。 系统每天生成一批内容，但没有判定门、没有使用记录、没有回退。这不是进化，这是批量生产。

L5 常见的自欺：把能看到的当全部。 只报告可归因部分的表现，不提未归因占比。这样做的报表很好看，却会让人误以为全部效果都算清了。

五种自欺的共同点，都是用一个容易达成的动作替代一个难达成的机制。本院建议在自评时对每一级问一句：这个动作如果不做，会有什么后果。答不出后果的，多半还没达到这一级。

9.5 从自评到行动

自评出等级之后，接下来的一步很具体：找出当前等级里最缺的那一个动作。

本院按等级给出建议的第一动作，一次只做一件。

当前等级	建议的第一动作	做到什么算完成
L1	定一批固定问法	问法表写下来，不再临时想
L2	给判断标准加版本号	改动有记录，能说清这一版改了什么
L3	定一个复用比例目标	下一轮生产里有一部分来自复用
L4	走通一次回退	退回去一次，业务没有中断
L5	把未归因占比放上报告	未归因部分单独一行，不摊派

一次只做一件的理由是：这几件事互相有依赖。标准没锁就定复用比例，复用会变成抄旧稿；回退没走通就上自进化，出错时接不住。

本院建议把选定的第一动作写进项目记录，并约定一个检查时间。没有检查时间的动作，通常会被日常事务挤掉。

读完这一章，建议先给自己定一个等级，而不是定一个目标等级。第 10 章给出从当前等级往上走的节奏。

实施节奏：先做什么，什么时候看到什么

10.1 本院不给工期承诺

先把这一点说清楚。AI 引擎的更新节奏不由任何机构控制，客户的内部数据接入也不由我们控制。给出“三十天见效”这类承诺，对双方都没有好处。

本院给的是四个阶段，每个阶段有一个可观察的完成标志。标志达到了才进入下一阶段。标志没达到时继续做当前阶段，不往下推。

10.2 第一阶段：补入口与定标准

这一阶段做两件事：把承接入口补齐，把判断标准固定成第一版。

完成标志有两个，都可以自己验证：

- 同一批固定问法，在相隔一段时间后重跑两次，结论一致。不一致说明标准还没固定住。
- 品牌在 AI 回答里出现时，能查到它引用的是哪个来源。

这一阶段最容易犯的错是先做内容。本院的建议是相反的：标准没定，内容做得越多越乱，因为不知道什么叫好。

10.3 第二阶段：把产出交回库里

这一阶段开始建立资产库，把句子、模板、判断标准、禁用表述存下来，并在下一轮生产里调用。

完成标志是：第二轮内容里，有一批约定比例的成果来自复用，而不是重新撰写。这个比例本院不设行业标准，由客户与本院在开始前约定，比如“至少三成”。

这一阶段的难点不在技术，在协作。写的人要愿意把东西交出来。本院建议把复用率公开在看板上，让省下来的时间被看见。

10.4 第三阶段：让系统改自己

这一阶段打开自进化机制，让系统从执行记录里生成改进候选，过判定门后投用。

完成标志有三个：

- 有候选成功过门并投入实际使用。
- 有候选被判定门挡回，并且挡回的理由可查。
- 完成过至少一次回退，回退后业务没有中断。

第三个标志是本院单独要求的。一次都没回退过，通常意味着两种情况：要么系统太保守，改动没有实质作用；要么回退路径没走通过，真出问题时接不住。两种情况都要在本阶段解决。

10.5 第四阶段：打开结算

这一阶段把私域数据接进来，打开留资与结算。

完成标志是：报告里出现留资质量分与回本进度，并且把未能归因到具体引擎的线索单列出来。

如果数据通道一直谈不下来，本院的建议是不硬上，处理方式如下：可以在报告里保留“数据不足”的状态，同时继续做前三个阶段。结算缺失会让优化的方向感变差，但不会让前面三个阶段白做。

10.6 四个关口的检查动作

第 5.5 节提到的四个关口，在每个阶段都有对应的检查动作。

关口	第一阶段	第二阶段	第三阶段	第四阶段
采集关	检查来源可追溯	检查来源库是否复用	检查异常是否被记录	检查数据授权范围
选题关	检查是否有真实问法支撑	检查是否命中已有资产	检查是否采纳了改进候选	检查选题是否靠近留资方向
放行关	检查表述是否合规	检查模板是否被使用	检查自动改动是否经过人确认	检查落地页参数是否完整
结算关	检查标准版本号	检查复用比例	检查通过率与回退记录	检查未归因部分占比

这张表可以当作交付验收的清单。本院建议每个阶段结束时逐行核对，核对结果写进验收记录。

10.7 什么时候该看到什么

最后说时间预期。本院按“先变、后变”的顺序描述，不给天数。

先变的是来源构成。 内容开始被新的渠道引用时，来源构成会先动。它动得快，也容易反复，看到变化时不要急着下结论。

随后变的是提及与前三提及。 内容真实覆盖了更多问法之后，品牌在回答里出现的稳定性会提升。

再往后才是留资。 从被引用到有人留电话，中间有客户的决策周期。

最后是成交。 这一段的滞后又比前一段更长，且受价格与案场因素影响。

本院的建议是把这四个阶段的变化分别挂上不同的观察周期：来源构成看周，提及看月，留资与成交看季度。混在一起看，一定会得出错误结论。如果连续几个周期都看不到来源构成变化，下一步应当回到第一阶段检查入口与标准，而不是加大内容投放量。

10.8 谁参与

四个阶段涉及的角色不同。本院建议在开始前把人定下来，避免中途换人导致上下文丢失。

阶段	必须参与的人	负责什么
第一阶段	市场负责人	确认判断标准与固定问法
第二阶段	内容负责人	决定哪些产出交回资产库
第三阶段	技术或产品负责人	确认回退路径可用
第四阶段	销售负责人	批准数据接入与授权范围

四类人里，最难协调的通常是销售负责人。数据接入要动他们的台账，他们会担心被拿来考核。本院建议在第四阶段开始前先把数据用途写清楚：用于结算与归因，不用于考核个人。

还有一个角色需要提前定：谁有权确认对外内容。第 4.6 节的表里写明这一步由人做，如果一直没定下来，放行关会卡住，整个节奏停在这里。本院建议在第一阶段就把这个人定下来，并写进项目记录。

适用边界、合规底线与数据纪律

11.1 四种不适用的情形

不适用说在前面，可以省掉双方很多时间。下面四种情形，本院建议不要采用本文的方法。

第一种，业务只做一次性投放。 如果内容做完就结束、不打算看后续，那资产库与结算层都用不上，剩下的部分也不需要这套机制。

第二种，没有人能确认对外内容。 第 4 章的表里写明，对外内容的最终编辑由人确认。如果客户内部找不到这个人，流程会卡在放行关。

第三种，不接受数据接入。 私域数据不接进来，结算永远算不出。前三个阶段的成果仍然有效，但“值不值”这个问题会一直空着。

第四种，要求承诺排名或成交额。 本院不做这类承诺，也不会把承诺换一种说法写进合同。

11.2 本院坚持的四条立场

第一条，不做隐藏信息。 不把文字藏在图片里、隐藏层里，或者用与背景同色的方式呈现。这样做的内容对用户无意义，也容易被平台判定为作弊。

第二条，不做批量同质内容。 第 5.7 节已经说明，互相相似的内容会互相稀释。本院不接受以“铺量”为主要手段的方案。

第三条，不伪造溯源。 不虚构来源、不假冒第三方身份、不伪造引用。引用他人观点时标注出处。

第四条，不承诺排名。 AI 回答由引擎生成，任何机构都无法保证位置。有同行以“进入前三”为卖点，请要求对方出具可复算的过程记录。

11.3 与 T/CAPT 026—2026 的对应

有一份国家标准值得客户了解。《生成式引擎优化（GEO）可信信息传播与信息生态治理规范》，编号 T/CAPT 026—2026，由新华网牵头，三十余家机构参与起草，2026 年 8 月 11 日发布并实施。

这里需要说明一件事：**智多研究院不是该标准的起草单位。** 本院见过把参与起草的名头写得含糊的做法，本院不做这种事。下面的对照是本院做法与该标准原则方向的一致性说明，不构成任何背书关系。

标准关注方向	本院的对应做法	本文位置
区分正当优化与作弊手法	四条立场，含隐藏信息与批量同质内容两条禁止	11.2
来源可追溯	每个对外数字标来源类别；引用他人观点标出处	0.3、附录 D
分级对待不同来源	数字分实跑值、估算值、公开值三类，不混列	0.3
内容不误导	不做绝对化表述、不做价格与升值承诺	11.4
数据使用合规	授权、最小必要、脱敏、可核对四条安全线	6.7

客户在选服务商时，可以拿这张表去核对对方。本院建议重点问两件事：数字的来源能不能复算，作弊手法有没有明确禁止清单。

11.4 房地产行业的五条红线

房地产是本院服务的主要行业，表述上的限制比一般行业更严。下面五条写进本院的作业规范，触线的内容在放行关会被拦下。

1. 不使用「最」「第一」「唯一」等绝对化表述。
2. 不做价格承诺，不做升值与投资回报承诺。
3. 不虚构学区、配套、交付时间与户型数据。
4. 不使用未取得同意的客户信息。
5. 不使用未经授权的真人形象与案例背书。

这五条之所以写成“绝不”而不是“尽量避免”，是因为它们对应的都是明确的违规风险，不是程度问题。第12.4节的作业红线与此处的五条配套使用。

11.5 数据纪律

本文第2.6节记录了本院自己的一次测量事故。为了让同类问题不再发生，本院在对外交付里固定三个动作。

动作一，标来源。 每个对外数字都要标出属于哪一类：实跑值、估算值还是公开值。

动作二，写假设。 凡属推算，写明推算的假设是什么。假设站不住的时候，数字作废，不做修饰。

动作三，算不出来就不出。 一个数字如果回答不了“它是怎么算出来的”，就不进报告。宁可显示“数据不足”。

这三个动作会让报告看起来不那么漂亮。本院接受这个代价，因为报告的作用是支撑决策，不是让人安心。

11.6 比选服务商时的十个问题

客户在选服务商时，本院建议按这十个问题逐个问。不看回答的措辞，看对方能不能给出具体做法。

1. 你们报的每个数字，能不能当场演示怎么算出来的？
2. 判断好坏的标准有没有版本号，改过几次，怎么通知客户？
3. 有没有明确写出不做什么？请给我一份禁止清单。
4. 系统自动改动的内容出错时怎么退回？到目前为止退回过几次？
5. 客户数据怎么脱敏？明文落在哪里？
6. 报告里有没有“数据不足”这一状态？出现时怎么处理？
7. 效果能不能拆到具体引擎？拆不到的那部分怎么记？
8. 你们会不会承诺 AI 回答里的位置？如果会，依据是什么？
9. 内容生产里有没有批量同质的部分？占多少？
10. 如果连续两个周期没看到来源构成变化，你们的下一步动作是什么？

十个问题里，第3、4、7三个最能区分服务方。能给出禁止清单，说明有底线；能给出回退记录，说明自动化是真的；能把未归因单列，说明数据没有被修饰。

本院把这十个问题公开，也欢迎客户拿它们来问本院。回答不上来的问题，我们会说明原因，不用含糊表述盖过去。

11.7 开放性

本院把方法与指标全部写在这份文档里，包括判定门的判据、质量分的权重、封顶规则与不出分的情形。这样做有一个明确的用意：客户可以自行复算，也可以提出质疑。

如果读者发现本文的数字、算法与您的实践不符，欢迎指正。附录 D 列出了全部数字来源，附录 B 给出了可复现清单。本院会把有效的修正写进下一版。

11.8 四种不适用情形的处理方式

第 11.1 节列出了四种不适用情形。客户处在其中一种时，并不是没有办法，只是要换一种做法。

情形一，只做一次性投放。 建议不做资产库与结算层，只做感知与内容两部分，并且按项目结算，不按年度服务。

情形二，没有人确认对外内容。 建议先把这个人定下来，再启动项目。本院不接受由服务方代签对外内容，这一条在第 12.5 节已经写明。

情形三，不接受数据接入。 建议先做前三个阶段，报告里保留“数据不足”状态，等内部条件成熟再打开第四阶段。这个顺序不影响前三个阶段的成果。

情形四，要求承诺排名或成交额。 建议直接找其他方向的服务方。本院不做这类承诺，也不会用别的名目包装它。

本院把这些写出来，是为了减少双方在比价阶段的消耗。判断匹配与否，看这四种情形就够，不需要先比较报价。

第 12 章讲本院的作业红线与对客户的承诺。如果正在做服务商比选，建议把第 12.4 节当作核对清单使用。

智多研究院的主张与作业红线

12.1 名字的含义

智多研究院是 JingYu Media 设立的方法研究机构，负责智多GEO引擎的方法论研究、指数与标准建设、以及对外交付材料的编写。

“智多”取的是“以智取胜、以多为备”的意思：判断要靠智，覆盖要靠多。研究院的定位是做方法，不做销售承诺。因此本院对外材料与商务材料是分开的，本文属于前者。

12.2 使命与愿景

使命： 让企业能算清自己在 AI 回答里的位置与收益。

愿景： 把 GEO 从一门靠经验的手艺，变成一套可复算、可交付、可回退的工程。

这两句话解释了一个选择：本院把大量篇幅花在“怎么算”和“怎么退回”上，而不是花在效果形容上。手艺靠人，工程靠机制。机制能被验收，手艺不能。

12.3 四条价值观

第一条，数字先于故事。 一个说不清算法的好消息，价值低于一个能复算的坏消息。

第二条，边界先于承诺。 先说清不做什么，再谈能做什么。本文第 3.4 节与第 11.1 节都是这个用法。

第三条，留痕先于自动。 每一次自动改动都要留记录，并且要能退回。做不到就先不要自动化。

第四条，复用先于新增。 库里已有的东西先用起来，再考虑新建。这条对策的是资产膨胀。

12.4 七条作业红线

本院把它写成红线而不是建议，因为触线的代价由客户承担。

序	红线	触线后果
1	不承诺 AI 回答里的排名或位置	终止合作
2	不使用绝对化用语与虚假表述	内容不予放行
3	不伪造来源、不假冒第三方身份	终止合作
4	不把内容藏在用户看不见的地方	内容不予放行
5	不用批量同质内容充数	方案退回重做
6	判断标准未定不启动测量	项目不进入测量阶段
7	推算值不写成实测值	报告退回重写

第 6 条与第 7 条是本院在实践中补上的，各来自一次教训：第 6 条来自标准没定就开始跑，跑出来的数字两次对不上，事后只能全部作废。第 7 条来自本文第 2.6 节记录的那次事故。

12.5 对客户四项承诺

与红线对应，本院给出四项可检验的承诺。

- 承诺一，每个对外数字都能给出算法与来源。** 客户可以要求对任意一个数字做复算演示。
- 承诺二，每一代能力资产都能回退。** 回退操作不需要本院介入，客户自己在系统里可以完成。
- 承诺三，报告显示“数据不足”时不虚构补充。** 本院不用“参考行业均值”之类的做法填补空缺。
- 承诺四，系统不代替客户对外发布。** 对外发布动作的最后一步由客户确认，本院与系统都不代签。

这四项承诺都可以写进合同，用法是：本院建议客户在签约前逐条确认，并在验收时逐条核对。如果某一条在实际执行中被绕过，请按附录 H 的联系方式反馈，本院会把处理结果写进下一版。

12.6 研究院不做什么

边界是方法的一部分。本院公开列出四件不做事。

- 不做纯内容外包。** 只写内容、不建资产库的项目，本院不接。这类项目做完，等于客户买了一批会过期的稿子。
- 不做无法复算的报告。** 数据来源说不清的项目不接，包括客户自己提供的、无法核实的行业数据。
- 不做承诺排名的合作。** 无论以什么名目包装，本院不签含位置承诺的条款。
- 不做数据转手。** 客户接入的数据只用于本项目结算，不用于其他客户，不对外发布。

本院把这些写出来，是为了让客户在接触初期就能判断是否匹配。四件事里有任何一件是客户的硬需求，本院建议直接找其他方向的服务方，不必进入比价流程。

12.7 研究院与客户的分工

最后把分工写清楚，避免执行中互相等待。

事项	研究院	客户
判断标准与问法库	提供与维护	确认
内容生产	承担主要部分	提供业务事实
对外内容确认	不代签	确认并署名
资产库维护	提供工具与规则	使用与反馈
数据接入	提供通道与脱敏	授权与配合
结算结论	出报告并标注不确定部分	用结论做决策

这张表里有两行需要客户留意。一行是“对外内容确认”，责任在客户，本院只做准备与检查。另一行是“结算结论”，本院保证算法透明、不确定部分如实标注，结论怎么用由客户决定。

分工清楚之后，双方各自的动作都能被检查。本院建议把这张表附在合同后面，作为服务范围的说明。

第 13 章是本文的收束。

结语

13.1 回到最初那个问题

第 2 章开头有一个没人答得上来的问题：这条 AI 提到，值多少钱。

本文给出的答案不是一句估价，而是一条路径。这条路径有四个站点：

站点	要做的事	做到之后拿到什么
一	把入口补齐，让看到答案的人有地方可去	说清楚人是从哪个引擎来的
二	把产出存下来，让下一次不用重新开始	一次投入，多次使用
三	让系统自己发现哪里不行，并且改掉	一份会自己变强的资产
四	把效果算到客户头上，分出哪一部分来自 AI 回答	一个能回答“值多少钱”的数字

四个站点走完，最初那个问题才有答案。只走前两个，得到的是效率和更漂亮的报告。走到第三个，得到的是会自己变强的系统。走到第四个，才谈得上算账。

本院在这份文档里反复做的一件事是划边界：说清楚哪些已经能跑、哪些还在建，说清楚哪些数字是实测、哪些是推算，说清楚哪些情形下我们不出分。这些边界看起来是在削弱主张，本院认为恰恰相反。一套能被验收的方法，才值得被长期使用。

13.2 更新策略

本院按年度做一次修订。出现下列情形之一时提前修订：

- 相关的国家或行业标准发布、修订或废止。
- AI 引擎的引用机制出现明显变化，导致本文的测量方法失效。
- 客户在实践中发现了本文的判据与实际不符。
- 本院自己的交付流程有实质性改动。

修订只改必要部分，不改的部分保留原文与编号，避免读者手里的旧版本失去参照。

13.3 本文没有做完的部分

诚实起见，本院列出这一版里明确还没做完的部分。

- 引擎跳转归因的三项改动仍在分批上线。** 第 7.5 节列出的三项，本文写作时还没有在全量客户上投用。文中把它们写成在建，不是写成本已具备。
- 到访比例的参考区间还没有公开数值。** 第 6.4 节的对照组分析目前只给出方法，没有给出跨客户的参考区间。样本量还不够支撑一个可对外引用的数值。
- 行业专版与 OEM 版尚未编写。** 两版会按本文框架实例化，判据不变。
- 这三项会在后续版本里补上。本院不会为了让文档看起来完整而提前写出结论。

13.4 致读者

如果你是企业方，建议从第 9 章的自评九问开始，先确定自己现在的等级，再挑一条最贴近业务的路径往前走。本院不建议一次做完四个阶段。

如果你是同行，欢迎对本文的判据提出质疑，包括「凭什么这样定」这一类问题。第 11.6 节说明了本院的方法与指标全部公开，可以自行复算。

如果你是研究者，附录 B 给出了可复现清单。

本文的下一步是行业专版与 OEM 版，会按本文的框架实例化，不改变判据。反馈与勘误请通过附录 H 提交，本院会在下一版中回应。

附录与术语说明

附录 A 术语表

正文里带「」的术语在此集中解释。表中的释义是本院固定下来的用法，与其他机构的同名提法不构成沿用关系。

术语	释义与使用场景
闭环	本文的核心提法。指系统把执行结果收回来、自己生成改进、过门后投用的机制。不是“流程走完一圈”的意思。
感知	三个部件之一。负责知道 AI 在各家引擎里怎么提我们、提得准不准。
资产	三个部件之一。指下次还能直接用的东西。
进化	三个部件之一。指系统对自身能力资产的自动改进。
结算	衡量效果的那一层。对象是留资、到访与成交。
能力组件	资产库的组成单元，共 102 个，分角色、模板、技能、判定标准、禁令五类。
问法库	2,420 条真实用户会问出口的完整问题。不是关键词。
覆盖网格	22 个决策角度乘 6 类购房人设乘 5 个决策阶段，共 660 格。用于衡量覆盖面。
执行轨迹	系统对一次执行过程的记录：走到哪一步、结果如何、卡在哪里。
改进候选	由执行轨迹生成、尚未过门的资产草稿。
判定门	候选能否进入正式资产库的四道判断，见第 4.3 节。
实测适应度	用真实使用记录算出的一条资产的可用程度，见第 4.4 节。
留痕	每一代资产的来源记录，用于追溯与排查。
回退	退回上一代资产的动作，由人确认。
标准版本	判断好坏的那套规则的版本号，见第 4.7 节。
触达基线	按引擎统计的内容触达次数。是先行指标，不等于客户量。
入口标记	分发时贴在入口链接上的短标识，用于区分内容来源，见第 7.2 节。
未归因	无法判断来自哪个引擎或渠道的线索。单列，不按比例摊派。
线索	漏斗第一段。留下可联系号码并通过有效性检查。
到访	漏斗第二段。客户实际到场，或完成线上深聊。
成交	漏斗第三段。签署认购或合同，金额写入记录。
留资质量分	由号码有效性、意向结构、线索到访率、到访成交率四部分构成的分，见第 6.5 节。
回本进度	成交金额与预设投入目标的比值。
区域决策切片	本院对楼盘相关数据规模的对外说法。按城市加板块切片，不披露精确条数。

附录 B 可复现清单

本附录给出本文关键结论的复现方式，供客户与研究者核对。

B.1 对外数字

四个对外数字分别为问法库 2,420 条、覆盖网格 660 格、能力组件 102 个，以及 22 个决策角度。它们同时出现在官网、产品界面与本文，四处数值一致。任一处不符即为我们的失误。

B.2 能力组件的构成

102 个组件分五类：角色、模板、技能、判定标准、禁令。五类相加等于 102。改动任何一类都要同步改动总数。

B.3 判定门的四道判断

结构、语义、重复、实战。前三道可写成规则，第四道依赖使用记录。判据见第 4.3 节，放行线与留任线的不对称见第 4.5 节。

B.4 小样本打折

小样本打折的取值如下：一次成功按 0.21 分计，三次全成功按 0.44 分计，十次全成功按 0.72 分计。这是成功比例的下置信界算法在本文所用的置信水平下的取值。

B.5 留资质量分的构成

号码有效性 0.30、意向等级结构 0.25、线索到访率 0.25、到访成交率 0.20。权重固定，不按客户调整。

B.6 到访归因的封顶

客户自述来源的覆盖比例低于约定线时，到访归因分最高按 70 分计，并在报告中注明“相关性而非因果”。覆盖比例过线时，按 AI 回答组与其他组的到访比例差值计算。

附录 C 合规与红线速查

类别	内容	正文位置
四条立场	不做隐藏信息、不做批量同质内容、不伪造溯源、不承诺排名	11.2
行业五条	无绝对化用语、无价格与升值承诺、不虚构配套、不用未授权客户信息、不用未授权形象	11.4
七条作业红线	见正文表格	12.4
数据三条安全线	授权、最小必要、脱敏，另加可核对	6.7
标准对应	T/CAPT 026—2026 五个方向的对应做法	11.3

附录 D 数据来源分类

本文出现的数字按下列三类归属。读者转述前请核对本表。

实跑值

数字	说明
2,420 条	问法库总条数
22	购房决策角度数
660 格	覆盖网格数

数字	说明
102 个	能力组件总数
264 条	第 2.6 节记录的历史事故涉及的批次条数，为该批次的实际记录数
0.35	实测适应度的地板值
70 分	无有效对照组时的到访归因封顶值
0.30 / 0.25 / 0.25 / 0.20	留资质量分四项权重
0.21 / 0.44 / 0.72	小样本打折在 n=1、n=3、n=10 下的取值，由 B.4 的方法算出
5 类 / 9 项	能力组件类别数与看板指标数

估算值

数字	假设
到访比例的相关区间	客户自述来源覆盖比例过线的前提下推演，未对外给出具体数值

公开值

数字	出处
引用来源构成在周与周之间的变化幅度	第三方公开统计
与绝对化用语、房地产广告相关的法定限制	《中华人民共和国广告法》及配套规定

附录 E 常见问题

- 1. 你们保证 AI 回答里排名靠前吗？** 不保证。AI 回答由引擎生成，任何机构都无法保证位置。我们能保证的是过程可复算、改动可回退。
- 2. 为什么报告有时显示“数据不足”？** 三种情形：数据授权没开、有效线索数太少、自述来源覆盖比例过低。少给一个分数比给一个会误导的分数好。
- 3. 没有对照组时，到访归因怎么算？** 按 70 分封顶，并注明这是相关性不是因果。数据够的时候，用 AI 回答组与其他组的到访比例差值来算。
- 4. 引擎级归因现在准不准？** 还在分批上线。已经能跑的是渠道级归因、平台触达基线与客户自述来源。归不到具体引擎的线索单列，不摊派。
- 5. 客户号码存在系统里安全吗？** 号码进入系统即加密，明文不留在库里；姓名只保留姓氏；采集前需要客户开启授权，授权没覆盖号码时该字段直接丢弃。
- 6. 内容是不是用 AI 批量生成的？** 不是。第 5.7 节说明了批量同质内容会互相稀释。生产量按方向做控制，允许复用资产。
- 7. 复用率多少算合格？** 本院不设行业标准，由客户与本院在开始前约定，例如至少三成。关键的是一次比一次高。
- 8. 系统会不会自己改我写好的文案？** 不会。机器只改机器产出。人工撰写的文案与人工确认过的表述，系统只给建议，不改动。

- 9. 怎么知道标准被改过？** 标准带版本号，升版需要人确认，历史数据标注它属于哪一版。对比两个周期的数字前，先确认它们是同一版算出来的。
- 10. 一次回退都没有，说明系统很稳吗？** 通常不是。连续零回退往往意味着两种情况：改动没有实质作用，或者回退路径没走通过。
- 11. 看板为什么不放“累计生成多少字”？** 总量型指标只增不减，读不出好坏。本院建议新增指标前先回答：它变差时我们要改什么。
- 12. 我们已经有内容团队，会冲突吗？** 不冲突。资产库交出的是句子与判断标准，不是人的位置。库越厚，写的人越省力。

附录 F 关于智多研究院

项目	内容
机构名称	智多研究院
出版主体	JingYu Media
官网	ZhiduoGEO.com
品牌	智多GEO引擎（正式全名） / 智多引擎（主用简称） / 智多智能GEO（平台名，简称智能GEO）
研究范围	生成式引擎可见性方法论、指数与标准建设、交付材料编写
本版文档	闭环型 GEO：Agent 自进化、内容资产与留资结算方法论，v2.0，2026-09-21

附录 G 常见质疑与回应

- 本附录收录本院在不同场合被问到的质疑，以及我们的回应。这些质疑大多有道理，本院不把它们说成误解。
- 质疑一：所谓自进化，是不是就是用大模型自动生成内容？**
- 不是。区别在于判定门与使用记录。没有门的自动生成只是产量，不是能力。第 4.3 节把两者的差别写清楚了。
- 质疑二：小样本打折会不会让系统太保守，好用的能力上不来？**
- 答案是会的，代价是本院主动接受的：早期偏保守，换来的是每一条入库能力的分数都能被解释。客户如果愿意承担更大的风险，可以在约定范围内放松放行线，但放松这个动作本身要留记录。
- 质疑三：没有对照组就说相关，那到访归因这个指标还有什么用？**
- 它的用途是排序，不是归因。它能回答“哪个方向更值得加投入”，不能回答“是这一条内容带来的”。本院在报告里会把这一层限制写出来。
- 质疑四：引擎跳转归因还没全量投用，为什么写进白皮书？**
- 因为它是一项已定义、可验收的机制。写出来，客户才能按它设计验收方式。把在建项写成本已具备，是不诚实的做法。
- 质疑五：把内容交给资产库，会不会让内容越来越像？**
- 存在这个风险。本院的对策有两个：资产库以判断标准与句子为主，不以成稿为主；保留人工撰写与人工确认的环节，不追求全自动生成。

质疑六：你们自己出过测量事故，凭什么谈数据纪律？

正因为出过。第 2.6 节把那件事的经过写出来了，它也是第 12.4 节两条红线的来由。本院不宣称自己没出过错，只宣称每个对外数字都能被复算。

附录 H 反馈方式

读者如果发现本文的数字、判据或引用有误，欢迎通过官网 ZhiduoGEO.com 联系我们。反馈时如果能附上具体位置与理由，会更容易被处理。

本院对每一条有效反馈的处理方式是：确认问题、给出结论、把结论写进下一版。已经发布的版本不追溯修改，改动在下一版体现，避免读者手里的旧版本失去参照。

需要数据来源说明、复算演示或转载授权，同样通过官网联系。本院建议在引用本文数字之前先核对附录 D，在比对本文判据与自身实践之前先阅读附录 B。